

UNIVERSITY OF NAPLES FEDERICO II

DEPARTMENT OF PHARMACY



PhD in Pharmaceutical Science – XXXI cycle

***Investigating non-canonical nucleic acid
structures and their interaction with potential
anticancer drugs***

PhD Thesis

Alessia Pagano

Tutor:
Prof. BRUNO PAGANO

PhD Coordinator:
Prof. MARIA VALERIA D'AURIA

ABSTRACT

The outstanding structural polymorphism of DNA allows for the formation of non-canonical secondary structures, such as G-quadruplexes (G4s), G-triplexes (G3s) and i-motifs. G4 and i-motif structures are found within important functional genomic regions such as telomeres and gene promoters. In particular, they are localized within the promoter regions of several proto-oncogenes, whose overexpression leads to malignant transformation, where they play a major role in the regulation of transcription. In many cases, the biological consequence of non-canonical secondary structure formation in the promoter element is gene silencing. For this reason, the search for ligands able to bind and stabilize G4s and/or i-motifs is pharmacologically very important to develop new anticancer strategies.

In this PhD thesis, a series of studies have been carried out with the aim of investigating these non-canonical nucleic acid structures and their interaction with potential anticancer drugs. Such studies have led to the discovery of new and selective G4 ligands (Chapter 3). Chapter 4 deals with the tandem application of virtual screening along with experimental investigations, that led to discover the first dual G-triplex/G-quadruplex stabilizing compound. In Chapter 5, biophysical techniques have been employed to demonstrate that some well-known G4 ligands are also able to interact with i-motif structure.

The last section of this PhD thesis deals with a study conducted in Dr. Vincenzo Abbate's laboratory at King's College London (UK). It concerns the design and synthesis of a new class of gallium chelator to be employed in the development of chemically-modified nucleic acid aptamers to be used as theranostics.

LIST OF PUBLICATIONS

Paper I

“Toward the Development of Specific G-Quadruplex Binders: Synthesis, Biophysical, and Biological Studies of New Hydrazone Derivatives”. J. Amato, R. Morigi, B. Pagano, A. Pagano, S. Ohnmacht, A. De Magis, Y. Tiang, G. Capranico, A. Locatelli, A. Graziadio, A. Leoni, M. Rambaldi, E. Novellino, S. Neidle, and A. Randazzo. *J. Med. Chem.*, **2016**, 59, 5706-5720. doi: 10.1021/acs.jmedchem.6b00129.

Paper II

“Targeting the *BCL2* Gene Promoter G-Quadruplex with a New Class of Furopyridazinone-Based Molecules”. J. Amato, A. Pagano, D. Capasso, S. Di Gaetano, M. Giustiniano, E. Novellino, A. Randazzo, and B. Pagano. *ChemMedChem*, **2018**, 13, 406-410. doi: 10.1002/cmdc.201700749.

Paper III

“Tailoring a lead-like compound targeting multiple G-quadruplex structures”. J. Amato, C. Platella, S. Iachettini, P. Zizza, D. Musumeci, S. Cosconati, A. Pagano, E. Novellino, A. Biroccio, A. Randazzo, B. Pagano, and D. Montesarchio. *Eur. J. Med. Chem.* **2019**, 163, 295-306. doi: 10.1016/j.ejmech.2018.11.058.

Paper IV

“Discovery of the first dual G-triplex/G-quadruplex stabilizing compound: a new opportunity in the targeting of G-rich DNA structures?”. J. Amato, A. Pagano, S. Cosconati, G. Amendola, I. Fotticchia, N. Iaccarino, J. Marinello, A. De Magis, G. Capranico, E. Novellino, B. Pagano, A. Randazzo. *Biochimica et Biophysica Acta – General Subjects*, **2017**, 1861, 1271-1280. doi: 10.1016/j.bbagen.2016.11.008.

Paper V

“Common G-Quadruplex Binding Agents Found to Interact With i-Motif-Forming DNA: Unexpected Multi-Target-Directed Compounds”. A. Pagano, N. Iaccarino, M. A. S. Abdelhamid, D. Brancaccio, E. U. Garzarella, A. Di Porzio, E. Novellino, Z. A. E. Waller, B., J. Amato, and A. Randazzo. *Frontiers in Chemistry*, **2018**, 6, 281. doi: 10.3389/fchem.2018.00281.

Other publications

1. “Characterization of monovarietal extra virgin olive oils from the province of Béjaïa (Algeria)”. F. Lainer, N. Iaccarino, J. Amato, B. Pagano, A. Pagano, G. Tenore, A. Tamendjari, P. Rovellini, S. Venturini, G. Bellan, A. Ritieni, L. Mannina, E. Novellino, and A. Randazzo. *Food Research International*, **2016**, 89, 1123-33. doi: 10.1016/j.foodres.2016.04.024.
2. “¹H NMR-based metabolomics study on follicular fluid from patients with Polycystic Ovarian Syndrome (PCOS)”. N. Iaccarino, J. Amato, B. Pagano, A. Pagano, L. D’Orlando, C. Alviggi, A. Randazzo. *Biochimica Clinica*, **2018**, 42, 26-31. doi: 10.19186/BC_2018.008.

LIST OF ABBREVIATIONS

A	Adenine
AMD	Age-related macular degeneration
BCL2	B-cell lymphoma-2
BJ	Human fibroblasts
BJ-hTERT	Human immortalized BJ fibroblasts
C	Cytosine
CD	Circular dichroism
CPG	Controlled pore glass
CTCF	Corrected total cell fluorescence
DAPI	4',6-diamidino-2-phenylindole
DDR	DNA damage response
DEAD	Diethyl azodicarboxylate
DIPEA	<i>N,N</i> -Diisopropylethylamine
DMEM	Dulbecco's modified eagle medium
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DMT	4,4'-dimethoxytrityl
DNA	Deoxyribonucleic acid
DSC	Differential scanning calorimetry
FACS	Fluorescence activated cell sorting
FAM	6-carboxyfluorescein
FBS	Fetal bovine serum
FID	Fluorescent intercalator displacement
FP	Furo[2,3- <i>d</i>]pyridazin-4(5 <i>H</i>)-ones
FRET	Fluorescence resonance energy transfer
G	Guanine
G3	G-triplex
G4	G-quadruplex

GE	Gel electrophoresis
HATU	1-[Bis(dimethylamino)methylene]-1 <i>H</i> -1,2,3-triazolo[4,5- <i>b</i>]pyridinium 3-oxid hexafluorophosphate
HDF	Normal human dermal fibroblasts
HeLa	Human adenocarcinoma cancer cell line
HepG2	Human liver cancer cell line
HIF-1 α	Hypoxia-inducible factor 1-alpha
hTERT	Human telomerase reverse transcriptase
IF	Immunofluorescence
INS	Insulin gene
IR	Infrared
LCAA	Long chain alkylamine
MCF-7	Breast cancer cell line
MST	Microscale thermophoresis
MTT	Thiazolyl blue tetrazolium bromide
NHE	Nuclease hypersensitive elements
NMR	Nuclear magnetic resonance
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PDGFR	Platelet-derived growth factor receptors
PET	Positron emission tomography
PI	Propidium iodide
POT1	Protection of telomere 1
PSMA	Prostate-specific membrane antigen
PVDF	Polyvinylidene fluoride
qPCR	Quantitative polymerase chain reaction
RAP1	Repressor/activator protein 1
RNA	Ribonucleic acid
RT-PCR	Reverse transcriptase-polymerase chain reaction
SDS	Sodium dodecyl sulfate
T	Thymine
TAMRA	6-carboxytetramethylrhodamine
TBA	Thrombin binding aptamer
TCA	Trichloroacetic acid

TEA	Triethylamine
TFA	Trifluoroacetic acid
THP	Tris(hydroxypyridinone)
TIF	Telomere induced foci
TIN2	TRF1- and TRF2-interacting nuclear protein 2
TO	Thiazole orange
TPP1	Tripeptidyl peptidase 1
TRF1	Telomere repeat factor 1
TRF1	Telomeric repeat-binding factor 1
TRF2	Telomere repeat factor 2
TSS	Transcription start sites
U	Uracile
U2OS	Human bone osteosarcoma epithelial cell lines
VEGF	Vascular endothelial growth factor
VS	Virtual screening

TABLE OF CONTENTS

ABSTRACT	I
LIST OF PUBLICATIONS	II
LIST OF ABBREVIATIONS	IV
CHAPTER 1 INTRODUCTION	1
1.1 General description of nucleic acids.....	1
1.2 Non-canonical nucleic acid structures.....	3
1.2.1 G-quadruplexes.....	4
1.2.2 G-triplex	7
1.2.3 i-Motif	8
1.3 Biological relevance of G-quadruplexes.....	10
1.3.1 Targeting G4 structure in cancer cells.....	12
1.4 Biological relevance of i-motif	15
1.5 DNA sequences as therapeutic agents	16
CHAPTER 2 METHODS	19
2.1 Circular dichroism.....	19
2.2 Fluorescence Spectroscopy	21
2.3 Electrophoresis	23
2.4 Differential scanning calorimetry.....	24
2.5 Nuclear Magnetic Resonance.....	27
2.6 Microscale Thermophoresis.....	32
CHAPTER 3 TARGETING G-QUADRUPLEXES	35
3.1 New Hydrazone derivatives as specific G-quadruplex binders (<i>Paper I</i>)	37
3.1.1 Introduction.....	37
3.1.2 Results and discussion.....	38
3.1.3 Conclusions	55
3.1.4 Experimental Section	56
3.2 Targeting the <i>BCL2</i> Gene Promoter G-Quadruplex with a New Class of Fuopyridazinone-Based Molecules (<i>Paper II</i>).....	62
3.2.1 Introduction.....	62

3.2.2 Results and discussion.....	66
3.2.3 Conclusions	76
3.2.4 Experimental Section	76
3.3 Tailoring a lead-like compound targeting multiple G-quadruplex	81
3.3.1 Introduction	81
3.3.2 Results and discussion.....	82
3.3.3 Conclusions	101
3.3.4 Experimental section.....	102
CHAPTER 4 TARGETING G-TRIPLEX	109
4.1 Discovery of the first G-triplex/G-quadruplex stabilizing compound (<i>Paper IV</i>)	109
4.1.1 Introduction	109
4.1.2 Results and discussion.....	111
4.1.3 Conclusions	126
4.1.4 Experimental Section	126
CHAPTER 5 TARGETING I-MOTIF-FORMING DNA	133
5.1 Common G-quadruplex binding agents found to interact with i-motif-forming DNA (<i>Paper V</i>)	133
5.1.1 Introduction	133
5.1.2 Results	135
5.1.3 Discussion	148
5.1.4 Conclusions	151
5.1.5 Experimental Section	152
CHAPTER 6 SYNTHETIZING A NEW THP DERIVATIVE SUITABLE FOR THE CONJUGATION TO NUCLEIC ACID SEQUENCES	156
CONCLUSIONS	169
REFERENCES	170
APPENDIX	183

Chapter 1

INTRODUCTION

1.1 General description of nucleic acids

Nucleic acids are the biological macromolecules that play fundamental roles in cells and ensure the normal development and functioning of an organism. They are composed of nucleotides, which are monomers made of three components: a pentose sugar, a phosphate group and a nitrogenous base. If the sugar is a ribose, the polymer is ribonucleic acid (RNA); if the sugar is a deoxyribose, the polymer is deoxyribonucleic acid (DNA). The bases of DNA divide into two groups: purines [adenine (A) and guanine (G)] and pyrimidines [thymine (T) and cytosine (C)]. In RNA, thymine is replaced by uracil (U). Purines are made of six- and five-membered nitrogen-containing rings fused together; pyrimidines consist of only a six-membered nitrogen-containing ring (Voet and Voet, 1995) (Figure 1.1).

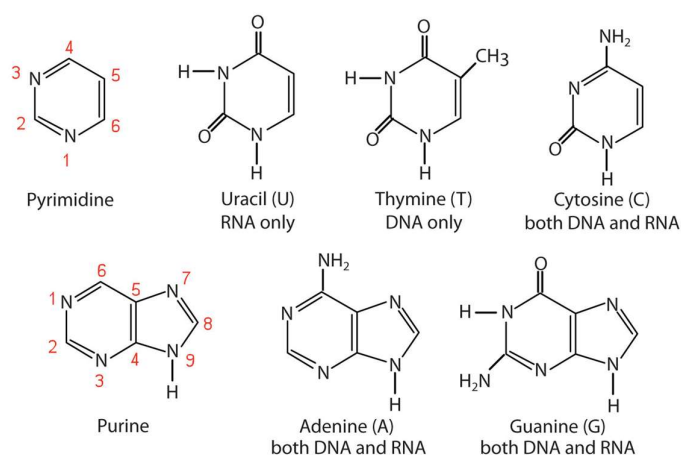


Figure 1.1. Structure of purine and pyrimidine bases.

A nucleoside is a molecule whose C1' of the pentose sugar is covalently bound by an N-glycosidic bond to the N9 or N1 of purines or pyrimidines, respectively. The rotations of this

glycosidic bond allows structural diversity to the DNA molecule. The glycosidic bond may be either *syn* or *anti* (Figure 1.2).

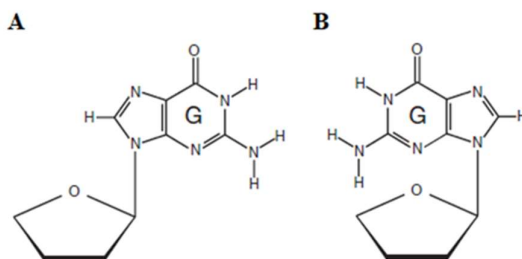


Figure 1.2. The glycosidic bond: A) *anti* conformation; B) *syn* conformation.

The *syn* conformation is formed when the C1'-O4' bond is *cis* to the N9-C4 purine bond, and *cis* to the N1-C2 of pyrimidine. This occurs when the bulk of the purine base faces the sugar or when the C2 carbonyl is on top of the sugar ring. The *anti* conformation is formed when the bulk of the base is rotated away from the sugar. This occurs when the C1'-O4' bond of the sugar is *trans* to the N9-C4 bond of the purine base. In pyrimidines, the *anti* conformation is formed when the C2 carbonyl faces away from the sugar. This occurs when the C1'-O4' bond is *trans* to the N1-C2 pyrimidine bond. The attachment of a phosphate group to the C5' carbon of the sugar converts a nucleoside into a nucleotide. A 5'-phosphate group of a nucleotide links to the 3'-hydroxyl group of the next nucleotide to form a *phosphodiester bond*. The connected nucleotides by *phosphodiester bonds* form single-stranded DNA.

The single-stranded DNA bases may pair non-covalent interactions, as first described for B-DNA by Watson and Crick back in 1953. (Watson and Crick, 1953). Adenine and thymine (A:T) base pairs are held together by two hydrogen bonds, whereas guanine and cytosine (G:C) base pairs are held by three hydrogen bonds (Figure 1.3).

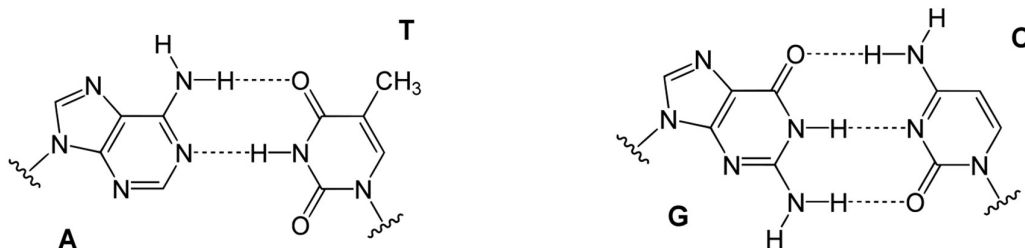


Figure 1.3. Watson and Crick base pairs.

Thus, all four base pairs fit neatly into the double helix. The two strands are antiparallel, with the 5'-end of one strand adjacent to the 3'-end of the other. The two strands coil around each other to form a right-handed double helix, with the base pairs in the center and the sugars and negatively charged phosphates forming the external hydrophilic backbone. The stability of the duplex derives from both base stacking and hydrogen bonding. The B-DNA has a wide major groove and a narrow minor groove running around the helix along the entire length of the molecule (Figure 1.4). Proteins interact with the DNA in these grooves (principally in the major groove), and some small drug molecules (e.g. netropsin and distamycin) bind in the minor groove.

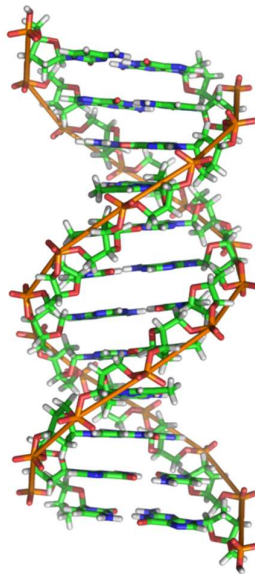


Figure 1.4. DNA fragment in the B-form.

1.2 Non-canonical nucleic acid structures

Under certain conditions, nucleic acids can adopt non-canonical conformations other than B-DNA. These unusual DNA structures can involve from one to four nucleic acid strands, that may arrange into hairpins, cruciform, parallel-stranded duplexes, triplexes (H-DNA), G-quadruplex (G4), i-motif and other non B-forms (Figure 1.5). It has been demonstrated that these structures are widely distributed throughout the human genome, and are enriched critical regions (Bacolla and Wells, 2009). In addition, their formation depends on the specific DNA sequence.

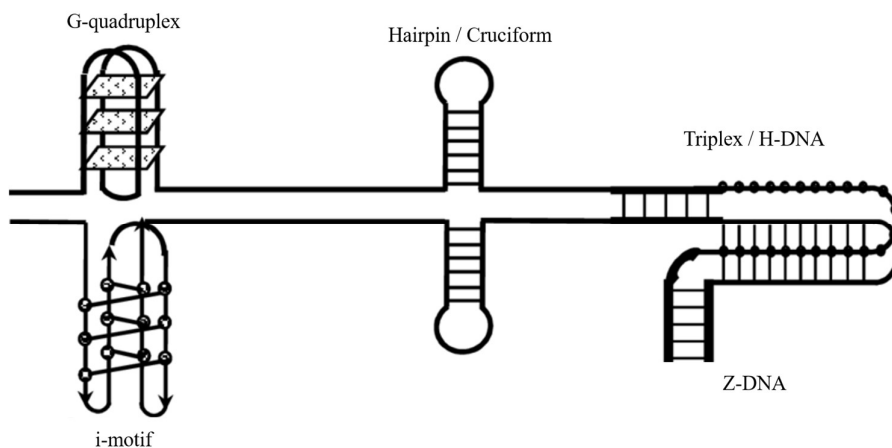


Figure 1.5. Schematic illustration showing some examples of non-canonical DNA structures.

1.2.1 G-quadruplexes

One of the most important non-canonical nucleic acid structures is the G-quadruplex (G4). It is a four-stranded structure formed by G-rich sequences, either DNA or RNA, which has at least two stacked G-tetrads (Majima and Tetsuro, 2011; Wang and Vasquez, 2006; Wells et al., 2005; Zhao et al., 2010). A G-tetrad is a planar square arrangement of four guanines held together by eight Hoogsteen hydrogen bonds (Huppert, 2010; Rhodes and Lipps, 2018). In it, the N7 and O6 of each guanine are hydrogen-bond acceptors from N2 and N1 of adjacent one, respectively. The π - π stacking between the G-tetrads further stabilizes this structure. The G-tetrads are linked by loops that can assume different conformations. Moreover, this arrangement delimits a central channel.

The formation and stability of G4s strongly depends on cations, with the central channel neutralizing the strong negative electrostatic potential from guanine O6 atoms (Figure 1.6) (Bochman et al., 2012; Burge et al., 2006; Williamson et al., 1989).

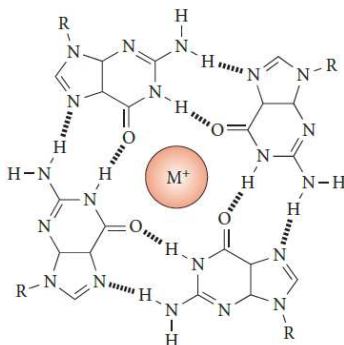


Figure 1.6. G-tetrad with monovalent cation [adapted from (Bochman et al., 2012)].

Cations that better stabilize the G4 structures are Na^+ and K^+ . The precise location of the cations depends on their nature: Na^+ ions are usually in-plane with the G-tetrads, while K^+ ions are always equidistant between two G-tetrads, forming a symmetric tetragonal bipyramidal configuration with the oxygen atoms. Switching between Na^+ and K^+ often induces a structural alteration of G4s, which indicates high conformational flexibility for these structures (Burge et al., 2006). From a biological point of view, the K^+ is far more biologically relevant due to its higher intracellular concentration (~ 140 mM) compared to Na^+ (5-15 mM).

The topology and stability of G4s depend on many factors: length and composition of the G4-forming sequence, strand stoichiometry and alignment, size of the loops, and nature of the binding cations.

G4s can be unimolecular, when a single strand folds back on itself; bi- or tetra-molecular, when formed by two or four strands, respectively (Figure 1.7) (Huppert, 2010).

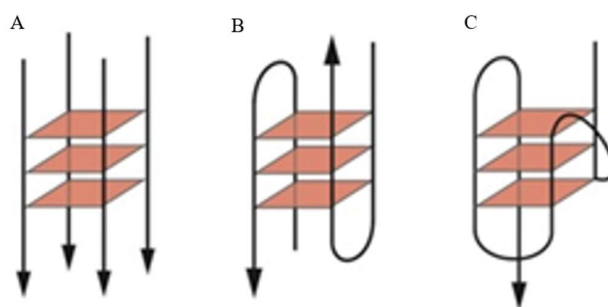


Figure 1.7. Schematic representation of A) tetrameric, B) dimeric and C) monomeric G4 structures.

G4s can be divided into four groups according to the relative strands orientation (Patel et al., 2007; Phan, 2010):

- Parallel-stranded core, in which four strands are oriented in the same direction (Figure 1.8A);
- Hybrid or (3+1) core, in which three strands are oriented in one direction and the fourth in the opposite direction (Figure 1.8B);
- Antiparallel-stranded core, in which two strands are oriented in one direction and the other two in the opposite one, with an *up-up-down-down* core (Figure 1.8C);
- Antiparallel-stranded core, in which two strands are oriented in one direction and the other two in the opposite one, with an *up-down-up-down* core (Figure 1.8D).

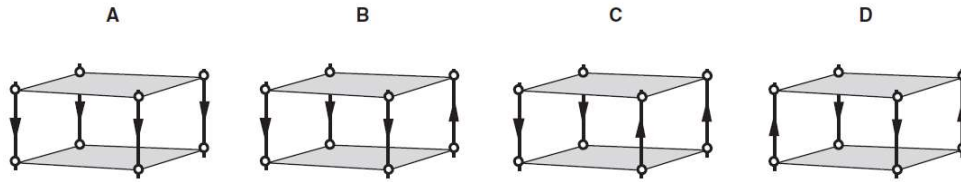


Figure 1.8. A) Parallel-stranded core; B) hybrid core; C) *up-up-down-down* core; D) *up-down-up-down* core [adapted from (Phan, 2010)].

Guanines in a G-tetrad may have two conformations: *anti* and *syn*. All parallel G4s have all the guanine glycosidic angles in the *anti* conformation, whereas antiparallel and hybrid G4s have both *syn* and *anti*.

G-tetrads are linked by loops, which can be classified in four groups (Bugaut and Balasubramanian, 2008):

- Lateral loop, connecting two adjacent antiparallel strands (Figure 1.9a);
- Diagonal loop, connecting two opposing antiparallel strands across the G4 plane (Figure 1.9b);
- Double-chain-reversal loop or propeller loop, connecting two adjacent parallel strands (Figure 1.9c);
- V-shaped loop, connecting two corners of a G-tetrad core in which a support column is missing (Figure 1.9d).

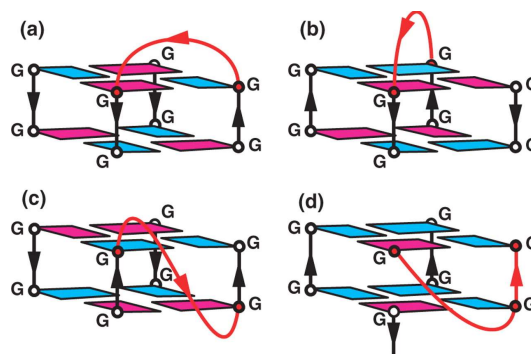


Figure 1.9. Schematic illustration showing (a) lateral loop; (b) diagonal loop; (c) propeller loop; (d) V-shaped loop [adapted from (Patel et al., 2007)].

Loops are usually short (1-7 nucleotides) and smaller loops result in more stable G4s (Bugaut and Balasubramanian, 2008; Gros et al., 2010).

All G4s have four grooves, which are the cavities delimited by the phosphodiester backbone. Groove dimensions vary and depend on the G4 overall topology and loops nature. If all strands are parallel, guanine angles are in the *anti* conformation, and the grooves are identical and medium-sized. If the glycosidic bonds orientations of the bases are *syn-syn-anti-anti*, two grooves are medium-sized, one narrow-sized and one wide-sized (Figure 1.10A). Finally, if the glycosidic bonds orientations of the bases are *syn-anti-syn-anti*, two grooves are narrow and the other two ones are wide (Figure 1.10B) (Burge et al., 2006).

Several studies have shown that in particular regions of the genome, G-rich sequences can form unimolecular G4s in cell (Biffi et al., 2013; Lipps and Rhodes, 2009). In contrast to bimolecular and tetramolecular G4s, intramolecular structures can be formed quickly and are more complex due to great conformational diversity, such as in folding topologies and loops conformations (Yang and Okamoto, 2010).

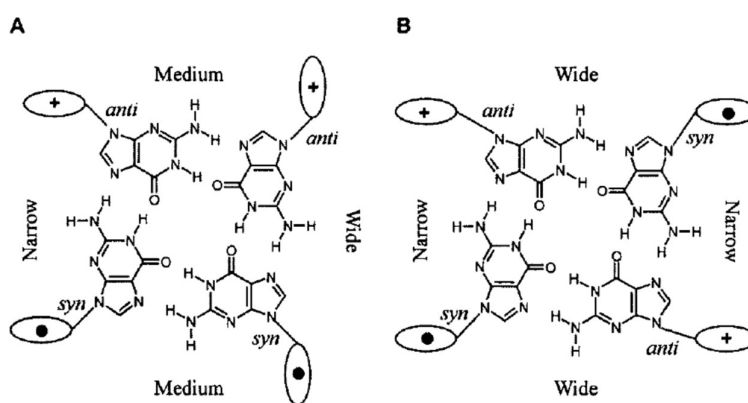


Figure 1.10. A) *syn-syn-anti-anti* glycosidic bonds orientation; B) *syn-anti-syn-anti* glycosidic bonds orientation.

1.2.2 G-triplex

Recent studies revealed that the folding and unfolding pathways of G4s proceed through a quite stable intermediate named G3. A G3 structure differs from the known triplex structures not only for the base pairing, but also for the structure. Indeed, this structure is characterized by the presence of G:G:G triad planes stabilized by an array of Hoogsteen-like hydrogen bonds similar to G4s. However, the lower number of hydrogen bonds and the smaller stacking surface of the triads with respect to the G-tetrads, make G3s much less stable than G4s and, therefore,

even more challenging to isolate both *in vitro* and *in vivo*. For all these reasons, it is clear that the discovery of molecules able to interact and stabilize G3s is highly appealing, also for the understanding of the putative biological and therapeutic importance of these intermediates.

Recent studies allowed to isolate and characterize a G3 structure in details. In particular, it was demonstrated that a truncated form of the G4-forming thrombin binding aptamer (TBA) forms a relatively stable G3. (Limongelli et al., 2013). This structure has two G-triads (G1:G6:G10 and G2:G5:G11 Figure 1.11A), characterized by a *syn-anti-syn* and *anti-syn-anti* arrangement of the residues, respectively. In this conformation, the metal ion seems to be placed at the center of the two G-triads in a way similar to that of the G4 structure (Figure 1.11B).

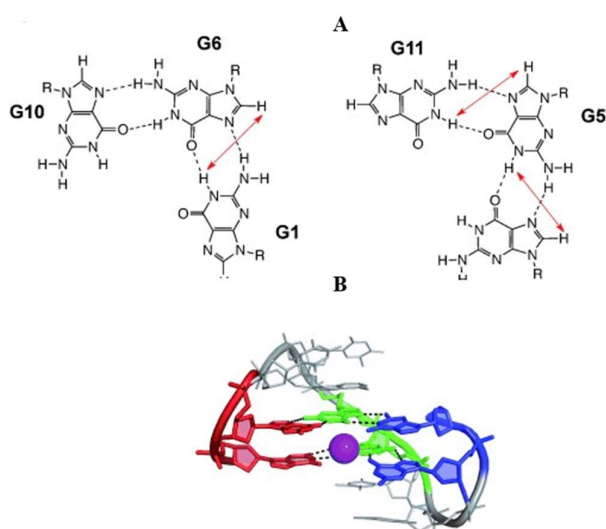


Figure 1.11. A) G-triads involved in the formation of the G-triplex; B) 3D representation of the G-triplex [adapted from (Limongelli et al., 2013)].

1.2.3 *i*-Motif

In 1993 Gehring, Leroy and Guéron observed that cytosine rich sequences can form four stranded structures under acidic conditions (Gehring et al., 1993). They found the d(TCCCC) oligomer to form a four-stranded structure under acidic pH, two base-paired parallel-stranded duplexes are intimately associated and, fully intercalated. The relative orientation of the duplexes is antiparallel, so that each base pair is face-to-face with its neighbors (Figure 1.12).

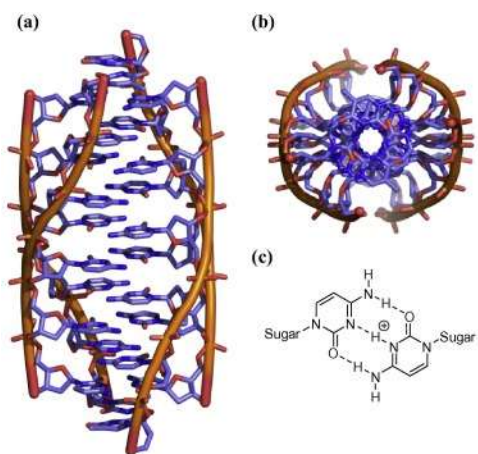


Figure 1.12. (a) Structure of the intermolecular i-motif formed by d(TC₅) (PDB ID: 225D), identified by Gehring et al. (b) Top view of the i-motif formed by d(TC₅). (c) A hemiprotonated cytosine–cytosine⁺ base pair [adapted from (Day et al., 2014)].

Due to the unusual nature of the configuration, they called this new non-canonical nucleic acid structure intercalated (i) motif. Sequences with two stretches of cytosines can also form dimeric i-motif structures where two hairpins intercalate (Pairs et al., 1997); and, more interestingly from a biological perspective, natural sequences with four stretches of cytosines separated by other bases can fold into intramolecular i-motif structures (Figure 1.13).

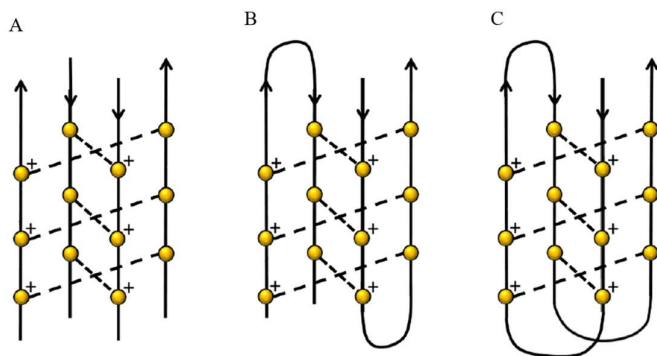


Figure 1.13. Schematic representation of tetrameric (A), dimeric (B) and monomeric i-motif structures (C). The yellow spheres represent the cytosines, plus signs indicate the protonated cytosines.

There are two different i-motif topologies depending on the intercalation: i) the 3'E, in which the outermost cytosine pair is from the 3' end; ii) the 5'E, where the outermost pair is from the 5' end (Guéron and Leroy, 2000; Nonin-lecomte and Leroy, 2001), and iii) the less stable T-

form, which features non-intercalated outermost C-C⁺ pair (Kanaori et al., 2001). The intramolecular i-motifs have also been classified according to the length of the loops (Brooks et al., 2010): ‘class I’ i-motifs have shorter loops, whereas ‘class II’ i-motifs have longer loops. In general, ‘class II’ i-motifs are thought to be the more stable due to extra stabilizing interactions within the longer loop regions.

As the formation of an i-motif requires the cytosine hemiprotonation, this structure is formed and is stable *in vitro* only under slightly acidic conditions (Zhou et al., 2010). However, it has been demonstrated that the negative superhelicity induced by transcriptional activity (Sun and Hurley, 2009), and cell-mimicking molecular crowding conditions (Rajendran et al., 2010) facilitate i-motif formation.

1.3 Biological relevance of G-quadruplexes

During the last decade, G4s have emerged from being a structural curiosity observed *in vitro* to being recognized as possible regulators of multiple biological processes. The sequencing of many genomes has revealed that they are rich in motifs with the potential to form G4s and that the location of those motifs is non-random, correlating with functionally important genomic regions.

The evidence of G4s formation in cell was first observed in the telomeric region, which is at the end of the eukaryotic chromosomes. The first evidence was obtained using a specific antibody for intermolecular telomeric G4 DNA of the ciliate *Stylonychia lemnae* (Schaffitzel et al., 2001). Further evidence was achieved observing that telomere end-binding proteins control the formation of G4 DNA structures *in vivo* (Paeschke et al., 2005). Bioinformatic sequence analyses have also showed that such putative G4-forming sequences are present in the genome of different organisms (Rawal et al., 2006; Todd et al., 2005). There is an estimate of ~700000 potential G4-forming sequences in the human genome (Chambers et al., 2015).

The telomeric structure and stability directly relate to aging (Bodnar et al., 1998), genome stability (Hackett et al., 2001), and cancer growth (Neidle and Parkinson, 2002). Human telomeric DNA comprises a long stretch of double-stranded tandem repeats and a short, single-stranded (TTAGGG)_n 3'-overhang (Hemann and Greider, 1999; Zakian, 1995). It has been demonstrated that truncations of this G-rich sequence can form G4 structures *in vitro* by X-ray crystallography and nuclear magnetic resonance (Parkinson et al., 2002; Phan et al., 2007b; Wang and Patel, 1993; Williamson et al., 1989; Zhang et al., 2010).

Telomeres regulate cell senescence as they progressively shorten at each replication round until apoptosis is triggered by genomic instability (Sfeir et al., 2005). Telomerase is the ribonucleoprotein complex with reverse transcriptase activity that elongates telomeres (Greider and Blackburn, 1985). Due to overactivation, it is responsible for cell “immortalization” in 80-85% of all human cancer (Kim et al., 2011). Telomerase recognizes the single-stranded telomeric DNA, but not the G4 structure. Therefore, the discovery of telomeric G4 ligands could be a way to induce the apoptosis in cancer cells. Furthermore, telomeric DNA also binds many other proteins like the shelterin complex. This complex is composed by six subunits: the telomere repeat factor 1 (TRF1), telomere repeat factor 2 (TRF2), protection of telomere 1 (POT1), repressor/activator protein 1 (RAP1), TRF1- and TRF2-interacting nuclear protein 2 (TIN2), and tripeptidyl peptidase 1 (TPP1). They can operate in subsets to regulate the length of or protect telomeres. In this context, small molecules competing with these proteins for the binding to the telomeric DNA would induce a DNA damage response, that quickly promotes senescence in cancer cells.

Other regions of genome that are particularly enriched in putative G4 motifs are the gene promoters (~1 kb upstream of transcription start sites - TSS). G4s within gene promoters are now proved to regulate transcription (Huppert and Balasubramanian, 2007; Patel et al., 2007). The first evidence supporting the existence of unusual DNA forms in gene promoters dates back to 1982, and it is based on the nuclease hypersensitivity element (NHE) in the chicken β -globulin gene promoter (Larsen and Weintraub, 1982; Wood and Felsenfeld, 1982), that was only later associated with guanine runs.

G4s have been reported within the promoter of several genes, e. g. *INS* (Hammond-kosack et al., 1992), *MYC* (Siddiqui-Jain et al., 2002; Simonsson et al., 1998), *VEGF* (Sun et al., 2005), *HIF-1 α* (Armond et al., 2005), *BCL-2* (Dai et al., 2006), *KIT* (Fernando et al., 2006; Rankin et al., 2005), *RET* (Zhou et al., 2010), *PDGF-A* (Qin et al., 2007), and *KRAS* (Cogoi et al., 2004). In 2008, Du *et al.* performed a comprehensive analysis of the relationship between potential G4 DNA motifs in the putative transcriptional regulatory region and gene expression level (Du et al., 2008). The *MYC* oncogene is a transcription factor that regulates the expression of a variety of genes and it is one of the most prevalent oncogenes found to be altered in human cancer, being dysregulated in about 50% of tumors (Delgado et al., 2013). The transcriptional regulation of *MYC* is tightly controlled by a complex mechanism involving four promoters (P1–P4), different transcription start sites (TSS) and nuclease hypersensitive elements (NHE). In

particular, the NHE III1, located just upstream the promoter P1 is responsible for the great majority of *MYC* transcription. It is composed of five consecutive runs of the sequence (G/A)G(G/A)AGGGGT that can form a G4 structure (González and Hurley, 2010). As a consequence, the possibility to inhibit *MYC* transcription through G4 stabilization has been actively pursued in several human cancer models using specific small molecules (González and Hurley, 2010).

A number of G4 structures from gene promoters has been determined by X-ray or NMR techniques (Ambrus et al., 2005; Hsu et al., 2009; Wei et al., 2012). In contrast to the tandem repeats in human telomeres, the G-rich sequences within gene promoters are often composed by G-tracts with an unequal number of guanines and a varying number of intervening bases. Each of these promoter sequences is unique in length, and type of G-tracts, and intervening bases (Qin and Hurley, 2008) and can form multiple G4s, i. e. mixtures of conformations. Notably, G4s from promoters exhibit the highly conserved G₃NG₃ motifs, which forms a robust parallel-stranded structure with a 1-nt, double-chain-reversal loop (Yang and Okamoto, 2010) (Figure 1.14).

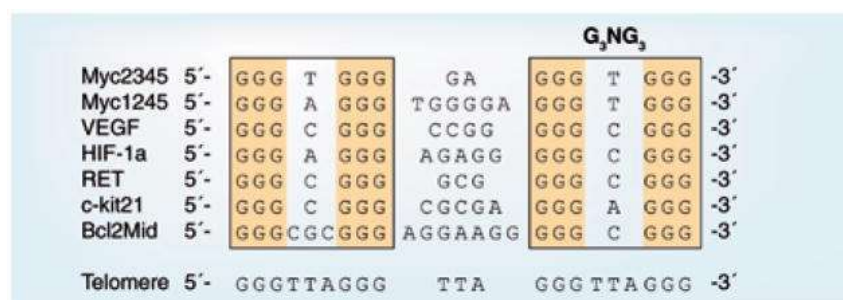


Figure 1.14. Comparison of G4-forming sequences in selected gene promoters. The telomeric sequence is also shown as a comparison. All the promoter G-rich sequences shown contain the G₃NG₃ motif; except for the *BCL-2* (*Bcl2Mid*) sequence, they have all been shown to form parallel-stranded G-quadruplexes.

1.3.1 Targeting G4 structure in cancer cells

Considering the crucial position of G4s in the genome and their important role in the carcinogenic processes, targeting them with small molecules could be a potential anticancer approach. In the last decade, a huge number of small molecules has been studied to evaluate their ability to stabilize G4 structures.

The natural product Telomestatin (Figure 1.15A) has been identified as a potent telomeric G4 stabilizer (Sun et al., 1997). It competes POT1 for the binding to the telomeric G4 DNA triggering DNA damage response and cellular senescence (Gomez et al., 2006; Temime-Smaali

et al., 2009). The same effect has been also observed for a panel of other G4-interacting small molecules. The triazine 12459 (Figure 1.15B) induces telomere shortening by POT1 uncapping, leading to cellular senescence (Gomez et al., 2003; Riou et al., 2002). The acridine BRACO-19 (Figure 1.15C), one of the very first promising synthetic G4-targeting small molecules, is a very potent telomerase inhibitor (Gunaratnam et al., 2007). RHPS4 and Pyridostatin (Figure 1.15D-E) induce telomere uncapping leading to a telomere-induced DNA damage response (Leonetti, 2004; Phatak et al., 2007; Rodriguez et al., 2008; Salvati et al., 2007).

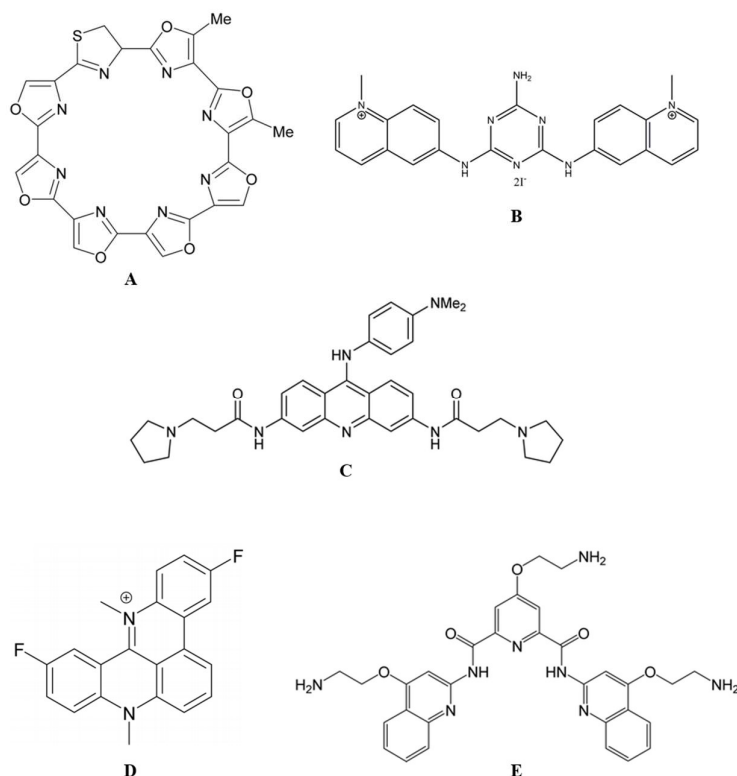


Figure 1.15. Structure of (A) Telomestatin, (B) 12459, (C) BRACO-19, (D) RHPS4, and (E) Pyridostatin.

Therefore, small molecules that bind and stabilize these structures are generally able to inhibit the telomerase thereby triggering DNA damage responses as with other mediators of telomere damage (d'Adda di Fagagna et al., 2003).

G4s could be also the potential targets in a new anticancer strategy based on search for G4-selective stabilizing ligands aimed at inhibition of oncogenes expression (Balasubramanian et al., 2011) (Figure 1.16).

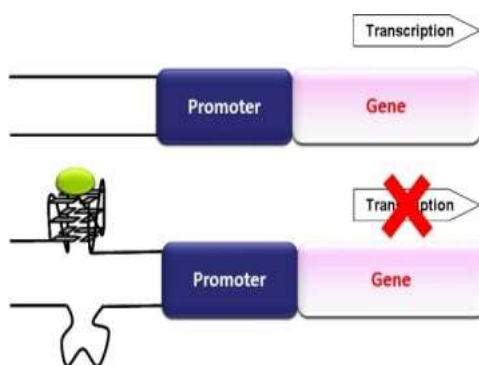


Figure 1.16. Scheme representing the targeting of a G4 to suppress oncogene transcription.

For example, it has been proved that the cationic porphyrin derivate TMPyP4 (Figure 1.17) can downregulate the prominent oncogene *MYC* by stabilizing a G4 structure in its promoter (Siddiqui-Jain et al., 2002). Interestingly, a G > T mutation in the G4-forming sequence within the *MYC* promoter sufficed to destabilize the G4 and resulted in increased transcription, which was only partially reduced by TMPyP4.

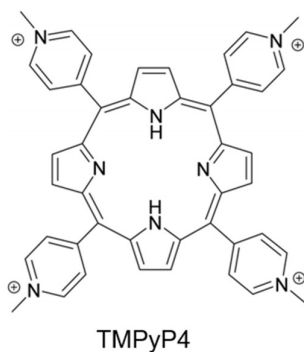


Figure 1.17. Chemical structure of TMPyP4.

Another example of proto-oncogene with a G4-forming sequence within this promoter is *KIT* gene, which encodes for a tyrosine protein kinase (Fernando et al., 2006; Rankin et al., 2005). It was demonstrated that 3,8,10-trisubstituted isoalloxazines stabilize G4s in the *KIT* promoter and inhibit its expression (Bejugam et al., 2007).

1.4 Biological relevance of i-motif

As yet, there has been no systematic analysis of the genome to find sequences capable of forming i-motifs. Due to the complementary nature of putative G4 and i-motif forming sequences, it is likely that there is some overlap in their prevalence. However, the stability and the dynamics of i-motif structures are likely to vary a lot, depending on the sequence. Initial investigations on natural i-motif forming sequences focused on those from the complementary C-rich strands of known promoter G4-forming sequences. A range of i-motif forming sequences has now been identified and characterized (Table 1.1).

Table 1.1. Examples of investigated i-motif forming sequences [adapted from (Day et al., 2014)].

Name	Sequence 5' → 3'
hTelo	d(CCCTAACCTAACCTAACCT)
c-MYC	d(CCCCACCTTCCCCACCTCCCCACCTCCCC)
bcl-2	d(CAGCCCCGCTCCCGCCCCCTTCTCCCGCGCCCGCCCT)
Rb	d(GCCGCCAAAACCCCCG)
RET	d(CCGCCCCGCCCCGCCCCGCCCTA)
VEGF-A	d(GACCCCGCCCCGGCCCGCCCCGG)
c-ki-ras	d(GCTCCCTCCCTCCCTCCTTCCCTCCCTCCC)
c-kit	d(CCCTCCTCCAGCGCCACCT)
PDGF-A	d(CCGCGCCCCTCCCCGCCCCGCCCCGCCCCCCCCCCCC)
c-myb	d(TCCTCCTCCTTCTCCTCCTCCTCGTGTCTCCTCCTCC)
hTERT	d(CCCCGCCCCGTCCCGACCCCTCCCGGGTCCCGGCCAGCCCCACCGGGCCCTCCAGCC CCTCCCC)
HIF-1 α	d(CGCGCTCCCGCCCCCTCCCCTCCCCGCGCGCCGAGCGCGCTCCGCCCTTGCCCCCCCC CTG)
c-jun	d(TAACCCCTCCCCCTCCCCCTTTAAT)
ILPR	d(TGTCCCCACACCCTGTCCCCACACCCTGT)
n-MYC	d(ACCCCTGCATCTGCATGCCCCCTCCACCCCT)

Furthermore, some proteins have been identified to bind C-rich stands in human telomeres with specificity (Day et al., 2014).

As for putative i-motif forming sequences in gene promoter regions, the NH3 III₁ of *MYC* proto-oncogene contains a C-rich strand capable of forming an i-motif. Another oncogene promoter that has a C-rich sequence prone to form i-motif is the P1 promoter of the *B-cell lymphoma-2 (BCL-2)* proto-oncogene. Unlike other proto-oncogenes, *BCL-2* promotes cell survival and proliferation via an anti-apoptotic function rather than affecting growth factors

(Joensuu et al., 1994). Interestingly, although *BCL-2* is overexpressed in some cancers, in neurodegenerative diseases such as Alzheimer's and Parkinson's it may be under expressed, leaving the cells more vulnerable to apoptosis (Bar-Am et al., 2005). Studies on the i-motif were previously limited, based on the assumption that, because they are stabilized in slightly acidic conditions, they were not physiologically relevant, despite a solid foundation of data indicating that these structures are detectable at neutral pH *in vitro* (Mergny et al., 1995).

In a recent study, Dzatko *et al.* provided the evidence for i-motif formation *in vivo* by in cell NMR experiments (Dzatko et al., 2018). Moreover, Zeraati *et al.* reported the generation and characterization of an antibody fragment (iMab) that recognizes i-motif structures with high selectivity and affinity, enabling to detect i-motifs in cells nuclei (Zeraati et al., 2018). This study also showed that the *in vivo* formation of such structures is also cell-cycle dependent. Thanks to iMab, it was demonstrated that i-motif structures are actually formed in regulatory regions of the human genome, including promoters and telomeric regions. These results support the notion that, as the G4s, i-motif structures could have some key regulatory role in the genome.

1.5 DNA sequences as therapeutic agents

Nucleic acids can be seen not only as a therapeutic target, but also as a new class of potential drugs. Among these are the so-called DNA and RNA aptamers. The term 'aptamer' derives from the combination of the Latin word *aptus* ('to fit') and the Greek word *meros* ('part').

Nucleic acid aptamers are short, single-stranded DNA (ssDNA) or RNA molecules that are selected for binding to a specific target (Morita et al., 2018). The first aptamer was isolated in 1990, thereafter thousands of aptamers have been generated against a wide range of target, including small metal ions, organic molecules, peptides, proteins, viruses, bacteria, whole cells and even targets within live animals.

Aptamers can fold into various non canonical nucleic acid structures (for example stem, loop, bulge, G4 and hairpin) (Mayer, 2009), and the peculiar three-dimensional structures allow them to specifically recognize their cognate target. The aptamer-target binding leads to a complex which is similar to an antibody-antigen complex. However, aptamers offer several advantages over traditional antibodies, including smaller size, quick and cheap production, versatile chemical modification, high stability, and low immunogenicity.

To date, Pegaptanib ([®]Macugen) is the only commercially available aptamer (Figure 1.18). It is a chemically-modified RNA molecule that binds to the isoform 165 of the anti-vascular endothelial growth factor (VEGF), a protein that plays a critical role in angiogenesis (the formation of new blood vessels) and increased permeability (leakage from blood vessels), two of the primary pathological processes responsible for the vision loss associated with neovascular age-related macular degeneration (AMD). Pegaptanib works as VEGF antagonist, hence reducing the growth of the blood vessels located within the eye and works to control the leakage and swelling.

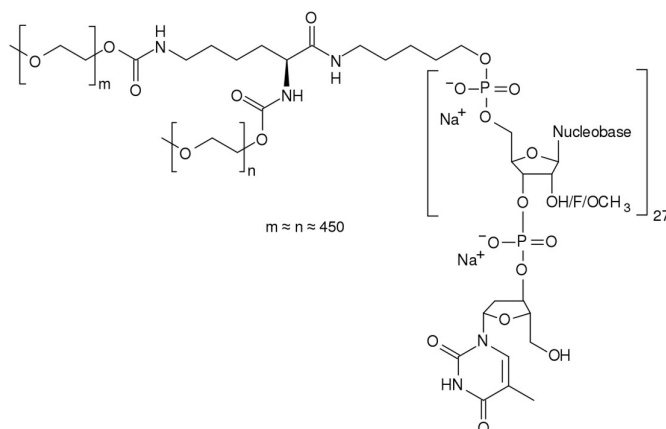


Figure 1.18. Chemical structure of Pegaptanib.

The thrombin binding aptamer (TBA) is further example. It is a 15-mer DNA d(GGTTGGTGTGGTTGG) able to fold into an antiparallel G4 (Bock et al., 1992; Nagatoishi et al., 2007). It binds to the exosite I of human alpha-thrombin, i. e. binding site of fibrinogen, and acts as an anti-coagulant agent by inhibiting the activation of fibrinogen as well as platelet aggregation. In addition, TBA shows good affinity and specificity against thrombin. TBA does not interact with other plasma proteins or thrombin analogues (e.g., gamma-thrombin) (Li et al., 2002) (Figure 1.19). As a result, TBA has been used as a short-term anti-coagulant in the coronary artery bypass graft surgery. Its optimized form (NU172) is now under the phase II of clinical trial by ARCA Biopharma (NCT00808964). In addition, due to its high affinity and specificity, a variety of sensors was coupled with TBA, and developed for thrombosis diagnostics.

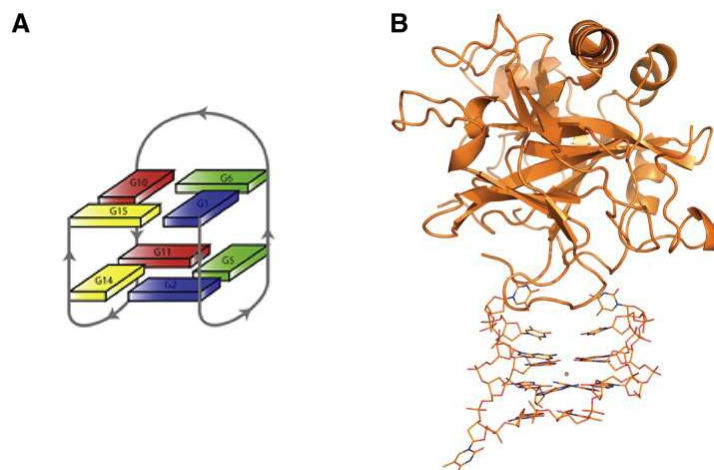


Figure 1.19. Schematic illustration showing (A) the G4 structure of the thrombin binding aptamer (TBA) and (B) its interaction with the thrombin exosite I according to the X-ray and NMR studies. Thrombin molecule is represented as cartoon, TBA molecule is represented as sticks [adapted from (Krauss *et al.*, 2012)].

One of the several therapeutic applications of aptamers is in oncological treatments, where they are used for their selective affinity against proteins that play fundamental roles in cancer cells.

AS1411, formerly named ARGO100, is a 26-nucleotide G4-forming oligonucleotide and it is the first aptamer in clinical trials for the treatment of human cancer (Soundararajan *et al.*, 2009). It was originally selected from a screen of antisense oligonucleotides with antiproliferative function, then pegylated to prolong its half-life for *in vivo* applications. Owing to its G4 structure, AS1411 is quite resistant to nuclease degradation. AS1411 shows high affinity to the external domain of nucleolin (Bates *et al.*, 1999; Soundararajan *et al.*, 2008), which is a *BCL2*-mRNA-binding protein involved in cell survival, growth and proliferation, and it is overexpressed on the surface of tumor cells (Berger *et al.*, 2015). After binding to nucleolin, and also thanks to macropinocytosis, AS1411 is rapidly internalized, thus leading to the induction of apoptosis. AS1411 can function as a tumor-targeting agent and it is a promising anticancer agent. It inhibited cancer in various preclinical models (Reyes-Reyes *et al.*, 2010), including breast, renal and lung cancer, and in 2007 it was reported to be well tolerated without serious side effects in an open-label phase I study on advanced solid tumors (Bates *et al.*, 2009) (ClinicalTrials.gov identifier NCT00881244). In 2009, a phase II trial assessed the efficacy and safety of AS1411 combined with cytarabine.

Chapter 2

METHODS

2.1 Circular dichroism

Circular dichroism (CD) spectroscopy is a of light absorption technique that measures the difference in absorbance between the right- and left-circularly polarized light, rather than the isotropic light absorption. It is widely used to study the conformations of biomacromolecules being very sensitive to their secondary structure, and also thanks to its other advantages, like:

- Very low sample amount required;
- Very sensitive detection of conformational changes due to temperature, pH and solvent variations.

The electromagnetic radiation is a complex wave that propagates at the speed of light. It is made by two waves that are at right angles to each other and perpendicular to the direction of wave propagation: the magnetic (B) and the electric (E) component. Because the two components are invariably perpendicular to each other, it is an optimal approximation to consider only one of them to simplify the wave description i.e. the E-component. The wave can oscillate in any direction perpendicular to the direction of propagation. When the wave passes through a polarizer, it is polarized. Linear polarized light occurs when the electric field vector (E) propagates in only one plane perpendicular to the direction of propagation, the direction of the vector stays constant and the magnitude oscillates (Figure 2.1A). In circularly polarized light, the E rotates around the propagation axis maintaining a constant magnitude (Figure 2.1B). Light can be circularly polarized in two directions: left and right. If the vector rotates counterclockwise when the observer looks down the axis of propagation, the light is left circularly polarized. If it rotates clockwise, it is right circularly polarized. The differential absorption of radiation polarized in two directions as function of frequency is called dichroism. For plane-polarized light this is called linear dichroism, while for circularly polarized light it is called circular dichroism.

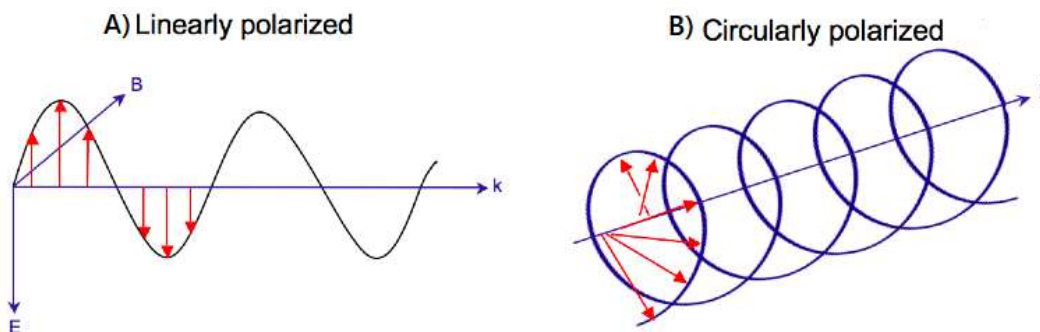


Figure 2.1. A) linearly polarized light; B) circularly polarized light.

Inherently asymmetric chromophores (uncommon) or symmetric chromophores in asymmetric environments interact with right- and left-circularly polarized light differently, resulting in two related phenomena. Circularly-polarized light will travel through an optically active medium with different velocities due to the different indices of refraction for right- and left-circularly polarized light. This is called optical rotation or circular birefringence. Right- and left-circularly polarized light will also be absorbed to different extents at some wavelengths due to differences in extinction coefficients for the two polarized rays called circular dichroism (CD):

$$CD = A_L - A_R$$

in which A_L is the absorbance of the left-polarized light, and A_R is the absorbance of the right-polarized one.

The absorbance is obtained by the Lambert-Beer law:

$$A = \epsilon l c$$

in which ϵ is the wavelength-dependent molar absorptivity coefficient ($M^{-1} \text{ cm}^{-1}$), l is the optical path length (cm) and c is the sample concentration (M).

During a CD measurement l and c are constant, hence the CD signal can be expressed as difference between molar absorptivity coefficients:

$$CD = \epsilon_L - \epsilon_R$$

An optically active sample, i.e. having different extinction coefficient toward right- and left-polarized light, introduces difference between the magnitude of the two components. The resultant E-vector no longer traces a circle over time but rather an ellipse, thus the light is elliptically polarized. Ellipticity is the unit of circular dichroism and is defined as the tangent of the ratio of the minor to major elliptical axis. The ellipticity is directly proportional to the difference between molar absorptivity coefficients of the two light components:

$$\theta = \frac{\pi\lambda(\epsilon_R - \epsilon_L)}{l}$$

in which l is the optical path and λ is the wavelength of the incident radiation.

A CD spectrum is a measure of the ellipticity vs the wavelength. The spectrum can be expressed in molar ellipticity ($[\theta]$) in order to compare spectra achieved at different concentrations:

$$[\theta] = \frac{\theta}{[c] \cdot l} = 32.98 \Delta\epsilon \text{ degrees M}^{-1} \text{ cm}^{-1}$$

A CD spectrum is recorded in the UV-Vis wavelength region. It is similar to the absorption spectrum, but it can have positive and negative bands being a difference between molar absorptivity coefficients.

The chromophores that are active to CD are intrinsically dissymmetric molecules. The dissymmetry can be structural or a consequence of the presence of a chiral center. CD can be applied only to samples that absorb in the UV-Vis region. Buffer baseline is subtracted from the CD spectrum of each sample.

2.2 Fluorescence Spectroscopy

Fluorescence spectroscopy is a technique used to analyze fluorescence emission of a sample. It is based on the use of UV or Vis radiation that excites the electrons of molecules in the sample and causes them to emit an electromagnetic radiation.

The fluorescence spectroscopy is:

- Rapid;
- Non-destructive technique;
- Relatively inexpensive;
- Sensitive: limits of sensitivity depend on intrinsic fluorescent of the sample; however this technique is usually able to detect very small concentrations of analytes (μM) and has a sensitivity higher than absorption spectroscopy;
- Specific: the number of fluorescent molecules is limited;
- Useful to study different types of molecules.

Moreover, the emission of fluorescence radiation is approximately linear in a wide range of concentrations without necessity of sample dilution.

Molecules have many accessible electronic states. The electron transition from the ground state to an excited state may be induced by electromagnetic radiation. The relaxation to the ground state can be spontaneous, because the molecule average lifetime in an excited state is limited. In this case the phenomenon is called spontaneous emission. Emission can also be induced by the interaction with an electromagnetic radiation and it is called stimulated emission. Fluorescence is a spontaneous emission of electromagnetic radiation.

The phenomenon of spontaneous emission can be explained using the Jablonski diagram (Figure 2.2):

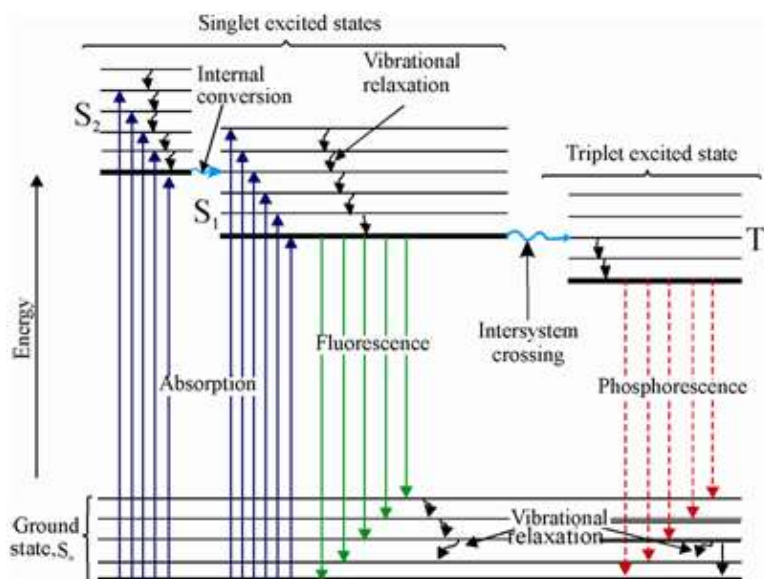


Figure 2.2. Jablonski diagram.

After molecule absorption of an UV-Vis radiation, the most probable transition occurs with conservation of spin multiplicity, i.e. it is the transition from the singlet ground state (S_0) to the first singlet excited state (S_1).

The molecule in the excited electronic state may be in any vibrational level and it has the tendency to return to the ground state. The first phenomenon is the transition from an excited to the ground vibrational level. It is called vibrational relaxation and involves the dissipation of energy from the molecule to its surroundings. The fluorescence is the emission of electromagnetic radiation due to the electron transition from S_1 to S_0 . The intersystem crossing is the electron transition from S_1 to the triplet excited electronic state (T_1). The phosphorescence is the emission of radiation due to the electron transition from T_1 to S_0 . When the molecule

returns to S_0 , it may be in any vibrational level and so successively a vibrational relaxation occurs.

Emission bands of a molecule have greater wavelengths compared to those of its excitation bands, it is a consequence of energy loss due to non-radiative phenomena. As a result, the fluorescence spectrum is similar to the absorption spectrum, but it is shifted to higher wavelengths and this shift is called Stokes shift.

A relevant aspect of the fluorescence is the impossibility to recover the total energy given to the molecule with the excitation. The fluorescence quantum yield (ϕ_F) provides an estimation of the energy recovered through the fluorescence process. It is defined as the ratio of the number of photons emitted to the number of photons absorbed. Possible ϕ_F values are between zero and one: the ratio is near to one if the molecule has a high fluorescence, ϕ_F is near to zero if the molecule has a low fluorescence. The fluorescent molecule is called fluorophore.

2.3 Electrophoresis

Electrophoresis is a general term that describes the migration and separation of charged particles under the influence of an electric field. Most of biological molecules (e.g. amino acids, peptides, proteins, nucleotides, and nucleic acids) has ionizable groups and therefore they can be present in solution as charged species in dependence of the pH.

Gel electrophoresis (GE) is a widely used technique to separate, purify, and identify ionic or ionizable macromolecules in solution. An electrophoretic system consists of two electrodes, i.e. anode and cathode, connected by a conducting medium called electrolyte, in which the electrophoresis gel is immersed. Sample is loaded in appropriate wells and species migrate towards either electrode according to their charge.

The electric force (F_e) is given by the equation:

$$F_e = E q$$

in which E is the electric field and q the charge.

There is a friction that opposes to the motion generated by the electric field. The friction force is given by:

$$F_f = \gamma v$$

in which γ is the friction coefficient and v the migration velocity.

At constant speed:

$$F_e = F_f$$

from which is drawn that:

$$v = \frac{E q}{\gamma}$$

The electrophoretic mobility (m) is defined as:

$$m = \frac{v}{E} = \frac{q}{\gamma}$$

It is noted that the electrophoretic mobility exclusively depends on the particle.

In the case of a globular molecule, γ is given by:

$$\gamma = 6 \pi \eta r$$

in which η is the medium viscosity and r the molecule radius.

Then m is:

$$m = \frac{q}{6 \pi \eta r}$$

In conclusion, the electrophoretic mobility of a molecule is determined by its charge, dimension and shape, as well as by the medium viscosity and gel pore size. The electrophoresis allows the separation of molecules from a mixture according to their charge and size.

2.4 Differential scanning calorimetry

Differential scanning calorimetry (DSC) is a calorimetric technique used to study conformational transitions of biomacromolecules and their thermal stability (Gill et al., 2010).

DSC major advantages are:

- Universal technique for its possible application in different research areas;
- Direct method to determine thermodynamic parameters;
- The absence of optical constituents allows using cloudy/colored solutions.

DSC is the main technique for measuring the thermal properties of biomacromolecules, and particularly to monitor stability and conformational transitions. It measures the heat capacity (C_p) of a sample as a function of temperature, in comparison to a reference.

In a basic DSC experiment, the sample cell (Figure 2.3) contains a solution with the molecule of interest, while the reference cell contains only the solvent. Temperature is raised identically in the two cells at constant rate. The difference in energy required to match the temperature of

the two cells is the heat either absorbed or released by molecules in the sample during a transition. These differences in input energy is a measure of molecule heat capacity.

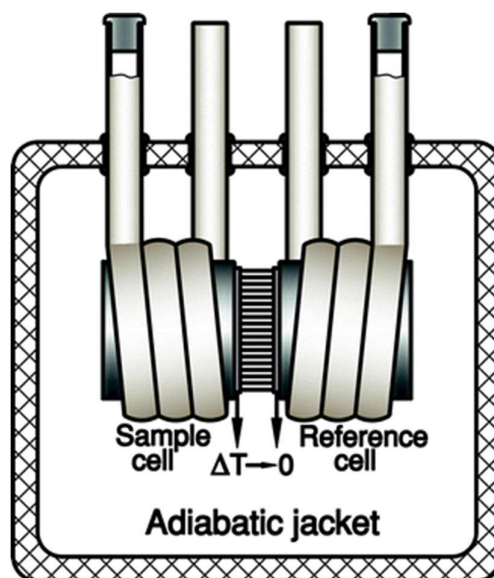


Figure 2.3. Schematic representation of a DSC cells.

An accurate buffer vs buffer (or solvent vs. solvent) baseline scan is required in this technique. A second baseline must be applied to the data to extrapolate pre- and post-transition DSC curve. The choice of this baseline is the most critical step in the analysis because it affects the value of thermodynamic parameters. Several baseline procedures are implemented in DSC software, but the simplest approach is a linear extrapolation.

After a scan sample vs. buffer, the result is a curve (Figure 2.4).

It is possible to program several consecutive heating/cooling cycles to check for reversibility. If the subsequent DSC scans of the same sample overlap, this means that the observed unfolding/folding process is fast compared to the scan rate employed in the experiment and can be considered, at each temperature, an equilibrium process; suggesting the full reversibility of the transition.

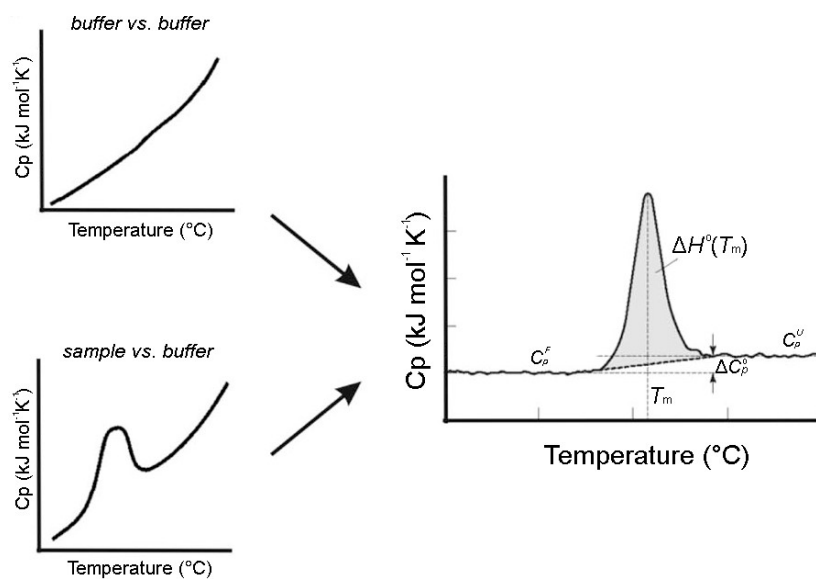


Figure 2.4. Buffer baseline scan, sample scan (on the left) and final DSC profile (on the right) after subtraction of the buffer baseline are shown. T_m is the melting temperature corresponding to the maximum of the curve, $\Delta H^\circ(T_m)$ is the unfolding enthalpy corresponding to the area under the curve and C_p^F and C_p^U are the heat capacities of folded (F) and unfolded (U) structures corresponding to the pre- and post-transition lines, respectively.

Thermodynamic parameters obtained from DSC experiments are quite sensitive to the structural state of biomolecules and associate with heat-induced macromolecular transitions.

The initial part of the curve provides the heat capacity (C_p^F) of the folded state; while, the final part provides the heat capacity (C_p^U) of the unfolded state instead; their difference represents the ΔC_p of the molecules transition. The peak is centered at the transition melting temperature (T_m). The melting temperature is considered as the temperature at which 50% of molecules owns its folded conformation and the rest remains unfolded. Higher T_m values are representative of a more stable molecule.

Calorimetric enthalpy (ΔH°_{cal}) is the total integrated zone below the peak, which indicates the total transition energy after suitable baseline correction:

$$\Delta H^\circ_{cal}(T_m) = \int_{T_i}^{T_f} C_p^o dT$$

in which T_i and T_f are the initial and the final temperature, respectively.

Van't Hoff enthalpy ($\Delta H^{\circ}_{\text{VH}}$) is an independent measurement of the transition enthalpy obtained assuming a two-state transition. $\Delta H^{\circ}_{\text{VH}}$ is determined through parameters derived from the experimental graph:

$$\Delta H^{\circ}_{\text{VH}} = kRT^2\Delta C_p(T_m)/\Delta H^{\circ}_{\text{cal}}(T_m)$$

in which k is a constant derived from transition stoichiometry, in detail it values 6 for nucleic acids and 4 for proteins.

Information on the transition can be obtained comparing $\Delta H^{\circ}_{\text{VH}}$ with $\Delta H^{\circ}_{\text{cal}}$. If $\Delta H^{\circ}_{\text{VH}}$ is equal to $\Delta H^{\circ}_{\text{cal}}$, the transition occurs in a two-state mode. When $\Delta H^{\circ}_{\text{VH}}$ exceeds $\Delta H^{\circ}_{\text{cal}}$, the macromolecule may be associated as dimer or multimer. When $\Delta H^{\circ}_{\text{VH}}$ is less than $\Delta H^{\circ}_{\text{cal}}$, one or more intermediates are expected.

The other thermodynamic parameters are derived from these calorimetric ones. The change in entropy (ΔS°) is obtained by:

$$\Delta S^{\circ}(T_m) = \int_{T_i}^{T_f} \frac{C_p^{\circ}}{T} dT$$

ΔG° must be calculated at a chosen reference temperature:

$$\Delta G^{\circ}(T) = \Delta H^{\circ}(T) - T\Delta S^{\circ}(T)$$

2.5 Nuclear Magnetic Resonance

Nuclear magnetic resonance (NMR) is a spectroscopic method based on the interaction between energy and matter but, differently from other spectroscopic techniques, it needs a strong static magnetic field.

The usefulness of NMR in chemistry was not appreciated until 1950 when the importance of the so-called *chemical shift* was discovered (Proctor and Yu, 1950). Indeed, it is in chemistry that NMR has attracted the greatest interest in the recent years.

Nuclear Magnetic Resonance is a property of the nucleus of an atom, which consists of protons and neutrons, related to what is known as *nuclear spin* (I). This is equivalent to the nucleus acting like a miniature bar magnet. The spin of a nucleus depends on the mass of the isotope and nuclei with even mass and even charge numbers have no spin angular momentum ($I=0$). These kinds of nuclei are called “NMR inactive” or “NMR silent” since the nuclear spin property is fundamental to enable NMR. Nuclei as hydrogen (^1H), carbon (^{13}C), fluorine (^{19}F) and phosphorus (^{31}P), have $I = \frac{1}{2}$, thus they can be analysed using NMR.

As showed in Figure 2.5, when the nuclei are not affected by external magnetic field (B_0), the spins are randomly oriented in all direction, while if nuclei with $I \neq 0$ are placed in a magnetic field, they will assume a possible number of different orientations that will correspond to specific energy levels; this number depends on the value of I , in particular it is equal to $2I+1$.

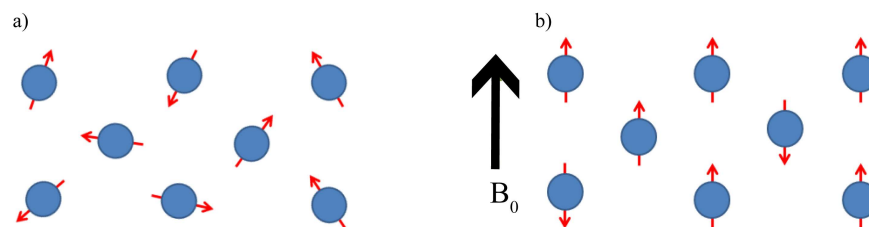


Figure 2.5. a) Randomly oriented spins in the absence of a magnetic field. b) Aligned spins in the presence of an applied magnetic field (B_0)

The proton (^1H) is the most abundant NMR nucleus and it has $I = \frac{1}{2}$, therefore when B_0 is applied, these nuclei can assume two possible orientations, α parallel ($I = \frac{1}{2}$) or β antiparallel ($I = -\frac{1}{2}$), each corresponding to an energy level.

The difference in the energy levels is equal to:

$$\Delta E = \gamma \cdot B_0$$

where γ is the *gyromagnetic ratio* that is constant for a given nucleus ($\gamma = 2.6752 \cdot 10^8/\text{T/s}$ for protons).

This means that, given a nucleus, the stronger the magnetic field the higher the gap in the energy levels, thus resulting in higher sensitivity in the NMR experiment.

The protons are distributed between the two energy states according to the Boltzmann distribution:

$$\frac{N_\alpha}{N_\beta} = \exp\left(\frac{\Delta E}{k \cdot T}\right)$$

Here N_α and N_β represent the protons populations in the lower and upper energy levels, respectively, k is the Boltzmann constant and T is the temperature.

The number of nuclei is not equal in the two states: a small excess of protons will occupy the lower energy state (α) since it is, indeed, the more favorable state from an energetic point of view. This gives rise to a net magnetization M_0 , aligned with the applied magnetic field B_0 .

As showed in Figure 2.6, when a radio frequency pulse is applied, the nuclei will absorb energy and undergo nuclear spin transitions from lower to higher energy levels, according to the following equation:

$$\Delta E = h \cdot \nu$$

where h is the Planck constant and ν is the frequency of the excitation pulse that induces the transition between the levels. This frequency is referred to as *Larmor frequency* and depends both on the nucleus and the magnetic field as showed in the following equation:

$$\nu = \frac{\gamma \cdot B_0}{2 \cdot \pi}$$

No NMR signal is observable when the sample is at equilibrium (in the static magnetic field), since the net magnetization vector has no component on the xy plane where the signal is detected by the detector coil. The duration of the pulse is usually measured in microseconds (μs), and it is called 90° pulse when it brings the magnetization on the xy plane, whereas a 180° pulse flips the net magnetization on the negative z-axis.

As illustrated in Figure 2.6, when the radiofrequency is switched off, the system will return to equilibrium. This return to equilibrium is referred to as relaxation and it causes the NMR signal to decay with time, producing the observed free induction decay (FID). The NMR signal is then Fourier transformed to be converted in the frequency domain.

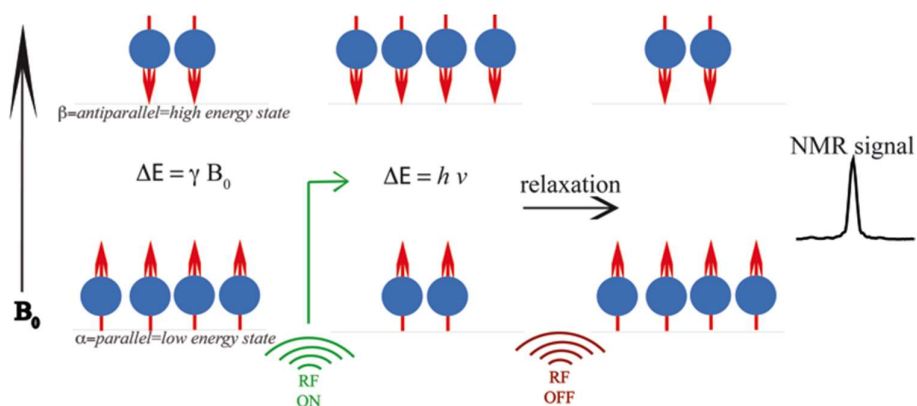


Figure 2.6. Schematization of the NMR experiment.

Protons are situated in different positions in a molecule, thus they do not experience the same magnetic field for a phenomenon known as *shielding*. Indeed, the nucleus is surrounded by electrons that start a rotational motion when an external magnetic field (B_0) is applied. This gives rise to a small local magnetic field (B_{loc}) that may oppose the external field, and, as a

consequence, the nucleus experiences a slightly reduced field, that can be called (B_{eff}). A specific parameter (σ , shielding constant) is used to indicate the density and the distribution of the electronic cloud that surrounds the nucleus. This dimensionless constant ranges from 10^{-6} , for lighter nuclei, to 10^{-3} for heavier ones. The changes in the σ value are given, for example, by the presence of functional groups in the vicinity of a nucleus. In particular, when an electronegative atom is present, it will withdraw electrons from the observed nucleus, reducing the density of the electronic cloud, thus causing a de-shielding effect. Thus, the nucleus will resonate at higher frequencies. Considering that B_{loc} is equal to $B_0 \sigma$ (Lenz rule), B_{eff} is given by:

$$B_{\text{eff}} = B_0 - B_{\text{loc}} = B_0 - B_0\sigma = B_0(1 - \sigma)$$

As a consequence, the relationship between the degree of shielding and the resulting resonance frequency is:

$$\nu = \frac{\gamma \cdot B_0}{2 \cdot \pi} * (1 - \sigma)$$

Therefore, protons situated in different chemical environments will experience different magnetic fields, thus meaning that they will resonate at different frequencies giving rise to the so-called chemical shift (δ). All the information retrievable from a NMR experiment is hidden in the chemical shift, thus it is crucial to only look at the frequencies of the different nuclei in a way that is independent from the magnetic field employed. For this reason, a conventional way to calculate δ has been established:

$$\delta = \frac{\nu \cdot \nu_{\text{ref}}}{\nu_{\text{ref}}} * 10^6$$

where ν is the frequency of the observed nucleus and ν_{ref} is the frequency of a reference compound. The most widely used reference compounds are tetramethylsilane, $\text{Si}(\text{CH}_3)_4$ (simply called TMS) for organic solvents and the sodium salt of trimethylsilyl propionic acid (TSP) for aqueous solutions. Both compounds have maximum shielding, showing higher σ than the nuclei usually analyzed via NMR. Therefore, the previous equation converts the chemical shift frequencies into parts per million (ppm), allowing to display the NMR signals on a new axis on which a given spin always shows the same value independently of the magnetic field employed. In this new system the reference compound has $\delta = 0$, while the sample resonances have positive δ values.

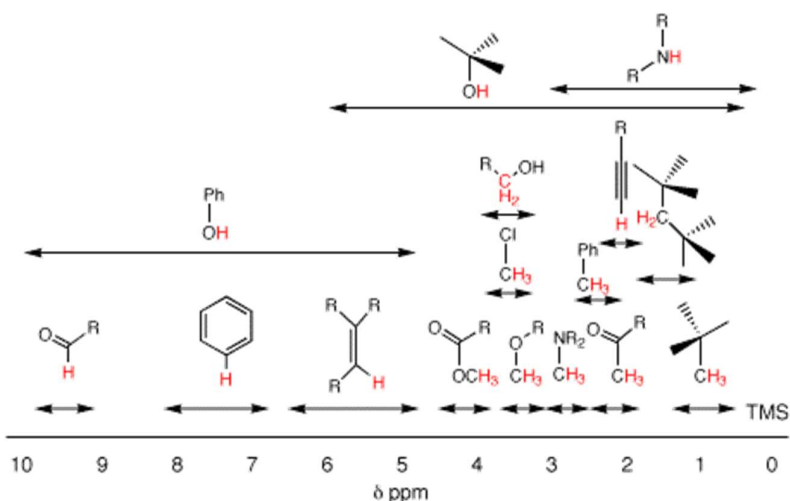


Figure 2.7. Approximate proton chemical shifts.

Chemical shift is not the only value to inspect to retrieve chemical information from the NMR experiment. Other valuable features are the *J coupling* and the *scalar coupling*. This phenomenon is due to the fact that the magnetic moments of the nuclei can influence each other in two possible ways: through space (dipolar coupling) or through chemical bonds (scalar coupling). The effect of the dipolar coupling can be neglected since the interactions average to zero because of the rapid molecular tumbling while the scalar coupling is visible from the NMR spectrum.

Indeed, when an atom is influenced by another, this results in the split of its resonance signal. When the perturbing nucleus is the observed one, it also exhibits signal splitting. In order to observe the splitting of the signal, the interacting nuclei must be bonded in proximity (e.g. vicinal and geminal positions) or be oriented in certain optimal configurations. Usually the signal splitting in proton spectra ranges from fractions of Hz to around 18 Hz.

After the nuclei excitation by means of a radiofrequency pulse, the net magnetization will flip on the *xy* plane (90° pulse). Here the spins start to *precess* (rotate) both clockwise and counter-clockwise, thus, a detection system with two simultaneous channels (reciprocally 90° out of phase), on each axis, is often employed in order to distinguish the spin with a frequency $-\nu$ and one with a frequency $+\nu$. This is called *quadrature detection*.

Therefore, after the pulsed excitation, the nuclei return to the equilibrium distribution between the two energy levels through a process called *relaxation*, with relaxation times that are specific for different nuclei. This process generates the free induction decay (FID). The FID is a time-domain representation of the superimposition of the frequencies of all the nuclei

present in the analyzed sample, thus it is not easy to interpret. *Fourier transform* is a mathematical operation that can be carried out on the final FID data to produce the familiar frequency spectrum (Figure 2.8).

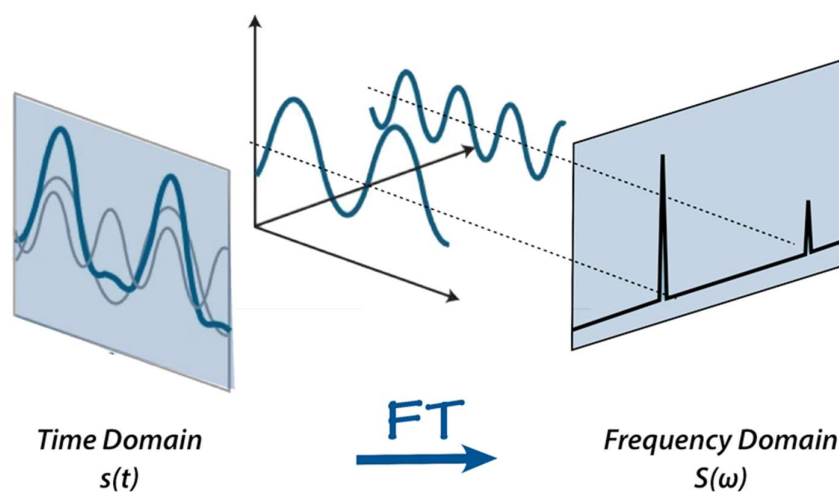


Figure 2.8. Illustration of the Fourier Transformation from the Time Domain $s(t)$ to the Frequency Domain $S(\omega)$.

2.6 Microscale Thermophoresis

MicroScale Thermophoresis (MST) is a powerful technique to quantify biomolecular interactions. It is based on thermophoresis, the directed movement of molecules in a temperature gradient, which strongly depends on a variety of molecular properties such as size, charge, hydration shell or conformation. Thus, this technique is highly sensitive to virtually any change in molecular properties, allowing for a precise quantification of molecular events independent of the size or nature of the investigated compound. During an MST experiment, a temperature gradient is induced by an infrared laser. The directed movement of molecules through the temperature gradient is detected and quantified using either covalently attached or intrinsic fluorophores. By combining the precision of fluorescence detection with the variability and sensitivity of thermophoresis, MST provides a flexible, robust and fast way to dissect molecular interactions.

The phenomenon of thermophoresis was first described by Ludwig and describes the directed motion of molecules through a temperature gradient. In recent years, MST emerged as

a revolutionary method to monitor the directed movement of fluorescent molecules through microscopic temperature gradients in μL -volumes, allowing for the precise analysis of binding events in a few microliter solution of virtually any molecule—independent of size or physical properties (Jerabek-Willemsen et al., 2011; Wienken et al., 2010). Therefore, MST can even detect events such as binding of small molecules to proteins, substrates to enzymes, or ligands to liposomes.

Thermophoresis describes a directed movement of particles in a temperature gradient. A temperature difference ΔT in space leads to a depletion of the solvated biomolecules in the region of elevated temperature, quantified by the Soret coefficient S_T :

$$\frac{C_{hot}}{C_{cold}} = \exp(-S_T \Delta T)$$

This thermophoretic depletion depends on the interface between molecule and solvent. Under constant buffer conditions, thermophoresis probes the size, charge and solvation entropy of the molecules. Typically, the thermophoresis of a protein significantly differs from the thermophoresis of a protein–ligand complex due to binding-induced changes in size, charge and solvation energy. Even if ligand binding could not significantly affect the size or charge of a protein, MST can still detect it, because of binding-induced changes in the molecules solvation entropy (Seidel et al., 2013).

Figures 2.9 A and B show the Monolith NT.115 instrument from NanoTemper Technologies GmbH, the capillary tray that can accommodate up to 16 capillaries, and the schematic representation of MST optics. MST is measured in capillaries with a total volume of $\sim 4 \mu\text{L}$. The fluorescence within the capillary is excited and detected through the same objective. A focused IR-Laser is used to locally heat a defined sample volume. Thermophoresis of fluorescent molecules through the temperature gradient is detected and it allows to get the signal of an MST experiment (Figure 2.9 C). Initially, the molecules are homogeneously distributed and a constant “initial fluorescence” is detected. Within the first second after activation of the IR laser, the “T-Jump” is observed, which corresponds to a rapid change in fluorophore properties due to the fast temperature change. Subsequently, thermophoretic movement of the fluorescently-labeled molecules out of the heated sample volume can be detected. Typically, the fluorescence change is measured for 30 s. After deactivation of the IR-Laser, an inverse T-Jump occurs, followed by the “back-diffusion” of molecules, which is solely driven by mass diffusion.

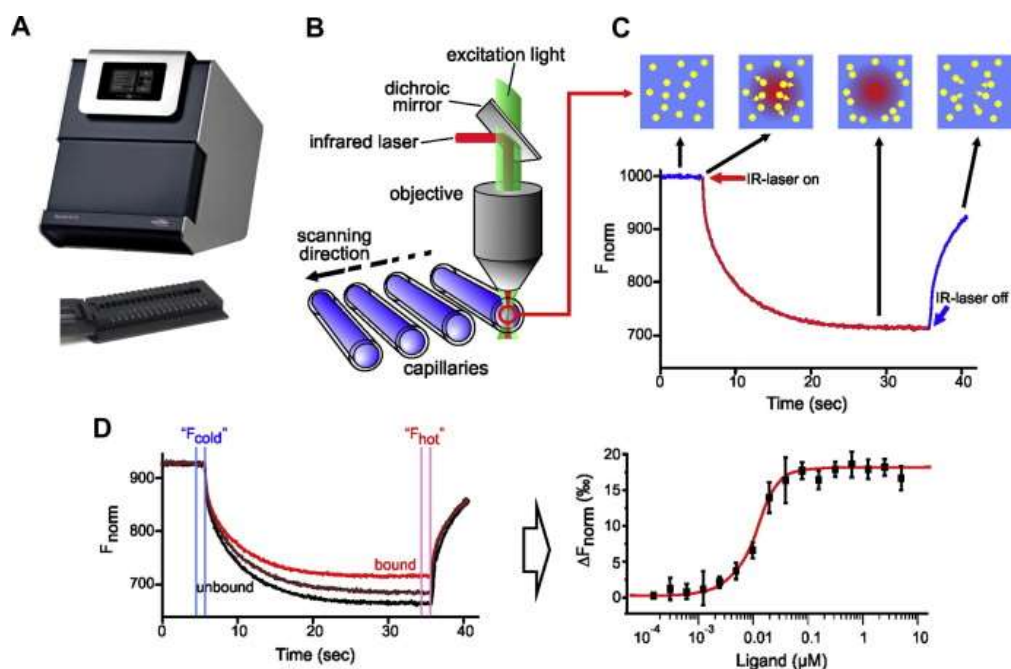


Figure 2.9. MST setup and experiments [adapted from (Jerabek-Willemsen et al., 2014)]

In Figure 2.9 D is represented a typical binding experiment in which the thermophoretic movement of a fluorescent molecule (black trace; “unbound”) changes upon binding to a non-fluorescent ligand (red trace; “bound”), resulting in different traces. For analysis, the change in thermophoresis is expressed as the change in the normalized fluorescence (ΔF_{norm}), which is defined as $F_{\text{hot}}/F_{\text{cold}}$ (F -values correspond to average fluorescence values between defined areas marked by the red and blue cursors, respectively). Titration of the non-fluorescent ligand results in a gradual change in thermophoresis, which is plotted as ΔF_{norm} to yield a binding curve, which can be fitted to derive binding constants.

In recent years, MST was successfully employed to analyze a large variety of bio-molecular interactions, ranging from oligonucleotide interactions (Baaske et al., 2010; Seidel et al., 2012), protein–DNA interactions (Strauß et al., 2012; Zillner et al., 2013) and protein–protein interactions (Arbel et al., 2012; Keren-Kaplan et al., 2013; Lin et al., 2012; Wilson et al., 2012; Xiong et al., 2013) to protein–small molecule (Gaffarogullari et al., 2013; Patnaik et al., 2012; Seidel et al., 2012; Wienken et al., 2010) and protein–liposome interactions (Van Bogaart et al., 2012; Van Den Bogaart et al., 2011). Besides the well-established capability of MST to accurately monitor molecular interactions and to derive dissociation constants, thermophoresis can also be used to obtain a number of other thermodynamic parameters (ΔH , ΔS , and ΔG), allowing the complete characterization of the interaction.

Chapter 3

TARGETING G-QUADRUPLEXES

Molecular recognition is a very important phenomenon in biology and chemistry, and the molecular basis of interaction between small molecules and biological macromolecules is the subject of numerous investigations aimed at the rational design of compounds with specific biological activities.

During the last years, after the identification of G4 motifs as biologically relevant elements, a certain number of studies about ligand/DNA G4 interactions have been reported (Amato et al., 2016; Ou et al., 2008). Since the formation of G4 structures by telomeric DNA inhibits the activity of telomerase and their ligand-induced stabilization leads to cellular senescence, small molecules that bind and stabilize such G4s could potentially be effective chemotherapeutic agents (Neidle, 2009), as well as compounds able to recognize oncogene promoter G4s thus suppressing their transcription. In this scenario, the identification of new ligands that are specific for the G4 structures is a promising approach to develop new anticancer drugs. Despite the fact that the G4 structures differ considerably from the double helix, the design of selective ligands is challenging (Burge et al., 2006). Nevertheless, a number of G4 binding agents has been proposed so far, and some of these have been demonstrated to have effective anticancer activity (Müller and Rodriguez, 2014).

Therefore, in this Chapter, I will present some studies about the discovery of new G4 binders endowed with biological activity.

In Paper I, starting from a promising lead, a small series of novel hydrazone-based compounds were synthesized and evaluated as G4 binders. The *in vitro* G4-binding properties of the synthesized compounds were investigated employing both human telomeric and oncogene promoter G4s with different folding topologies as targets. Such investigation led to the identification of potent G4 stabilizers with high selectivity over duplex DNA and preference for one G4 topology over others. Among them, selected derivatives have been shown to trap

G4 structures in the nucleus of cancer cells. Interestingly, this behavior correlated with efficient cytotoxic activity in human osteosarcoma and colon carcinoma cells (Amato et al., 2016).

In Paper II, the possibility of synthesizing through a multicomponent reaction a new series of furo[2,3-*d*]pyridazin-4(5*H*)-one (FP) derivatives as selective binders of *BCL2* oncogene promoter has been exploited. Biophysical studies showed that two derivatives could effectively bind to bcl-2 G4 with good selectivity. Moreover, one such ligand was found to appreciably inhibit *BCL2* gene transcription, with a substantial decrease in protein expression levels, and also showed significant cytotoxicity toward the Jurkat human T-lymphoblastoid cell line (Amato et al., 2018).

During my PhD, I also contributed to study a library of analogs of a lead-like G4 targeting compound (**4c**), sharing a furobenzoxazine naphthoquinone core and differing for the pendant groups on the N-atom of the oxazine ring, with the aim of developing more potent and selective ligands (Paper III) (Amato et al., 2019). These molecules have been tested vs. topologically different G4s by the G4-CPG (Controlled Pore Glass) assay, an affinity chromatography-based method for screening putative G4 ligands. The obtained results showed that all these compounds were able to bind several G4 structures, both telomeric and extra-telomeric, thus behaving as multi-target ligands, and two of them fully discriminated G4 vs. duplex DNA. Biological assays proved that almost all the compounds produced effective DNA damage, showing marked antiproliferative effects on tumor cells in the low μM range. Combined analysis of the G4-CPG binding assays and biological data led us to focus on a compound (**S4-5**), proved to be less cytotoxic than the parent compound **4c** on normal cells. An in-depth biophysical characterization of the binding of **S4-5** to different G4s showed that the selected ligand has higher affinity for the G4s and higher ability to discriminate G4 vs. duplex DNA than **4c**. Molecular docking studies, in agreement with the NMR data, suggested that **S4-5** interacts with the accessible grooves of the target G4 structures, giving clues for its increased G4 vs. duplex selectivity.

3.1 New Hydrazone derivatives as specific G-quadruplex binders (*Paper I*)

3.1.1 Introduction

Many research groups have attempted to identify bioactive G4-binding compounds and, to date, over 1000 ligands have been reported, representing a wide range of chemotypes (Monchaud and Teulade-Fichou, 2008; Ohnmacht and Neidle, 2014). Some of these have been identified in order to target the grooves of G4 structures, although most have been devised to primarily stack on the external G-tetrads of G4s. The structural features that appear to facilitate the latter binding mode are: (i) the presence of an extended planar aromatic scaffold that can efficiently stack on terminal G-tetrads through π - π stacking, (ii) an electron-deficient aromatic core that can enhance stacking interactions on G-tetrads via cation- π interaction, and (iii) the presence of cationic side-chains. These positively charged tethered substituents can interact with the grooves and loops of a G4 and with the negatively charged phosphate backbone.

In 2010, the bis-guanylhyazone derivative of diimidazo[1,2-*a*:1,2-*c*]pyrimidine **1a** (in order to avoid confusion, compounds **1-10** in Paper I here appear as **1a-10a**) (Figure 3.1) was reported to be a potent stabilizer of several G4 DNAs, whereas no significant interaction with duplex DNA was found (Sparapani et al., 2010). Its high affinity and selective G4-stabilizing ability led us to consider **1a** as a promising lead compound and to design a small series of analogues. For this reason, in collaboration with Dr. Rita Morigi and her research group, (Department of Pharmacy and Biotechnology, University of Bologna, Italy) the lead compound was chemically modified in order to obtain new derivatives to be tested as G4 ligands (Amato et al., 2016). Structural modifications of **1a** involved in particular the guanyl chains, the substituents at positions 2 and 8 and the core chromophore itself (Figure 3.1).

The guanyl groups have been modified by introducing a more rigid frame (**2a**) or an aromatic system (**4a**). Together with compound **2a**, the monohydrazone **3a** was isolated from the reaction mixture and has also been investigated. The phenyl rings at positions 2 and 8 have been replaced with the bioisostere thiophene to obtain compound **5a**. The tricyclic core has been substituted with different systems: the pyrroloindole nucleus to obtain compound **6a** or a chloroindole moiety (**7a**, **8a**), to achieve core simplification, and finally, core expansion was explored by introducing the planar bisindole system of indigo, which was functionalized with iminoguanyl or hydrazinylpyridyl groups in order to obtain compounds **9a** and **10a**.

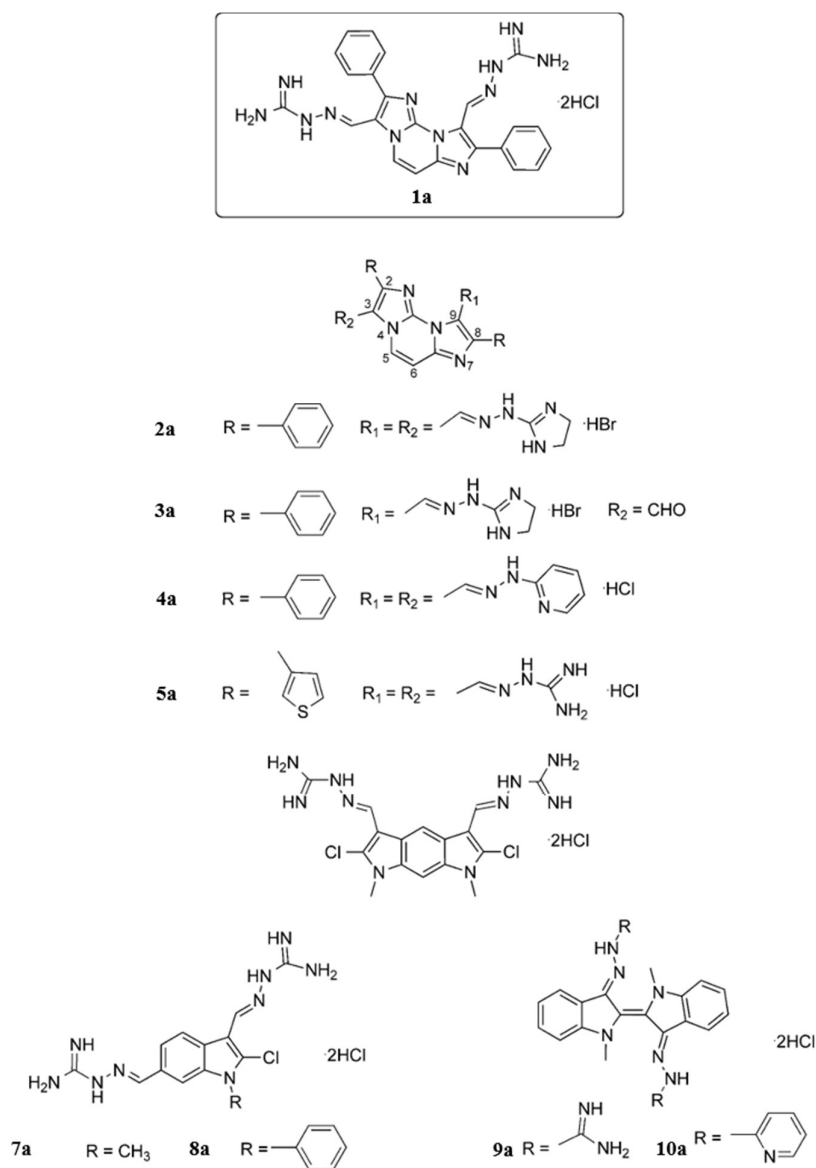


Figure 3.1. Compound lead 1a and structural modifications (in order to avoid confusion, compounds 1-10 in Paper I here appear as 1a-10a).

3.1.2 Results and discussion

Circular Dichroism studies

To evaluate the interaction of compounds **1a–10a** with G4 structures, several diverse G4-forming sequences able to form parallel, antiparallel, and hybrid G4 structures were selected for this study. It is known that human telomeric DNA G4s can adopt different topologies depending on the selected sequence and experimental conditions (Dai et al., 2008). In this study,

two human telomeric DNA sequences, namely Tel23 and Tel26, were used. They are known to form the so-called hybrid-1 and hybrid-2 folds, respectively. In addition, because several studies suggest the parallel G4 fold (Parkinson et al., 2002) as the prevalent one in the overcrowded solution conditions present inside a cell (Xue et al., 2007), we also prepared a Tel23 sample at high DNA concentration conditions in order to promote the parallel G4 conformation (hereafter referred to as Tel23-p). Further, to determine the affinity of the ligands toward other G4 structures, the two G4-forming sequences from the nuclease hypersensitive region of the *KIT* promoter (c-kit1 and c-kit2) and one from the *MYC* promoter (c-myc) were also used. The structures adopted by each G4 sample were verified by circular dichroism (CD) spectroscopy. Tel23-p, c-kit1, c-kit2, and c-myc displayed a positive band at 264 nm and a negative one around 240 nm in the CD spectrum (Figure 3.2). These bands are characteristic of parallel-stranded G4 topologies (Karsisiotis et al., 2011; Masiero et al., 2010). On the other hand, Tel23 and Tel26 showed very similar CD spectra having a positive band at 289 with a shoulder at ca. 268 nm and a weak negative band at around 240 nm (Figure 3.2). These data are consistent with the presence of hybrid structures as major conformations. CD experiments were also performed to examine the potential of compounds **1a–10a** to alter the native folding topology of these G4s. Thus, DNA/ligand complexes were obtained by adding ligands (4 mol equivalents) to the folded G4 structures. No significant variations of CD signal were observed for any of the structures analyzed here (Figure 3.2), suggesting an overall preservation of their G4 architectures upon the addition of each ligand. The structure of a duplex DNA (ds12), which is formed by a self-complementary 12-nt sequence, was also verified by CD spectroscopy because it has been used in subsequent experiments to evaluate the selectivity of the ligands for G4 over duplex DNA. In the presence of K^+ , the CD spectrum of ds12 was characterized by a positive band at around 280 nm and a negative one at 250 nm, typical of values observed for duplex DNA (Figure 3.2). These bands were not modified upon ligand addition.

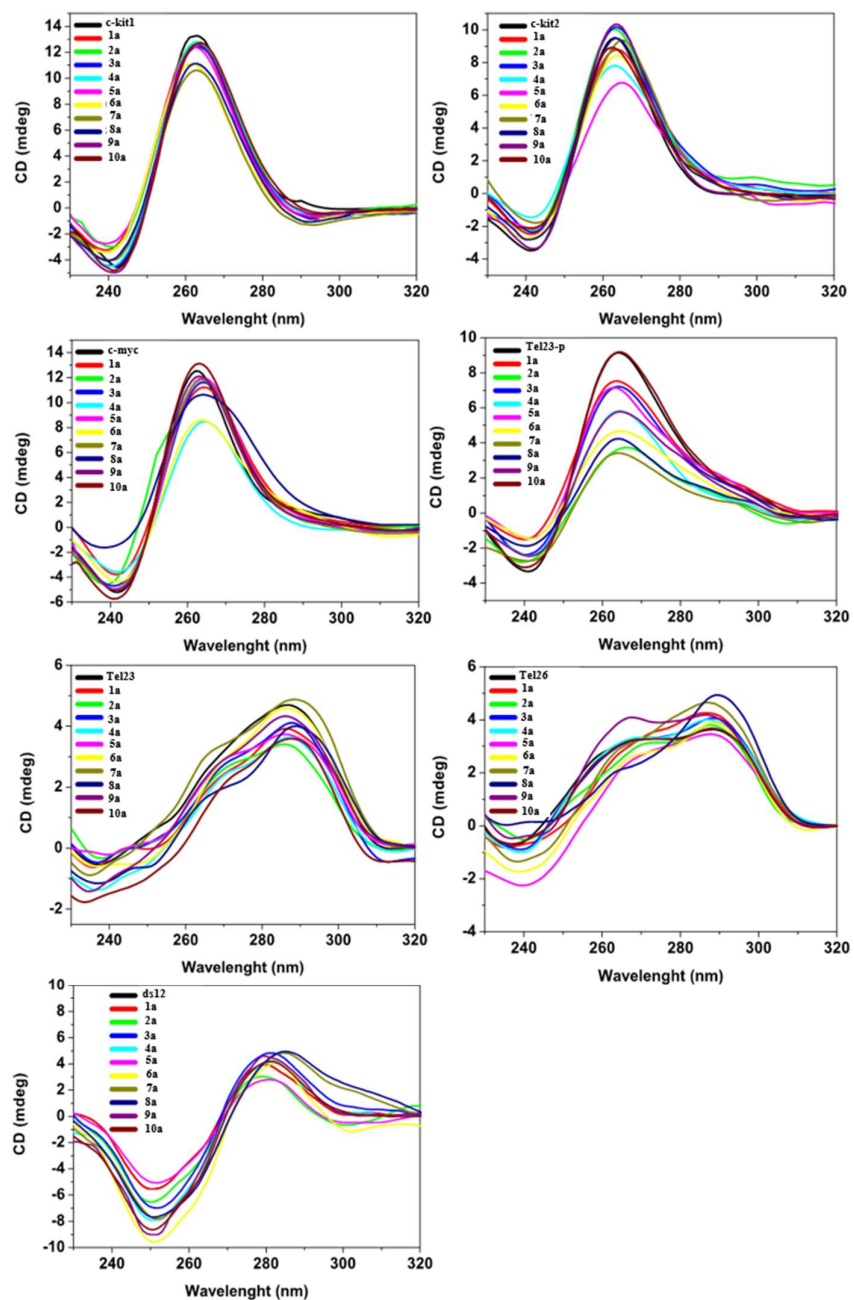


Figure 3.2. CD spectra of the investigated DNAs in the absence and in presence of 4 mol equivalents of compounds 1a-10a.

The stabilizing properties of the compounds were evaluated by CD-melting experiments measuring the ligand-induced change in the melting temperature (ΔT_m) of G4 and duplex structures. CD-melting curves of DNAs in the absence and presence of each ligand were

obtained by following the variations of the intensities of CD signals at the wavelength of 264, 290, and 280 nm for parallel G4s, antiparallel G4s, and duplex, respectively (Table 3.1).

Table 3.1. Ligand-induced thermal stabilization of G4 and duplex DNAs measured by CD melting experiments.

compd	ΔT_m (°C) ^a						
	c-kit1 ^b	c-kit2 ^c	c-myc ^c	Tel23-p ^b	Tel23 ^b	Tel26 ^b	ds12 ^b
1a^d	>15.0	>20.0	>20.0	>20.0	-10.0	-4.5	0.1
2a	1.6	>20.0	>15.0	10.8	-14.0	-8.9	1.3
3a	2.7	9.5	>20.0	9.1	-2.4	-3.0	-0.8
4a	0.4	0.0	-0.2	1.3	-1.0	-1.6	0.9
5a	>15.0	>20.0	>20.0	>20.0	-4.9	-6.5	3.5
6a	>15.0	>15.0	>20.0	>20.0	3.0	1.3	ND ^e
7a	>15.0	>20.0	>20.0	>20.0	6.0	5.8	ND ^e
8a	>15.0	>15.0	>20.0	>20.0	5.6	5.1	ND ^e
9a	0.9	8.2	11.0	7.7	-0.5	0.2	-0.8
10a	1.5	2.0	1.6	4.0	-2.5	-2.0	-0.7

^a ΔT_m represents the difference in melting temperature [$\Delta T_m = T_{m(\text{DNA}+4 \text{ ligand equiv})} - T_{m(\text{DNA})}$]. The T_m values of DNAs alone are c-kit1 = 69.8 ± 0.5 °C, c-kit2 = 62.5 ± 0.5 °C, c-myc = 72.9 ± 0.5 °C, Tel23-p = 68.5 ± 0.5 °C, Tel23 = 65.2 ± 0.5 °C, Tel26 = 59.5 ± 0.5 °C, ds12 = 71.3 ± 0.5 °C. All experiments were duplicated, and the values reported are average of two measurements. ^b100 mM KCl buffer. ^c20 mM KCl buffer. ^dLead compound. ^eThese compounds significantly increase the thermal stability of ds12, however, ΔT_m values cannot be accurately determined.

For the first CD-melting experiments, a 4:1 ligand/DNA ratio was used so that all possible G4 binding sites were available. Results of these experiments are reported in Table 3.1. On the basis of these data, ligands have been ranked in the following categories: (i) compounds that exhibit poor G4-stabilizing properties toward any G4 ($\Delta T_m < 5$ °C) (compounds **4a** and **10a**), (ii) compounds showing medium G4-stabilizing effect ($\Delta T_m = 5\text{--}15$ °C) (compound **9a**), and (iii) potent G4 stabilizers ($\Delta T_m > 15$ °C) (compounds **2a**, **3a**, **5a–8a**). As with the lead compound **1a**, most of the compounds investigated here showed high selectivity for G4 over duplex DNA. However, compounds **6a**, **7a**, and **8a**, having a different core from compound **1a**, were the only ligands to significantly increase duplex stability and therefore were not considered to be selective. Interestingly, compounds **2a**, **3a**, and **5a** showed some preference for parallel G4s over antiparallel ones. In particular, compound **3a** stabilized to a large extent only the c-myc G4 ($\Delta T_m > 20$ °C). Because we were mainly interested in finding G4 ligands

specific for one topology or, even better, for one structure, further studies were only performed on compounds **2a**, **3a**, and **5a**. Therefore, we carried out CD melting experiments to investigate the effect of increasing concentrations of **2a**, **3a**, and **5a** on the stability of both c-kit2 and c-myc G4s (Figure 3.3). Three different [ligand]/[DNA] ratio were examined (1:1, 2:1, and 4:1). In the case of c-kit2, we observed a dose dependent increase of thermal stability for all three compounds although with some differences between them. In particular, compounds **2a** and **5a** increased the T_m of the G4 with every addition of ligand, while **3a**, which had shown a modest increase in c-kit2 stability, did not induce any further increase of T_m beyond the 2:1 [ligand]/[DNA] ratio. In the case of c-myc, increasing thermal stability of the G4 with compound **3a** also ceased at a 2:1 ratio but after a considerable increase of T_m . On the other hand, concentration dependency was observed for **2a** (although slight compared to c-kit2), while the high G4-stabilizing effect of **5a** on c-myc did not enable us to assess any dose-dependent effects. Overall, these results indicate that six out of nine synthesized ligands are potent G4 stabilizers. The diimidazo[1,2-*a*:1,2-*c*]pyrimidine derivatives having a positive charge on the side chain(s) in the used experimental conditions (i.e., **2a**, **3a**, and **5a**) showed high affinity toward some of the investigated G4 structures. Conversely, compound **4a**, whose side chain protonation is less favorable at pH 7.0, showed generally poor G4-stabilizing properties. In any case, beyond the overall charge of these ligands, their duplex-binding activity is negligible. Concerning the compounds with a different core, the positively charged ones (**6a**, **7a**, and **8a**) showed affinity for some of the G4s investigated as well as for duplex DNA. On the other hand, compounds **9a** and **10a**, whose side chain protonation is less likely to occur at pH 7.0, showed medium to low G4-stabilizing properties and negligible effects on duplex. In summary, these results indicate that the overall charge of compounds is important in DNA binding but, at the same time, suggest that the charge is necessary but not sufficient to ensure potent and selective interaction with a specific DNA structure. In addition, CD melting data suggest that two molecules of **3a** could bind and stabilize the G4 target (the addition of more than two equivalents of ligand does not lead to a further enhancement of G4 stability), possibly interacting in an end-stacking mode at the terminal G-tetrads of the G4, a common binding mode exhibited by the majority of effective G4-targeting ligands. It has to be noted that a significant destabilizing effect on Tel23 and Tel26 G4s occurs with some ligands. In particular, this phenomenon occurs with compounds **1a**, **2a**, and **5a**, namely the bishydrazone derivatives of diimidazo[1,2-*a*:1,2-*c*]pyrimidine. The mechanism of destabilization of these G4 structures

by such derivatives is unclear and would need further study. However, we hypothesize that it may involve nonspecific interactions of compounds with the DNA molecule that might favor single stranded DNA, resulting in a shift of folded-unfolded equilibrium toward the unfolded form during the melting experiment, which, in turn, may result in a decrease of T_m .

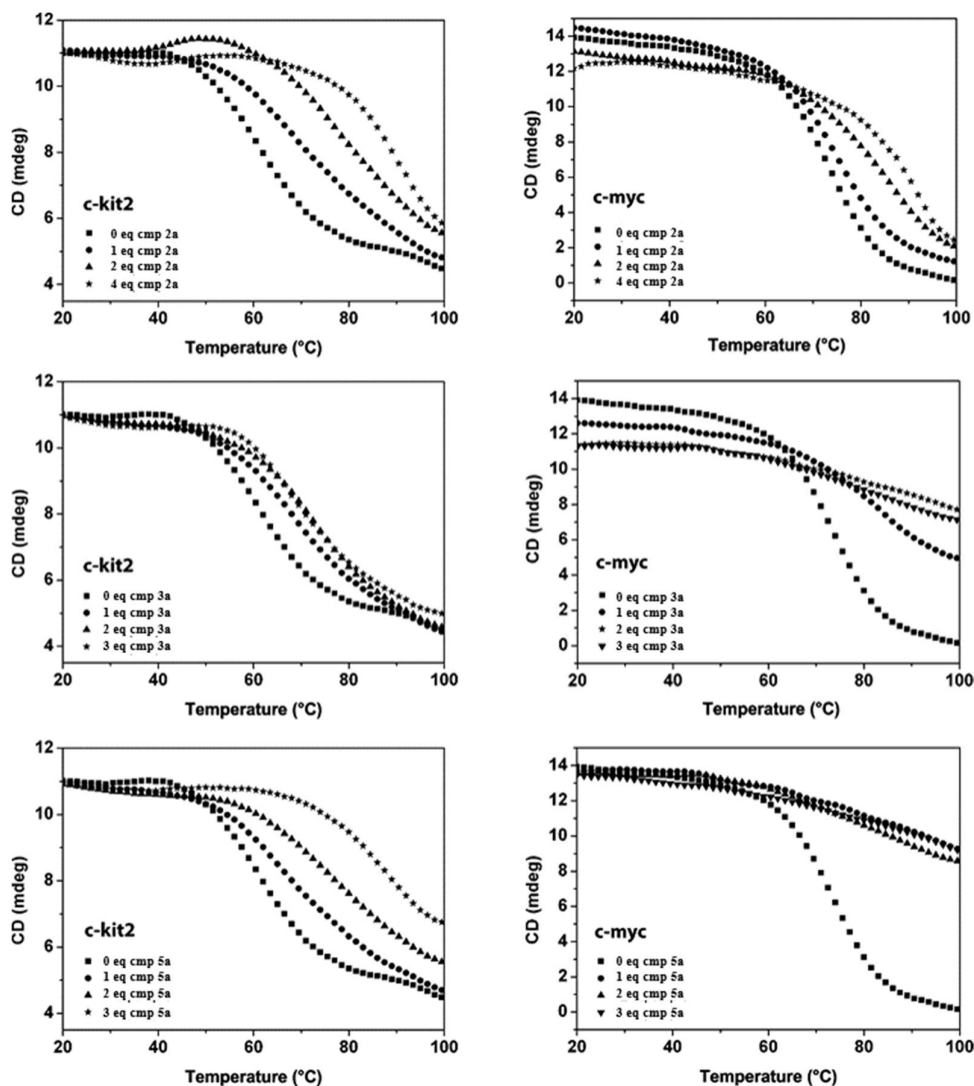


Figure 3.3. CD melting experiments of c-kit2 (left column) and c-myc (right column) G4s with increasing amounts of compounds 2a, 3a, and 5a.

CD has been also used to investigate the ability of the best ligands (2a, 3a, and 5a) to induce G4 formation from unfolded single-stranded DNAs by performing CD titration experiments in the absence of metal ions. In particular, the sequences c-kit1, c-kit2, Tel23, and c-myc in 50 mM Tris buffer were investigated. The CD spectra of c-kit1 and c-kit2 sequences in the absence

of monovalent metal cations showed a positive band at around 262 nm and a negative one around 240 nm that, even if small, are indicative of the presence of a small percentage of a parallel G4 structure in solution. Upon addition of ligands to G4s, a dose-dependent increase of the CD bands was observed in some case, suggesting the induction of the parallel G4 structure formation. In particular, as far as c-kit2 is concerned, compounds **2a** and **5a** caused a major increase of ellipticity compared to **3a** (Figure 3.4). Similarly, the addition of **2a** and **5a** to the c-kit1 sequence also caused a dose-dependent increase in CD signals (more intense for **5a**), whereas the addition of **3a** did not affect the CD profile, thus indicating that **3a** is unable to induce c-kit1 G4 formation. The c-myc G4 DNA sequence exists in a parallel topology, in the absence of any added monovalent metal cation, showing a positive band at 263 nm and a negative one around 240 nm. The addition of ligands **2a**, **3a**, and **5a** to c-myc G4 did not perturb the parallel G4 structure already formed in Tris buffer. On the other hand, the addition of compounds **2a**, **3a**, and **5a** to the Tel23 sequence did not lead to any increase in CD ellipticity, thus suggesting that none of them is able to induce any telomeric G4 structures.

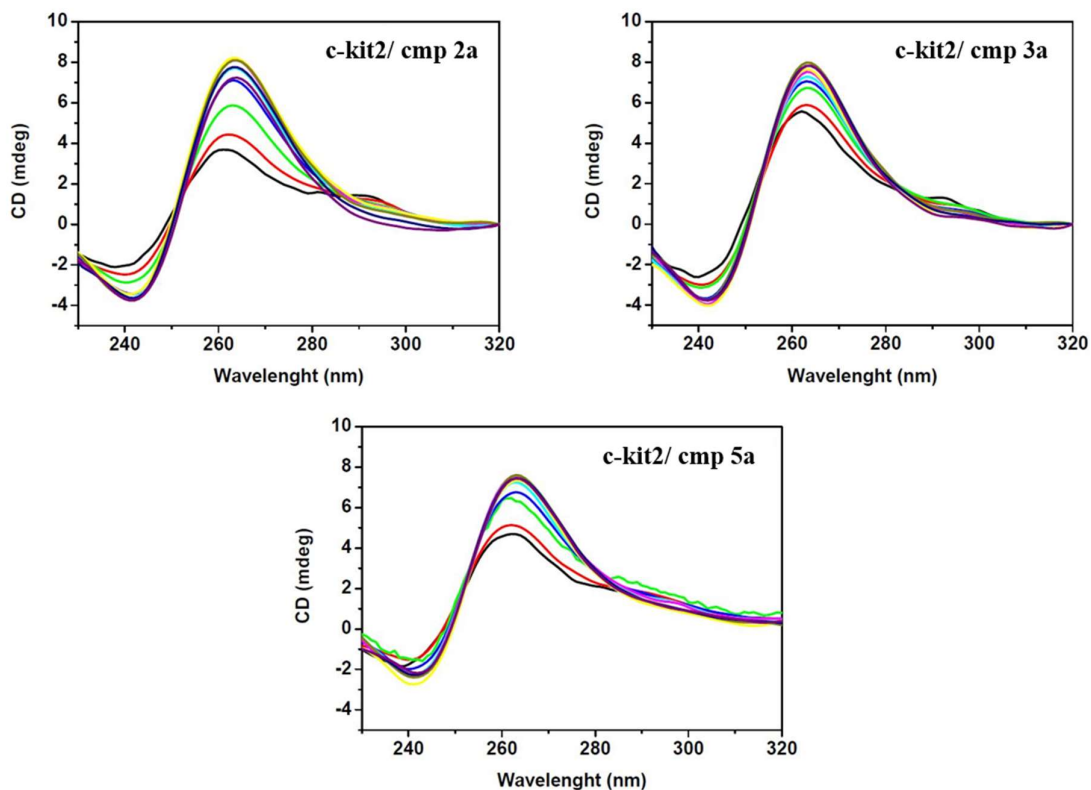


Figure 3.4 CD titration spectra of c-kit2 with compounds **2a**, **3a** and **5a** in the absence of metal cations (50 mM Tris buffer).

Fluorescence resonance energy transfer (FRET) melting studies

The FRET methodology was also used to evaluate the investigated compounds and to estimate their G4 vs duplex selectivity (Mergny and Maurizot, 2001). The compounds were first screened with the derivatized human telomeric G4 sequence F21T and a duplex DNA sequence (T-loop) by using various concentrations of ligands. Similarly to the lead **1a**, compounds **2a**, **3a**, and **5a–8a** showed potent G4 stabilizing capabilities, while compounds **4a**, **9a**, and **10a** produced no significant change in G4 melting temperature (Figure 3.5). Among the effective G4 stabilizers, compounds **6a**, **7a**, and **8a** showed a significant degree of duplex stabilizing ability, confirming once again their low selectivity for G4 over duplex DNA. On the other hand, FRET results also confirmed that compounds **2a**, **3a**, and **5a** are effective and selective stabilizers of the G4 structures.

However, the results of these FRET melting experiments cannot be directly compared with those obtained from CD melting studies because of differences in DNA sequences and in experimental conditions. Therefore, to more fully compare the results of CD melting experiments obtained for compounds **2a**, **3a**, and **5a**, additional FRET melting experiments were performed on these ligands under the same buffer conditions used for CD, using as targets the labeled G4-forming oligonucleotides from the *KIT* (Fckit2T) and *MYC* (FcmcyT) promoter regions and the telomeric G4-forming sequence annealed at high concentration to promote the parallel conformation (F21T-p) (Figure 3.5).

The results of FRET melting experiments, shown in Table 3.2, are in good qualitative agreement with those obtained from CD melting studies. The FRET ranking order reflects that compared to the CD melting data; the fact that the ΔT_m values themselves are somewhat different is likely to be a consequence of different techniques and does not have particular significance.

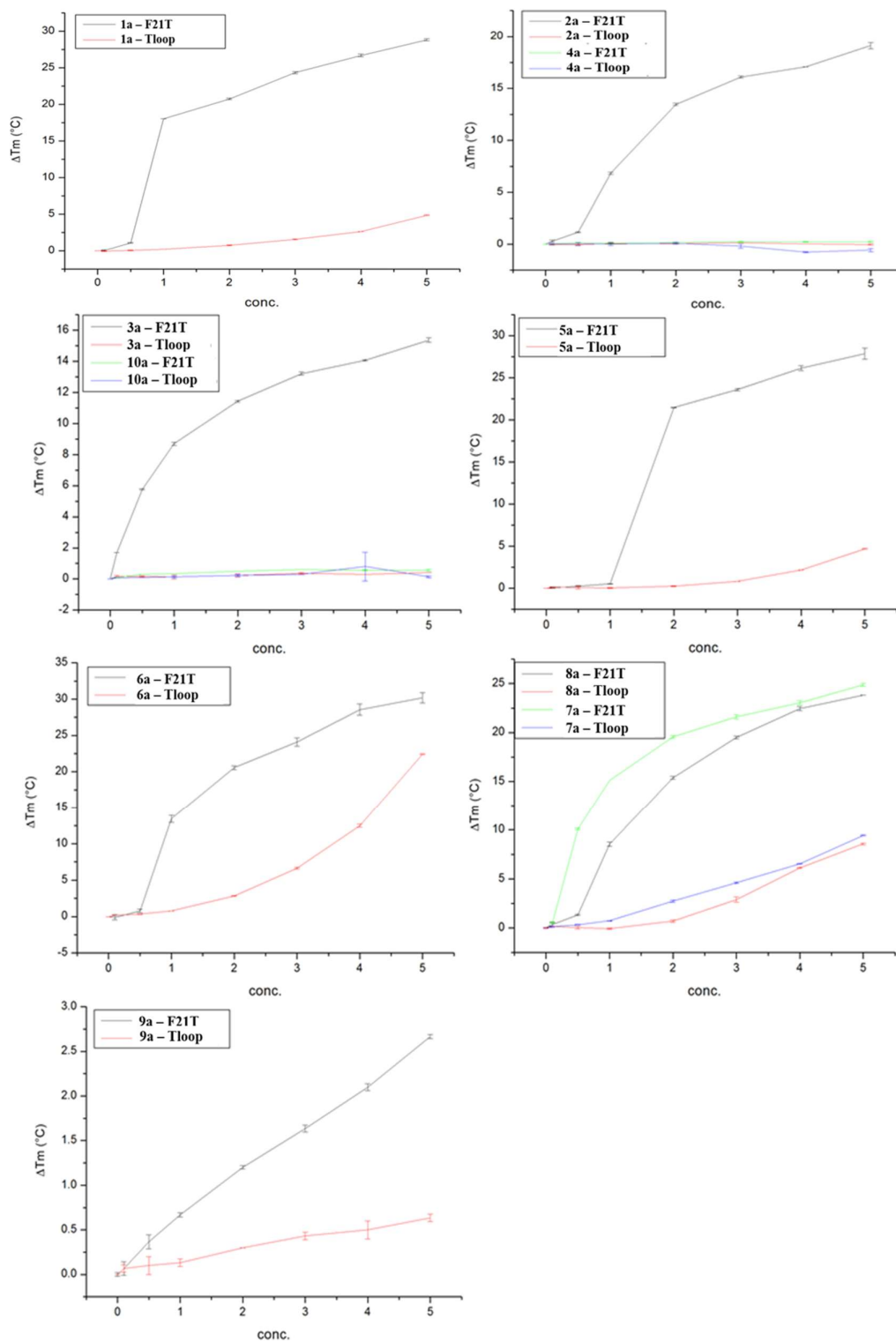


Figure 3.5. Thermal stabilization (ΔT_m) of F21T G4 and hairpin ds-DNA (T-loop) induced by the investigated ligands (1a-10a) at various concentrations (μM) measured by FRET-melting experiments.

Table 3.2. Ligand-induced thermal stabilization of labeled G4-forming sequences evaluated by FRET melting analyses.

compd	ΔT_m (°C) ^a		
	Fc-kit2T ^b	Fc-mycT ^b	F21T-p ^c
2a	>20.0	>15.0	>15.0
3a	≈15.0	>20.0	7.0
5a	>20.0	>20.0	>20.0

^a ΔT_m represents the difference in melting temperature [$\Delta T_m = T_m(\text{G4+4 ligand equiv}) - T_m(\text{G4})$]. T_m values of G4s alone: Fc-kit2T = 66.3 ± 1.0 °C, Fc-mycT = 68.6 ± 1.0 °C, F21T-p = 67.1 ± 1.0 °C All experiments were duplicated, and the values reported are average of two measurements. ^b20 mM KCl buffer. ^c100 mM KCl buffer.

Nuclear Magnetic Resonance studies

NMR spectroscopy was employed in order to obtain structural information about the binding mode of compound **3a**. As stated above, this compound is a potent and specific stabilizer of c-myc G4. Because the wild-type G4 sequence of c-myc is not suitable for study by NMR (its ¹H NMR spectrum is affected by the presence of minor conformations), the well-characterized modified sequence (myc22) containing two G-to-T mutations at position 14 and 23 was used for the NMR titration experiments. The spectral regions of the imino and aromatic protons in the absence and presence of increasing amounts of compound **3a** are shown in Figure 3.6A. According to the literature, under the experimental conditions used, the myc22 DNA sequence forms a single G4 conformation characterized by 12 well-resolved imino protons peaks, corresponding to the 12 guanines involved in the three G-tetrad planes (Ambrus et al., 2005). Upon addition of compound **3a** to the myc22 solution, both imino and aromatic protons became broad even at higher drug equivalence, indicating a dynamic binding process of **3a** to myc22 with a medium exchange rate on the NMR time scale. The titration was virtually complete at a ligand/DNA ratio of 2:1. Major chemical shift changes were observed for the imino protons belonging the 5' G-tetrad (G7–G11–G16–G20) as well as for G9 of the 3' G-tetrad (G9–G13–G18–G22), while the remaining imino protons of G13, G18, and G22 at the 3'-end tetrad and those of middle tetrad (G8–G12–G17–G21) were less affected. The origin of this different behavior could be due to inherent structural features associated with the 3'- and 5'-faces as well as the flanking sequences. The 5' side is more accessible for ligand stacking and the only 5'-

flanking residue particularly affected by the binding turned out to be the T4, as shown by the corresponding aromatic proton shift. By contrast, the 3'-face is less accessible. A25 from this side folds back to form a base pair with T23 to cover the external G-tetrad, while A24 stays above the T23:A25 base pair. However, these latter flanking residues are not able to completely prevent the binding. Upon addition of **3a**, the aromatic protons of T23, A24, and A25 were strongly perturbed, suggesting that the ligand can be positioned between the four guanines of the 3'-tetrad and the 3'-flanking bases. The stability of the 3'-end bound ligand complex with myc22 could be related to the possibility of specific H-bond interactions between the imidazole group of **3a** and the 3'-flanking bases (as discussed in the docking section). For comparison, we have also examined the binding behavior of **3a** to the c-kit2 G4. Because the c-kit2 G4 wild-type sequence gives poorly resolved imino proton NMR spectra, we used the modified sequence c-kit2_{T12/T21} (in which G12 and G21 are replaced by two T residues), which is known to fold into a monomeric parallel G4 structure with well-resolved imino proton signals (Hsu et al., 2009; Kuryavyi et al., 2010). Figure 3.6B shows the spectral regions of the imino and aromatic protons of c-kit2_{T12/T21} in the absence and presence of increasing amounts of compound **3a**. Interestingly, the imino signals belonging to the G-tetrad on the 5'-side (G2, G6, G14, and G18) were equally affected with respect to those belonging to the 3'-terminal G-tetrad, as well as the aromatic protons of C1 and A13 on the 5'-face and those of C9 and T21 on the 3'-face. These results clearly suggest an involvement of both external G-tetrads in binding, despite the potential pairing between residues C1 and A13 on the 5'-side (Kuryavyi et al., 2010). Thus, we hypothesize that **3a** is able to interact with both the 5' and 3' G-tetrads, possibly by displacing the C1 and A13 base pairing at the former. Overall, these findings suggest that compound **3a** binds to the 5'- and 3'-terminal G-tetrads of both myc22 and c-kit2_{T12/T21} G4s. However, some differences in the binding to the two G4s were observed that could explain the difference in the thermal stabilization of the two G4s upon binding observed by CD analysis. Particularly, we speculate that this different binding behavior is due to the possibility of **3a** forming specific H-bond interactions with myc22 (see Molecular Docking section), thus ensuring that compound **3a** binds more effectively to the c-myc G4 rather than to the c-kit2 G4.

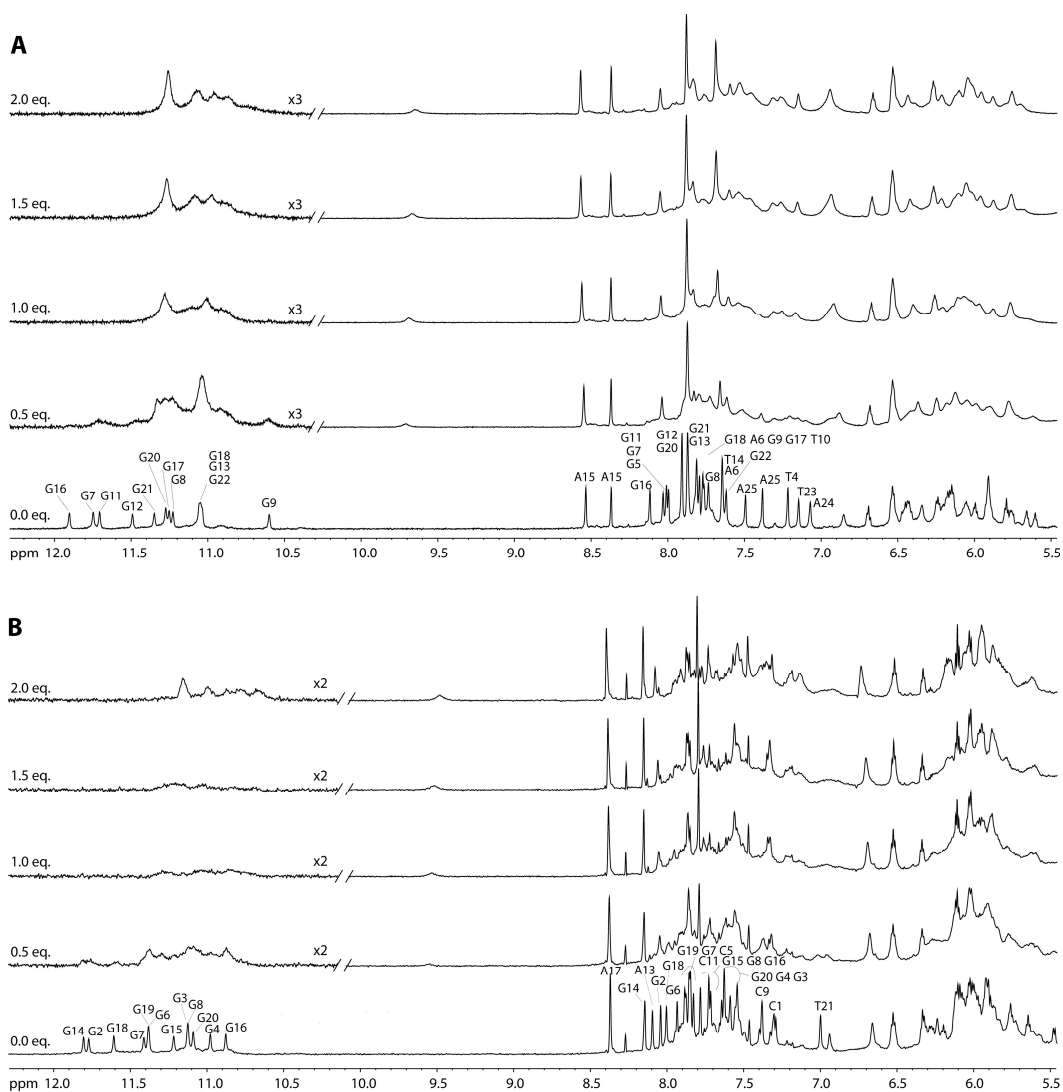


Figure 3.6. Imino and aromatic proton regions of (A) *MYC* promoter G4 (*myc22*) in 25 mM phosphate, 70 mM KCl solution (pH 7.0), and (B) *KIT* promoter G4 (*c-kit2*_{T12/T21}) in 5 mM phosphate, 20 mM KCl solution (pH 6.8) titrated with compound **3a**.

Molecular Docking

Molecular docking calculations were performed to provide a possible molecular basis for the G4 binding of **3a**. We docked the ligand to the three-dimensional structures of *c-kit2* and *c-myc* G4s determined by solution state NMR spectroscopy (Dai et al., 2011; Hsu et al., 2009). Taking into account the experimentally suggested binding stoichiometry, 2:1 (ligand/DNA) complex models were built with the G4 targets. Both 5'- and 3'-terminal G-tetrads were

considered as potential binding sites for the ligand and were therefore examined in the docking procedures. For each binding site, the most favorable complex was selected from the docked structures on the basis of the calculated binding energies. The low-energy binding poses for the ligand are shown in Figure 3.7. The shape of **3a** effectively fills much of the available space in the plane of the two G-tetrads in the c-myc G4. The tricyclic diimidazo[1,2-*a*:1,2-*c*]pyrimidine moiety is stacked onto the planes at a 3.4 Å separation at both 5'- and 3'-faces and overlaps part of two guanine bases (G11 and G16 at the 5' G-tetrad; G9 and G13 at the 3' G-tetrad), making extensive π -stacking interactions (Figure 3.7A). In both cases, one of the two attached phenyl rings overlap part of a guanine of the tetrad, contributing to the stabilization, while the other one is twisted out of the diimidazo[1,2-*a*:1,2-*c*]pyrimidine plane. The 2-hydrazino-2-imidazoline groups of the two ligand molecules make close hydrogen bonds with the oxygen atoms of the backbone phosphate groups of the G4. It is noteworthy that the formyl group of **3** is, in both cases, hydrogen bonded to the first flanking base (A6 and T23 at 5'- and 3'-end, respectively) that stacks over the G-tetrad, thus forming a plane capping the 5' and 3' external G-tetrads of the c-myc G4. In the case of c-kit2 G4, different binding poses were observed for the ligand on the two external G-tetrad surfaces (Figure 3.7B). Indeed, the molecular model of **3a** bound to the c-kit2 G4 shows that at the 3'-end, the planar aromatic scaffold of the ligand makes π -stacking interactions with the 3' G-tetrad, while the 2-hydrazino-2-imidazoline group is positioned to participate in a hydrogen bond with a phosphate group. However, by contrast with the c-myc G4, compound **3a** does not form any additional interaction to further stabilize the complex. On the other hand, the docking calculations at the 5'-face of the c-kit2 G4 structure reveal that to make hydrogen-bonding/electrostatic contacts between the charged 2-hydrazino-2-imidazoline group and the backbone phosphate groups of the G4, the side chain of compound **3a** needs to be located in the deep groove formed by the long loop 2 of the G4. As a consequence, the diimidazo[1,2-*a*:1,2-*c*]pyrimidine core of **3a** results in not being parallel to the plane of the terminal G-tetrad. This is probably due to the 5' -3' strand polarity generated by the propeller loop topology of this G4, which does not expose the phosphodiester backbone on this side, thus making only the 3'-end more favorable to interact with ligands with positively charged side chains. Overall, the docking results reveal that **3a** has the appropriate shape and electronic structure to form specific interactions. Interestingly, the binding of **3a** could involve the first flanking bases to form a plane covering the external G-tetrads and contributing greatly

to the G4 stabilization. This could also explain the selectivity of **3a** in increasing the thermal stability of the c-myc G4.

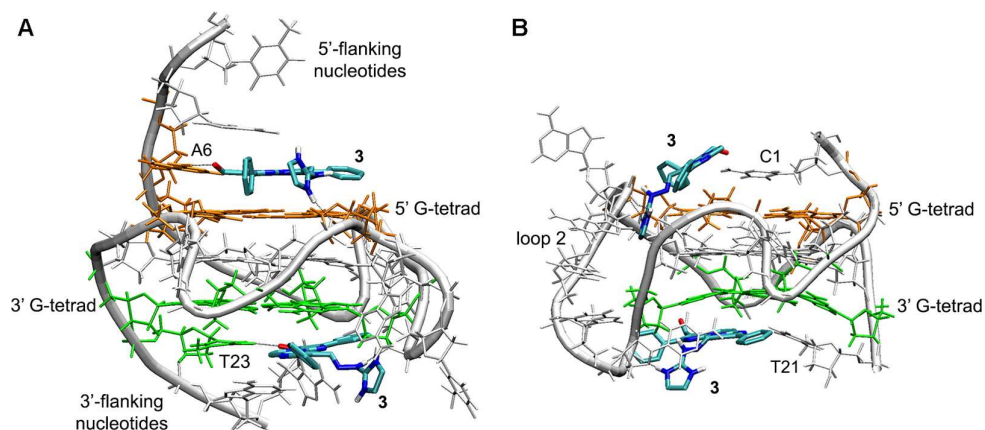


Figure 3.7. Predicted binding poses upon docking of compound **3a** onto the (A) c-myc and (B) c-kit2 G4s. Ligand molecules are represented as sticks. The phosphate backbone and nucleosides of G4s are represented in tube and stick representations, respectively. Black dashed lines indicate hydrogen bond interactions.

Biological assays

Biological assays were performed in collaboration with Prof. Giovanni Capranico (Department of Pharmacy and Biotechnology, University of Bologna, Italy). specificity for different G4 structures, we next determined the effects of selected compounds at cellular levels. We selected compound **3a** as it was the most selective agent in vitro, and we compared its effects with those of **1a** and **14a**, used as a negative control because it is unable to bind G4s in vitro) (Figure 3.8) (Sparapani et al., 2010).

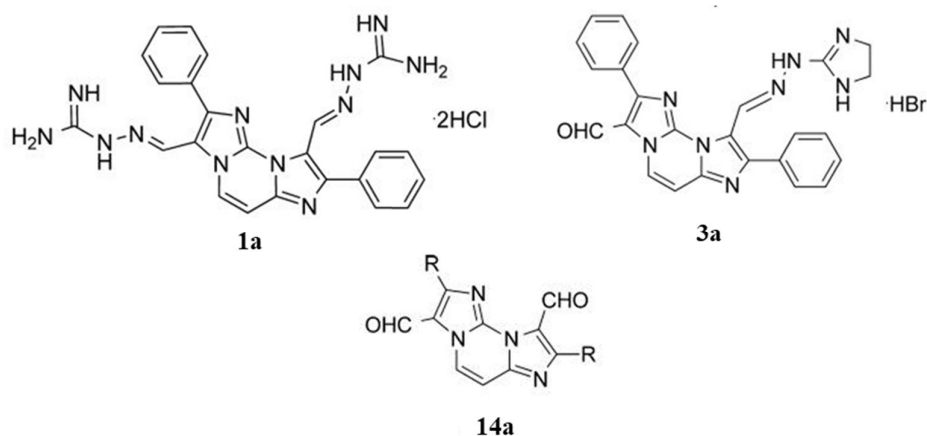


Figure 3.8. Structures of the compounds **1a**, **3a**, **14a**.

First, we determined the cell killing activity of the three selected molecules with the thiazolyl blue tetrazolium bromide (MTT) assay in two human cancer cell lines: the Human Bone Osteosarcoma Epithelial Cell line (U2OS) and colon cancer HCT116 lines. In this assay, after exposure to the compounds, cells were allowed to recovery in drug-free medium for 24–48 h before adding MTT to measure surviving cells.

HCT116 cells were more sensitive than U2OS cells to the tested compounds, however, both cell lines were more sensitive to compound **3a** than other agents (Table 3.3)

Table 3.3. Cytotoxic activity of selected hydrazone derivatives in human U2OS and HCT116 cells^a.

compd	IC ₅₀ (μM) 1 h		IC ₅₀ (μM) 24 h	
	U2OS	HCT116	U2OS	HCT116
1a ^d	101.0 ± 1.63	35.07 ± 1.06	15.97 ± 1.41	5.12 ± 1.05
3a	5.73 ± 1.48	4.18 ± 1.07	0.845 ± 0.001	0.407 ± 0.106
14a	45.23 ± 2.05	16.79 ± 1.06	9.16 ± 1.44	4.06 ± 1.07

^aThe concentrations killing 50% of cells (IC₅₀) are means ± SE of two independent experiments performed in triplicates.

In particular, compound **3a** showed high cell killing activity in U2OS cells after 24 h of treatment (IC₅₀ = 0.845 μM). In addition, **3a** was 20- and 12-fold more cytotoxic than compounds **1a** (IC₅₀ = 15.97 μM) and **14a** (IC₅₀ = 9.16 μM), respectively. Similarly, **3a** was more cytotoxic than compounds **1a** and **14a** after 1 h of cell treatment (Table 3.3). The sensitivity pattern was very similar in HCT116 cells, with compound **3a** showing highest cell

killing activity as compared with compounds **1a** and **14a** both after 24 and 1 h of treatment (Table 3.3). Therefore, the results established that compound **3a** is more cytotoxic than the other tested compounds, at least in the two cell lines used here. Next, to show whether **3a** and **1a** stabilized G4s in living cells, we visualized nuclear G4 structures in human U2OS cells by immunofluorescence microscopy using the BG4 antibody, a known specific antibody against G4s (Biffi et al., 2013). U2OS cells were treated for 24 h with compound concentrations very close to the established IC₅₀ values (Table 3.3). Immediately after treatments with tested compounds, cell death was minimal, as determined by cell fluorescence activated cell sorting (FACS) analyses and microscopy visualization. After incubation with BG4 antibody, the fluorescence signal was generated by a final incubation with a fluorochrome-labeled antibody and measured and analyzed with ImageJ. BG4 showed nuclear staining in control, untreated cells indicating that G4 structures were present at discrete chromatin sites in the nucleus (Figure 3.9A), in agreement with previous studies (Biffi et al., 2013). Fluorescence spots were of different sizes, and larger spots apparently consisted of a few smaller spots. The fluorescence patterns were similar in cells treated with the three selected hydrazone derivatives. However, the compound effects were very different because compounds **3a** and **1a**, but not **14a**, markedly increased the number, size, and fluorescence intensity of nuclear spots, thus indicating that they could stabilize G4 structures in the nucleus of living cells. We then measured the total fluorescence signal per cell for each sample, and the results indicated that **3a** was the most potent as an increase of G4 signal was detected even at 0.2 μM (a dose 4-fold lower than IC₅₀), and further increased at 1 μM, of **3a**. Compound **1a** was able to increase the nuclear G4 signal at 10 μM at comparable levels as **3a** at 0.2 μM (Figure 3.9A). In contrast, **14a** did not show any increase in fluorescence intensity at the tested concentrations significant increase of G4 after 0.2 μM ($p = 2.05 \times 10^{-32}$) and 1 μM ($p = 1.82 \times 10^{-55}$) of compound **3a** and 10 μM ($p = 2.44 \times 10^{-38}$) of **1a**. Taken together, these data demonstrate that hydrazone derivatives can stabilize G4s in cultured human cancer cells and that **3a** is the most potent agent among the studied derivatives. Although we detected an association of cytotoxic potency with G4 stabilization in living cells among the tested compounds, further investigations at the cellular level are needed to establish a clear link between G4 stabilization and biological activity of these hydrazone derivatives.

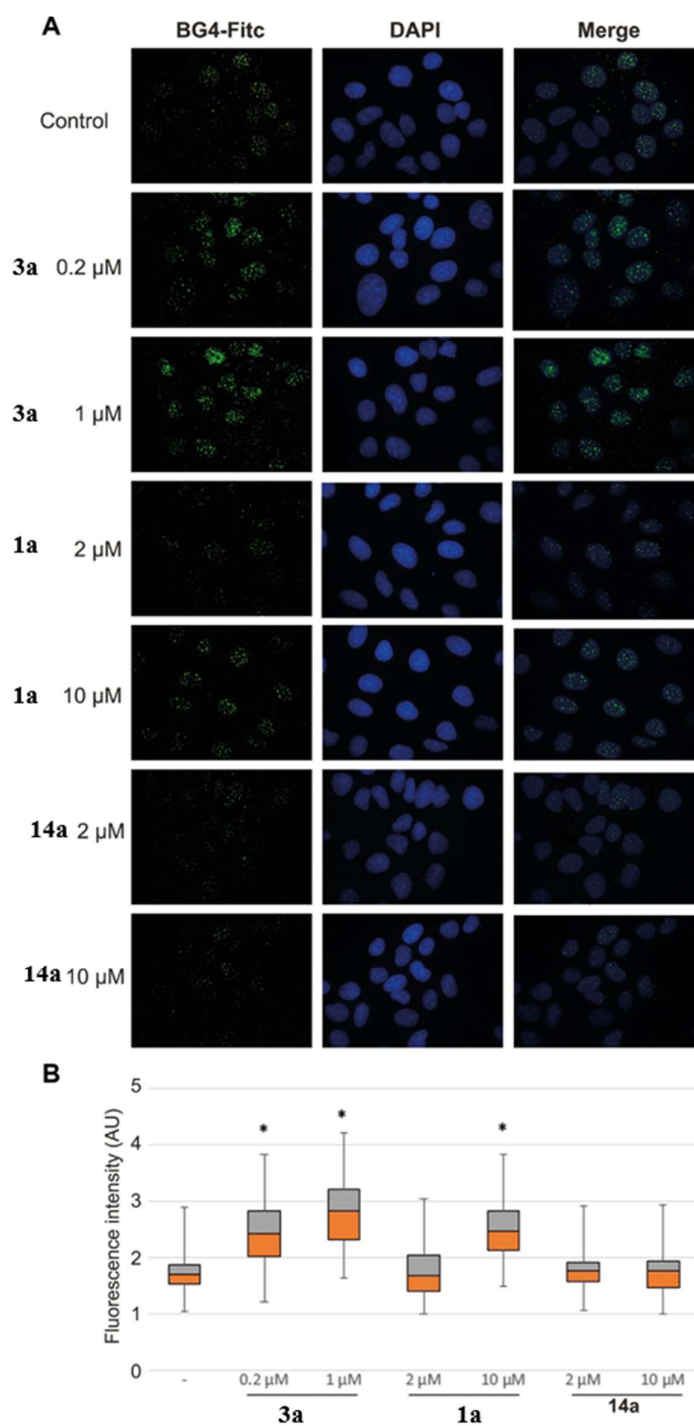


Figure 3.9. (A) Immunofluorescence showing BG4 foci (green) in human osteosarcoma U2OS cell nuclei (blue). Cells were treated with G4 ligands for 24 h and then fixed. G4s and DNA were stained with BG4 antibody and DAPI, respectively. (B) Box Plot analysis of G4 signal in the nucleus of control and treated cells. At least two experiments were considered, and 150 nuclei were counted per sample. Statistical significance was determined with the *t* test, and * indicates a *p* value less than 10^{-32} .

3.1.3 Conclusions

Starting from the lead compound **1a**, a series of analogues with structural modifications of the substituents and the core have been designed, synthesized, and evaluated as effective and selective G4 binding ligands. The results from CD assays showed that six out of nine synthesized ligands are potent G4 stabilizers that retain the pre-existing architecture of the target G4s upon interaction. However, the compounds having a different core from that of compound **1a** (**6a**, **7a** and **8a**) were also shown to increase the stability of duplex DNA and therefore they are not selective for G4s. On the other hand, compounds **2a**, **3a**, and **5a** displayed a marked preference for binding parallel G4s over duplex and antiparallel G4s.

Therefore, it was investigated in detail in terms of binding mode. NMR experiments suggested an end-stacking binding mode at the terminal G-tetrads of the G4 with a 2:1 drug/DNA binding stoichiometry. Molecular docking results provided further insight into the recognition between **3a** and the G4 DNA, highlighting the possible key structural elements involved in the interaction. Interestingly, compound **3a**, which is the only monohydrazone of the series, appears to have the appropriate structure to stack on the planar G-tetrad surfaces of c-myc G4 with the tricyclic diimidazo[1,2-*a*:1,2-*c*]pyrimidine core and, at the same time, to make (i) close hydrogen-bonding/electrostatic contact between the charged 2-hydrazino-2-imidazoline group and the oxygen atoms of the backbone phosphate of G4 and (ii) a hydrogen bond interaction between the formyl group at position 3 and the hydrogen-bond donor groups of the first flanking base that stacks on the G-tetrad.

In line with the *in vitro* assays, biological experiments demonstrated that hydrazine derivatives can effectively trap G4 structures in the nuclei of cells. Remarkably, derivative **3a** is more effective than the lead compound **1a** in inhibiting human U2OS and HCT116 cancer cell growth as well as in stabilizing G4s in osteosarcoma cells. Altogether, the reported results indicate that this molecular scaffold could actually form the basis for the development of highly selective drug-like G4 ligands with superior biological activity to accomplish the desired result of achieving new and effective anticancer drug candidates.

3.1.4 Experimental Section

Oligonucleotide synthesis and sample preparation

The DNA sequences were synthesized using standard β -cyanoethylphosphoramidite solid phase chemistry on an ABI 394 DNA/RNA synthesizer (Applied Biosystem) at the 5 μ mol scale. DNA detachment from support and deprotection were performed by treatment with concentrated ammonia aqueous solution at 55 °C for 12 h. The combined filtrates and washings were concentrated under reduced pressure, dissolved in water, and purified by high-performance liquid chromatography (HPLC) on a Nucleogel SAX column (Macherey- Nagel, 1000–8/46), using buffer A consisting of 20 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ aqueous solution (pH 7.0), containing 20% (v/v) CH_3CN , buffer B consisting of 1 M KCl, 20 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ aqueous solution (pH 7.0), containing 20% (v/v) CH_3CN , and a linear gradient from 0% to 100% B for 30 min with a flow rate 1 mL/min. The fractions of the oligomers were collected and successively desalted by Sep-pak cartridges (C-18). The isolated oligomer was proved to be >98% pure by NMR. In particular, the following oligonucleotides have been synthesized: d[$\text{TAGGG}(\text{TAGGG})_3$] (Tel23) and d[($\text{TAGGG})_8\text{TT}$] (Tel26) that are two different truncations of human telomeric sequence, two sequences that occur in the nuclease hypersensitive region of the promoter of *KIT* oncogene d($\text{AGGGAGGGCGCTGGGAGGAGGG}$) (c-kit1) and d($\text{CGGGCGGGCGCGAGGGAGGGG}$) (c-kit2), the modified c-kit2 promoter sequence d($\text{CGGGCGGGCGCTAGGGAGGGT}$) (c-kit2_{T12/T21}), the *MYC* promoter sequence d($\text{TGAGGGTGGGGAGGGTGGGGAAGG}$) (c-myc), the modified *MYC* promoter sequence d($\text{TGAGGGTGGGTAGGGTGGGTAA}$) (myc22), and the self-complementary duplex forming dodecamer d(CGCGAATTCGCG) (ds12). The concentration of oligonucleotides was determined by UV adsorption measurements at 90 °C using appropriate molar extinction coefficient values ϵ ($\lambda = 260$ nm) calculated by the nearest neighbor model (Cantor et al., 1970). Samples were heated at 90 °C for 5 min and then gradually cooled to room temperature overnight. Parallel arrangement of telomeric sequence (Tel23-p) was prepared as previously described (Pagano et al., 2015).

Circular dichroism spectroscopy

Circular dichroism (CD) experiments were recorded on a Jasco J-815 spectropolarimeter equipped with a PTC-423S/15 Peltier temperature controller. All the spectra were recorded at

20 °C in the wavelength range of 230–360 nm and averaged over three scans. The scan rate was set to 100 nm/min, with a 1 s response time and 1 nm bandwidth. Buffer baseline was subtracted from each spectrum. For the CD experiments in the presence of metal cations, 10 μM G4 DNA and 15 μM duplex DNA were used. CD spectra of DNA/ligand mixtures were obtained by adding 4 molar equivalents of ligands (stock solutions of ligands were 6 mM in DMSO). DNAs were prepared in 10mM potassium phosphate (for c-kit1, Tel23, and Tel26) or 10mM lithium phosphate (for Tel23-p and ds12) buffers (pH 7.0) containing 100 mM KCl. Because the T_m value of the c-myc G4 alone is above 80 °C in buffer containing 100 mM KCl, it is very difficult to evaluate the stabilizing effects of adding ligands, thus a buffer solution containing lower amounts of K^+ (5 mM potassium phosphate, 20 mM KCl) was used for the experiments on c-myc G4. As for c-kit2, it was shown that this sequence forms a dimeric G4 in 100 mM K^+ -containing solution (Kuryavyi et al., 2010), with a completely different fold compared to the monomeric G4 that is formed in buffer solution containing low amounts of K^+ (Hsu et al., 2009; Kuryavyi et al., 2010). Therefore, 5 mM potassium phosphate buffer containing 20 mM KCl was also used in this case. All spectra were baseline corrected and analyzed using Origin 7.0 software. For the studies in the absence of metal cations, oligonucleotides (12–15 μM) were prepared in 50 mM Tris buffer and the CD spectra were recorded 10 min after each ligand addition (stepwise additions of 0.5 mol equivalents). All spectra were baseline corrected and analyzed using Origin 7.0 software. CD melting were carried out in the 20–100 °C temperature range at 1 °C/min heating rate by following changes of CD signal at the wavelengths of the maximum CD intensity. CD melting experiments were recorded in the absence and presence of ligands (4 mol equivalents) added to the folded DNA structures. For selected compounds, CD melting experiments were also performed by varying ligand concentration. Three different [ligand]/[DNA] ratio were examined (1:1, 2:1 and 4:1). The melting temperatures (T_m) were determined from curve fit using Origin 7.0 software. ΔT_m values were determined as the difference in melting temperature between the G4-DNA with and without ligands.

FRET melting studies

The ability of the compounds to stabilize DNA sequences was investigated using a fluorescence resonance energy transfer (FRET) assay modified to be used as a high-throughput screen in a 96-well format. The studied sequences were the G4 forming sequence 5'-FAM-d(GGG[TTAGGG]₃)-TAMRA-3' (F21T) (Mergny and Maurizot, 2001), and the duplex-

forming sequence 5'-FAM-d(TATAGCTATA-HEG-TATAGCTATA)-TAMRA-3' (T-loop) (HEG linker: $[(-\text{CH}_2-\text{CH}_2-\text{O}-)_6]$). The labeled oligonucleotides had attached to them the donor fluorophore FAM, 6-carboxyfluorescein, and the acceptor fluorophore TAMRA, 6-carboxytetramethylrhodamine. The FRET probe sequences were diluted from stock to the correct concentration (400 nM) in a 60 mM potassium cacodylate buffer (pH 7.4) and then annealed by heating to 85 °C for 10 min, followed by cooling to room temperature in the heating block. The compounds were stored as a 10 mM stock solution in dimethyl sulfoxide (DMSO); final solutions (at 2× concentration) were prepared using 10 mM HCl in the initial 1:10 dilution, after which 60 mM potassium cacodylate buffer (pH 7.4) was used in all subsequent steps. The maximum HCl concentration in the reaction volume (at a ligand concentration of 20 μM) is thus 200 μM, well within the range of the buffer used. Relevant controls were also performed to check for interference with the assay. 96-Well plates (MJ Research, Waltham, MA) were prepared by aliquoting 50 μL of the annealed DNA into each well, followed by 50 μL of the compound solutions. Measurements were made on a DNA Opticon Engine (MJ Research) with excitation at 450–495 nm and detection at 515–545 nm. Fluorescence readings were taken at intervals of 0.5 °C in the range 30–100 °C, with a constant temperature being maintained for 30 s prior to each reading to ensure a stable value. Final analysis of the data was carried out using a script written in the program Origin 7.0 (OriginLab Corp., Northampton, MA). The advanced curve-fitting function in Origin 7.0 was used for calculation of ΔT_m values. All determinations were performed in triplicate. Esds in ΔT_m are ± 0.2 °C. For selected compounds, additional FRET melting studies were performed by using the following G4 forming sequences: the 5'-FAM-d(GGGCGGGCGCGAGGGAGGGG)-TAMRA-3' (Fckit2T) and 5'-FAM-d(TGAGGGTGGGTAGGGTGGGTAA)-TAMRA-3' (FcmycT) oligonucleotides from the *KIT* and *MYC* promoter regions, respectively, and the telomeric F21T G4-forming sequence were annealed at high concentrations to promote the parallel conformation (F21T-p). Labeled oligonucleotides were purchased from Biomers (Germany) and purified employing standard HPLC protocols. Such studies were performed under the same buffer and experimental conditions used for the corresponding CD experiments. The FRET melting assay was carried out on a FP-8300 spectrofluorometer (Jasco) equipped with a Peltier temperature controller accessory (Jasco PCT-818) with excitation at 492 nm and detection at 522 nm. Both excitation and emission slit widths were set at 5 nm. A sealed quartz cuvette with a path length of 1 cm was used. The final concentration of the G4s was 0.2 μM. The fluorescence melting of the G4s

was monitored at 1 °C/min with and without ligands (at a 4:1 ligand/DNA ratio). Final analysis of the data was carried out using Origin 7.0.

Nuclear Magnetic Resonance

NMR experiments were performed on a 700 MHz Varian Unity INOVA spectrometer, with data recorded at 25 °C. One-dimensional proton spectra of the sample in H₂O were recorded using pulsed-field gradient DPGSE for H₂O suppression. DNA samples were prepared at 0.2–0.5 mM strand concentration in 0.6 mL of (H₂O/D₂O 9:1) buffer solution. NMR experiments on G4 forming oligonucleotides from *MYC* (myc22) and *KIT* (c-kit_{T12/T21}) promoters were performed by employing the same buffers and experimental conditions as used for the determination of their 3D structures (Ambrus et al., 2005; Kuryavyi et al., 2010) in order to avoid different G4 folds or conformational heterogeneity in solution, which is detrimental to such experiments. The solution was either 25 mM KH₂PO₄, 70 mM KCl, 0.2 mM EDTA, pH 7.0 (for Myc22), or 5 mM KH₂PO₄, 20 mM KCl, 0.2 mM EDTA, pH 6.8 (for c-kit_{T12/T21}). Aliquots of a stock solution of **3a** in DMSO-d₆ were added directly to the DNA solution inside the NMR tube. The NMR data were processed on an iMAC running iNMR software (www.inmr.net).

Docking simulations

The solution structures of G4 formed by the human c-kit2 promoter sequence (PDB ID: 2KQH) (Hsu et al., 2009) and of c-myc G4 bound to quindoline molecule (PDB ID: 2L7V) (Dai et al., 2011) were used as the targets for docking studies. The ligand found in the c-myc complex was removed from the structure to leave empty binding sites. After optimizing the ligand and assigning partial atomic charges, docking calculations were performed with the AutoDock 4.2 program using the Lamarckian genetic algorithm (Morris et al., 2009). The all-parallel-stranded arrangement with propeller loops linking adjacent parallel strands on opposite surfaces of the G4s results in accessible outer planar G-tetrad surfaces that, being the potential binding sites for the ligands, were defined as the grid boxes for the docking. The size of the boxes was constrained to 18 Å × 18 Å × 18 Å in the x, y, and z dimensions. Grid maps were generated for each atom type in the ligand using AutoGrid. An active site box was created with a grid spacing of 0.375 Å. The maximum number of energy evaluations was set to 2.5 × 10⁶, the maximum number of genetic algorithm operations was set to 2.7 × 10⁴, the number of individuals in a

population was set to be 300, and the rates of mutation and crossover were set to 0.02 and 0.8, respectively. When searching the conformational and orientation spaces of ligand with rotatable bonds having full flexibility, the structure of the G4 was kept rigid. The most favorable binding poses were selected on the basis of both the calculated energies and visual inspection. All the figures were drawn using VMD 1.9 software (www.ks.uiuc.edu/Research/vmd).

Cell culture and Immunofluorescence

Human osteosarcoma U2OS and colon cancer HCT116 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Carlo Erba), supplemented with 1% L-glutamine and 10% fetal bovine serum (FBS). U2OS cells were seeded in 35 mm dishes at a concentration of 100000 cell/mL. Twentyfour hours after seeding, cells were treated with **1a** (2 and 10 μ M), **3a** (0.2 and 1 μ M), or **14a** (2 and 10 μ M). Cells were then fixed in methanol:acetic acid (3:1), permeabilized with 0.1% triton-X100/PBS, and blocking in 2% milk/PBS. Immunofluorescence was performed using standard methods with BG4, anti-FLAG (Cell Signaling Technology), and antirabbit Alexa 488-conjugated (Invitrogen) antibodies. Nuclei were stained with 4',6-diamidino-2-phenylindole (DAPI) (Sigma-Aldrich), and coverslips were mounted with Mowiol 4-88 (Sigma-Aldrich). Fluorescence signal was determined using ImageJ software with the following formula: corrected total cell fluorescence (CTCF) = integrated density - (area of selected cell \times mean fluorescence of background readings). Significance was determined by parametric tests "Student's *t* test". BG4 antibody was obtained by transfection of BG4 plasmid (kindly obtained by S. Balasubramanian) in BL21 Escherichia coli cells. Then BG4 protein expression was induced by the autoinduction method as described by Studier (Studier, 2005). BG4 was purified by using silica-based resin (Protino Ni-IDA) precharged with Ni²⁺ ions, eluted with 250 mM Imidazole/PBS pH 8.0. The eluted antibody was concentrated with Amicon Ultra-15 centrifugal filter units (Millipore), and imidazole was finally removed by buffer exchange with PBS pH 8.0 with Amicon Ultra-15 centrifugal filter units.

MTT Cell Proliferation Assay

U2OS and HCT116 cells were cultured in DMEM as above and were seeded in 24-wells at a concentration of 40000 cell/mL. Twenty-four hours after seeding, cells were treated with the compound at the indicated concentration and time. After 1 or 24 h of treatment, agents were

removed and cells were further cultured in complete drug-free medium for 1–2 days. Then, MTT (Sigma-Aldrich) was added to each well and incubated for 1 h at 37 °C. Then medium was removed and 300 µL of DMSO (Sigma-Aldrich) were added and incubated for 1 h at room temperature. Then 100 µL of the solution was put in 96-well, and absorbance at 595 nm was measured using a multiplate reader. The linear regression parameters were determined to calculate the IC₅₀ (GraphPad Prism 4.0, Graph Pad Software Inc.).

3.2 Targeting the *BCL2* Gene Promoter G-Quadruplex with a New Class of Fuopyridazinone-Based Molecules (Paper II)

3.2.1 Introduction

Apoptosis is regulated by many factors, which are involved in the activation and execution of related events (Hengartner, 2000). The Bcl-2 (Bcell lymphoma-2) family of proteins is the main regulator of apoptotic process, acting either to promote or inhibit it (Adams, 1998). Indeed, this family includes both pro- and anti-apoptotic proteins that should function in harmony for a controlled regulation of apoptosis pathways, and their relative levels are critical for cells. Overexpression of *BCL2* gene, which encodes the antiapoptotic Bcl-2 protein, greatly contributes to the resistance of cancer cells to apoptosis, and it has also been reported to play a role in the resistance to conventional cancer treatments (Yip and Reed, 2008). Targeting the *BCL2* gene to inhibit protein expression may be an effective way to prevent the evasion of apoptosis in cancer cells and to increase chemotherapeutic efficacy. The human *BCL2* gene has two promoters: P1, identified as the principal transcriptional promoter, and P2 (Seto et al., 1988). The P1 promoter and the region upstream thereof, contain G-rich elements that have unequivocally been shown to form G4s (Dexheimer et al., 2006).

A G4-forming sequence that plays a major role in the regulation of *BCL2* transcription is located directly upstream (~30 bases) from the P1 promoter (Dai et al., 2006). This sequence forms a predominant G4 structure (bcl-2) with a distinctive hybrid (3+1) topology having two lateral loops (of 3 and 7 nucleotides, respectively), one single-nucleotide propeller loop, and four grooves of different widths (Dai et al., 2006) (Figure 3.10).

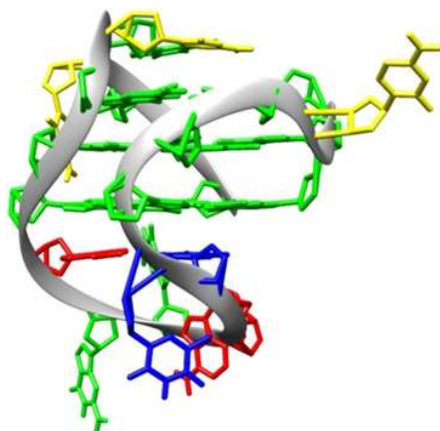


Figure 3.10. Representative model of bcl-2 G4 NMR structure (guanine, green; adenine, red; thymine, blue; cytosine, yellow).

Several organic molecules have been shown to be able to bind and stabilize bcl-2 G4 with subsequent down-regulation of Bcl-2 transcription and expression levels (Feng et al., 2016; Jana et al., 2017; Micheli et al., 2016; Wang et al., 2010), thereby demonstrating the real potential of the G4-targeting therapeutic approach, which also has the great advantage of circumventing the problem of acquired resistance, which is the main limitation of current Bcl-2-targeted therapeutic strategies.

Most of the bcl-2 G4 ligands identified so far are characterized by a large, planar aromatic core with the capability of stacking on the external G-tetrads, a common element in G4 recognition (Amato et al., 2014a; Neidle, 2016; Pagano et al., 2015). Therefore, although these ligands have shown good G4 over duplex selectivity, they do not exhibit specificity for bcl-2 versus other G4 structures. Moreover, most of them share unfavorable chemical properties such as high molecular weight and/or hydrophobic nature, which result in very poor drug-like properties. Therefore, despite being excellent ligands *in vitro*, the unfavorable pharmacokinetics and toxicity due to poor selectivity versus different G4 structures critically hamper their advancement in chemotherapy.

The next goal in G4-based drug design is to enhance the drug-likeness of the ligands and their selectivity for a specific structure, so as to permit pathway-specific targeting. The peculiar structural features of bcl-2 G4 suggest that specific targeting of such a promoter G4 may, in principle, be possible. To discover structurally novel compounds that do not show the aforementioned drawbacks, it is necessary to explore a larger chemical space. In this perspective, compounds without a large planar aromatic core may be more promising to achieve the desired drug-likeness and G4 specificity.

With the aim of identifying compounds with these properties, in collaboration with Dr. Mariateresa Giustiniano (Department of Pharmacy, University of Naples “Federico II”, Italy) we decided to exploit the innate ability of multicomponent reactions to generate diversity and to explore a large chemical space in short times, for searching new G4-targeting ligands. In particular, we designed and synthesized a small series of furopyridazinone-based molecules. The pyridazinone nucleus was selected because it is present in a plethora of compounds endowed with various biological activities (antihypertensive, antithrombotic, anti-inflammatory, and anticancer) (Asif, 2012; Prime et al., 2011), and it is considered a privileged structure that is currently experiencing a renewed interest in medicinal chemistry. Starting from this nucleus, we recently presented an expeditious multicomponent one-pot reaction

methodology for generating a small collection of fully substituted furo[2,3-*d*]pyridazin-4(5*H*)-ones (FPs) (Giustiniano et al., 2014, 2015). The facile and efficient synthetic route for the FP derivatives is shown in Figure 3.11.

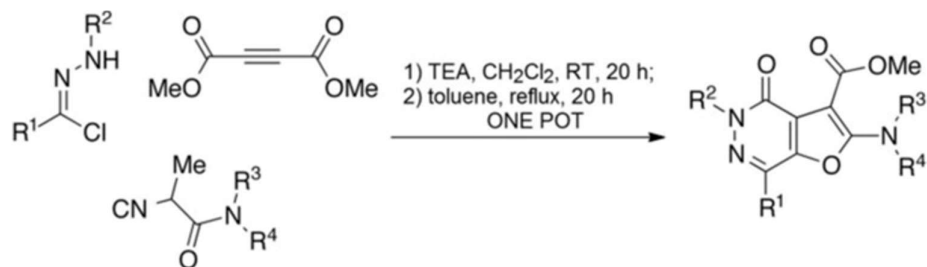


Figure 3.11. General synthetic route for the FP derivatives.

This strategy allowed us to readily produce a variety of analogues (**1b-11b**, Table 3.4) with different side arms (in order to avoid confusion, compounds **1-11** in Paper II here appear as **1b-11b**). It is known that side arms can be important in modulating the affinity of G4-targeting ligands, as well as in influencing their binding specificity toward different conformations (Amato et al., 2016). In addition, their structural features should prevent intercalation into the duplex DNA.

Table 3.4. Structures of the furopyridazinone derivatives.

Compound	R ₁	R ₂	-NR ₃ R ₄	R ₅
1b	Me-			-COOMe
2b		Cl-		-COOMe
3b	I-			-COOMe
4b	PhO-			-COOMe
5b				-COOMe
6b				-COOMe
7b	Cl-	MeO-		-COOMe
8b		MeO-		-COOMe
9b		MeO-		-COOMe
10b				-COOMe
11b	I-			-COO-Li ⁺

3.2.2 Results and discussion

Circular Dichroism studies

Circular dichroism (CD) spectroscopy and CD-melting experiments were carried out to investigate the *in vitro* binding properties of synthesized FP derivatives for bcl-2 versus other G4 structures (Pagano et al., 2012). In addition to bcl-2, two G4-forming sequences from *KIT* (c-kit1) and *MYC* (c-myc) oncogene promoter regions, and a G4 from the human telomere (Tel24) were used in these experiments. The self-complementary duplex-forming dodecamer d(CGCGAATTCGCG) (ds12) was also used to estimate the G4 over duplex selectivity of the ligands. The structure adopted by each DNA sample was first verified by CD measurements. In agreement with the presence of parallel G4 topologies, c-kit1 and c-myc showed a positive band at 264 nm and a negative one at 243 nm in the CD spectra. Otherwise, bcl-2 and Tel24 showed two positive bands at around 265 and 290 nm and a weak negative band at around 240 nm, in agreement with the presence of hybrid structures as major conformation (Figure 3.12). The structure of ds12 was also verified by CD, showing the typical spectrum of a duplex DNA (Figure 3.12). CD experiments were then performed to determine if compounds **1b–11b** alter the native folding topology of the investigated DNA structures. Upon addition of each compound (4 mol equivalents) to the DNAs (G4s and duplex), no significant alterations of CD spectra were observed, thus suggesting an overall preservation of their architectures (Figure 3.12).

The DNA stabilizing properties of **1b–11b** were evaluated by CD-melting experiments measuring the ligand-induced change in the melting temperature (ΔT_m) of both G4 and duplex structures. The results of these experiments (Table 3.4) show that none of the derivatives significantly increase the stability of c-kit1, c-myc and Tel24 G4s, as well as that of ds12 duplex, thus suggesting their poor ability to bind to those DNA structures ($\Delta T_m \leq 2.6$ °C). However, compounds **2b** and **3b** showed an appreciable enhancement of the stability of bcl-2 ($\Delta T_m \geq 3.5$ °C).

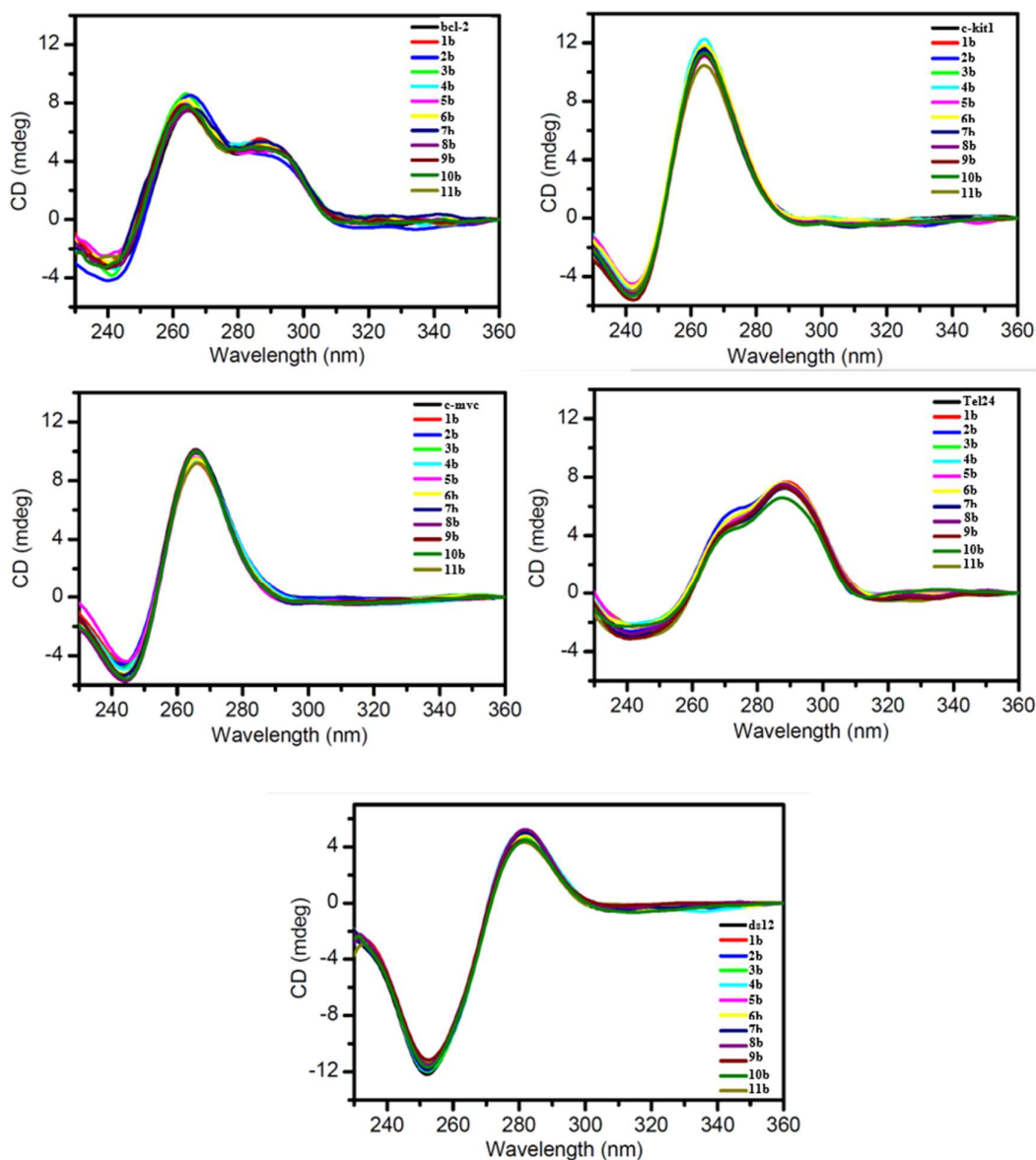


Figure 3.12. CD spectra of the investigated DNAs in the absence and in presence of 4 mol equivalents of compounds 1b-10b.

Notably, a non-negligible destabilization of the ds12 duplex occurs with some of these compounds. We hypothesize that it may be due to nonspecific interactions of compounds with the DNA molecule that might favor single-stranded DNA, resulting in a shift of the folded–unfolded equilibrium toward the unfolded form during the melting experiment, which, in turn, would result in a slight decrease in T_m .

Table 3.5. Ligand-induced thermal stabilization of G4 and duplex DNAs measured by CD melting experiments.

Compd	ΔT_m (°C) ^a				
	bcl-2	c-kit1	c-myc	Tel24	ds12
1b^d	1.9	1.1	0.2	0.0	-0.6
2b	4.0	1.3	0.5	1.4	-1.3
3b	3.5	0.9	0.0	0.1	-0.4
4b	2.9	0.8	0.5	0.0	-2.0
5b	2.3	0.4	1.2	0.0	-0.7
6b	2.2	0.8	2.0	0.1	-1.2
7b	2.8	1.1	0.8	0.2	-0.5
8b	2.3	0.7	1.0	0.1	-2.0
9b	2.3	1.0	2.6	0.0	-2.0
10b	1.6	1.0	1.4	0.0	-1.6
11b	2.9	0.4	2.0	0.1	-1.4

^a ΔT_m represents the difference in melting temperature [$\Delta T_m = T_m$ (DNA+4 ligand equiv) - T_m (DNA)]. The T_m values of DNAs alone are bcl-2 = 64.8 ± 0.5 °C, c-kit1 = 59.8 ± 0.5 °C, c-myc = 89.3 ± 0.5 °C, Tel24 = 59.5 ± 0.5 °C, ds12 = 69.5 ± 0.5 °C.

Despite compounds **2b** and **3b** not having shown very potent thermal stabilization effects, they exhibited a clear preference for the bcl-2 over the other G4 motifs investigated in this study. This suggests that the aromatic furo[2,3-*d*]pyridazin 4(5*H*)-one core can, if appropriately substituted, actually represent a new molecular scaffold for the specific targeting of *BCL2* gene promoter.

Fluorescence Studies

To obtain quantitative data regarding the affinity of **2b** and **3b** for bcl-2 G4, fluorescence titration experiments were performed (Giancola and Pagano, 2013). Fluorescence emission spectra of the ligands in the absence and presence of increasing amounts of G4 were recorded. The fluorescence intensity of the ligands decreased gradually with addition of DNA until it reached saturation (Figure 3.14). The binding curves were obtained by plotting the fraction of bound ligands (α), calculated following fluorescence changes at the emission maximum, as a function of G4 concentration. The curves were fitted using an independent and equivalent binding sites model, by means of nonlinear regression algorithm, giving binding constants (K_b) of $1.0 (\pm 0.4) \times 10^6$ and $0.6 (\pm 0.2) \times 10^6$ M⁻¹ for compounds **2b** and **3b**, respectively, and a stoichiometry of 1:1 in both cases.

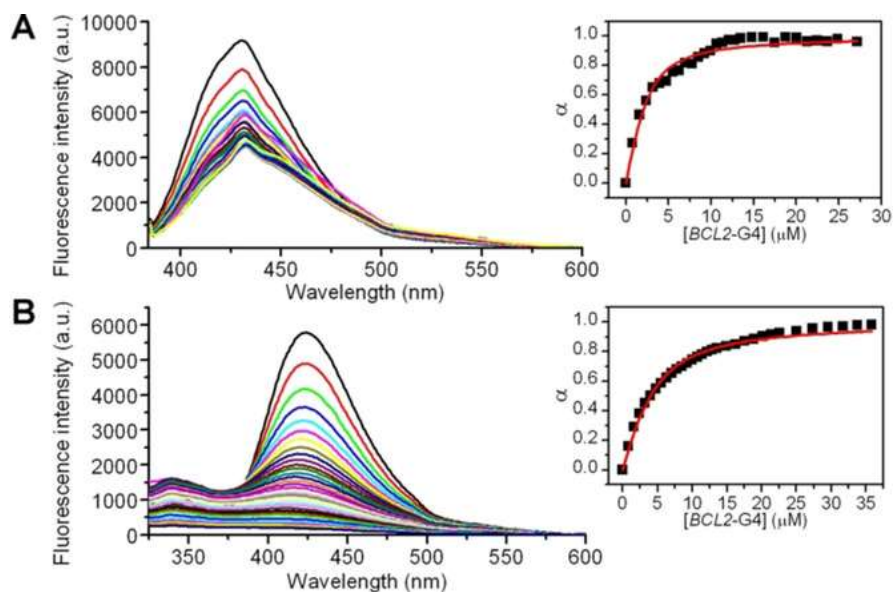


Figure 3.13. Representative fluorescence emission spectra of A) **2b**, and B) **3b** (2.5 μM) in the absence and presence of stepwise additions (5 μL) of bcl-2-G4 (165 nM) at 25 $^{\circ}\text{C}$. Insets show the titration curves obtained by plotting the fraction of bound ligand (α) versus DNA concentration

Nuclear Magnetic Resonance studies

NMR spectroscopy was used to obtain indications about the binding mode of two selected compounds to bcl-2 (Pagano et al., 2012). According to the literature, under the experimental conditions used, the investigated sequence forms a single G4 conformation characterized by 12 well-resolved imino proton peaks, corresponding to the 12 guanines involved in the three G-tetrad planes (Figures 3.15 and 3.16) (Dai et al., 2006). Upon addition of increasing amounts of **2b** and **3b** to the bcl-2 solutions, gradual chemical shift changes for some G4 proton signals were observed. Some signals of the compounds were also detected in the spectra. These peaks only grew in intensity, without showing any significant change in chemical shift values by increasing ligand concentration, clearly suggesting a rapid binding process on the NMR timescale.

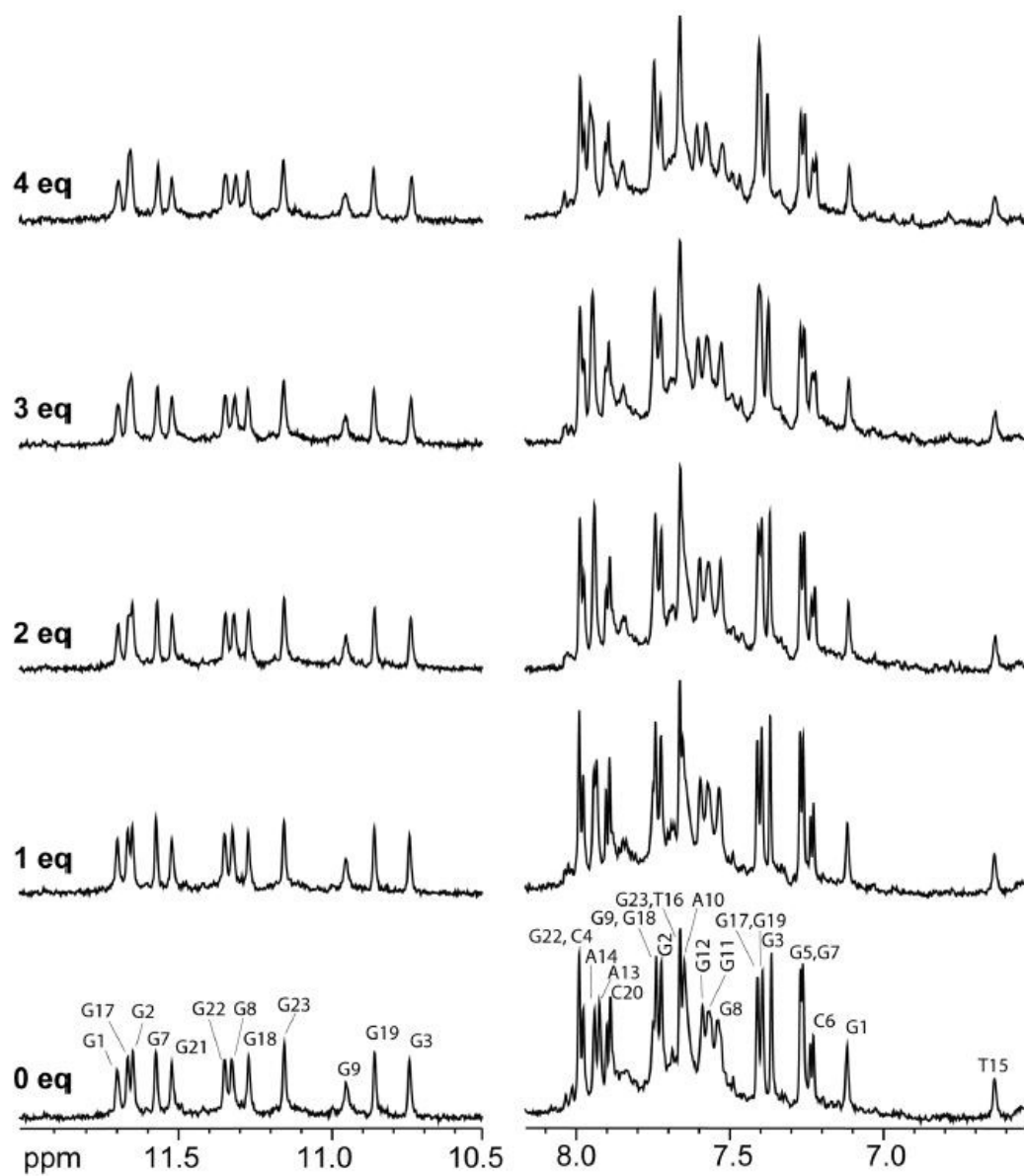


Figure 3.14. Imino (left) and aromatic (right) regions of the 1D ^1H NMR spectra of bcl-2 G4 titrated with 2b. The ligand equivalents are shown on the left of the spectra.

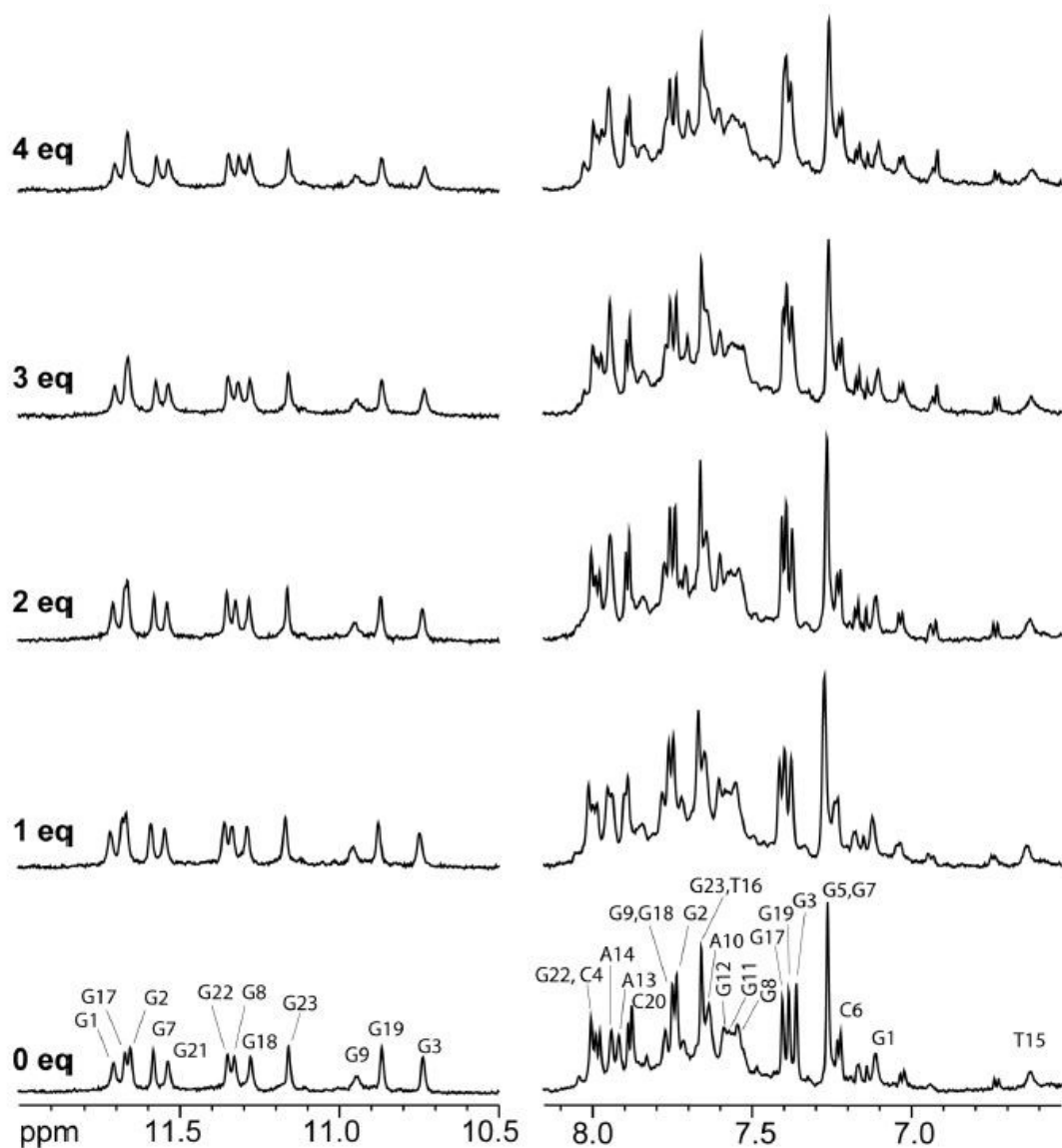


Figure 3.15. Imino (left) and aromatic (right) regions of the 1D ^1H NMR spectra of bcl-2 G4 titrated with **3b**. The ligand equivalents are shown on the left of the spectra.

To evaluate the DNA residues involved in the interaction with **2b** and **3b**, the chemical shift variations ($\Delta\delta$) of bcl-2 were calculated (Figure 3.16). The residues of the G4 structure are highlighted with different colors to outline different ranges of $\Delta\delta$ values (green, orange, and red for $|\Delta\delta| \leq 0.01$, $0.01 < |\Delta\delta| \leq 0.02$, and $|\Delta\delta| > 0.02$, respectively). Interestingly, both ligands were able to induce significant changes in the chemical shift of the aromatic protons of residues A10–T16 (except for T15), which form the 7-nucleotide lateral loop of the G4. Moreover, the binding of **2b** also caused the perturbation of the aromatic signals of G9 and G17, that are very

close to the loop. On the other hand, **3b** also induced variations of the aromatic protons of G1–G3 and G22–G23 residues that face into the same groove of bcl-2. All these findings give indications that **2b** and **3b** bind to the target G4 mainly through interactions at the long lateral loop and in the groove between the fourth and first parallel G-strands.

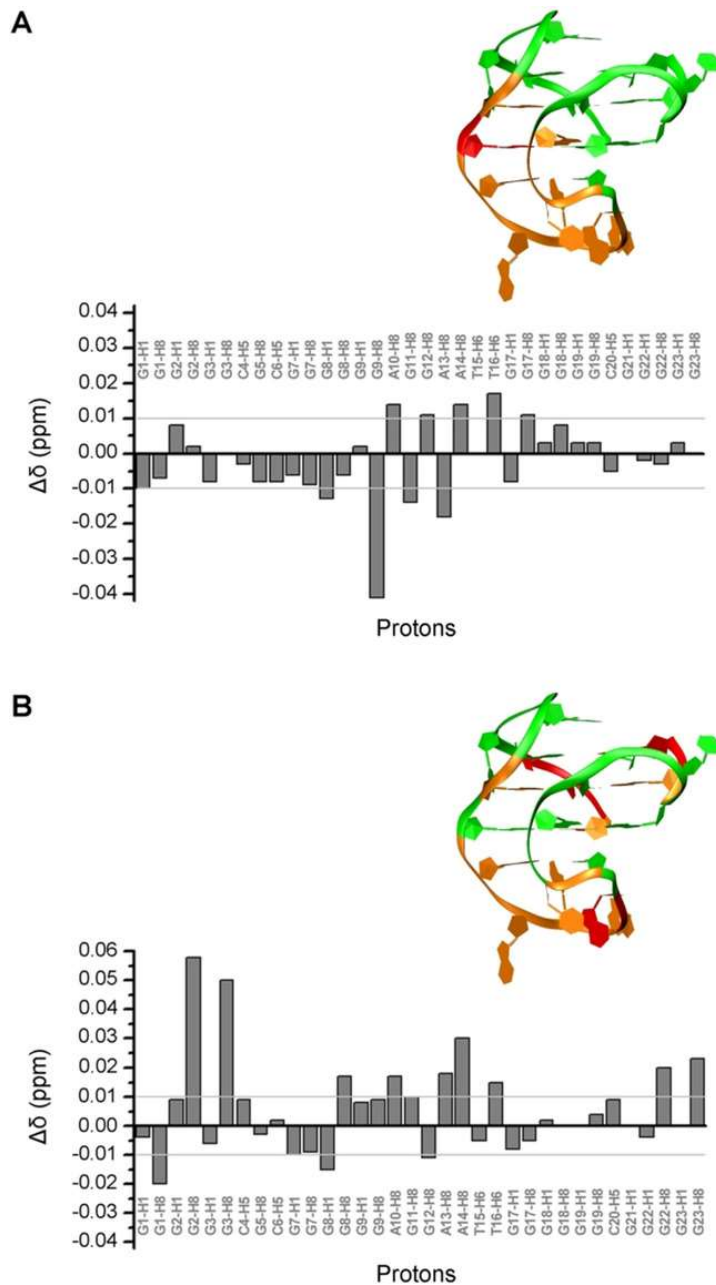


Figure 3.16. Chemical shift variations ($\Delta\delta$) of selected signals of bcl-2 G4 upon addition of compounds (A) **2b** and (B) **3b**, and three-dimensional structure of bcl-2 colored according to $\Delta\delta$ values (green, orange, and red for $|\Delta\delta| \leq 0.01$, $0.01 < |\Delta\delta| \leq 0.02$, and $|\Delta\delta| > 0.02$, respectively).

Biological assays

In collaboration with Dr. Sonia Di Gaetano and Dr. Domenica Capasso (Institute of Biostructures and Bioimaging, CNR, Naples, Italy) biological assays were also performed to evaluate the cellular effects of compound **3b**, and to see whether it could also interact with the G4-forming sequence of *BCL2* promoter region in cellulo, and thus decrease gene transcription and expression levels. Compound **3b** was selected for such further studies because of its higher solubility and bcl-2 selectivity over compound **2b**. RHPS4, one of the gold-standard ligands for G4 affinity and G4 over duplex selectivity (Cheng et al., 2008; Gavathiotis et al., 2001), was also tested for comparison. First, the cytotoxic activity of the compound was evaluated *in vitro* on different human cancer cell lines (HepG2, MCF-7, HeLa, WM266, Jurkat) and normal human dermal fibroblasts (HDF). The experiments, performed at 10 μM ligand concentration for 24 h, showed an interesting cytotoxic activity on all the cancer cell lines examined (Figure 3.17), with a higher effect (~30% inhibition) observed on Jurkat cells (treated with 10 or 25 μM ligand concentration), where *BCL2* is known to play a critical role in regulating the survival and tumorigenicity of cells (Reed et al., 1990). Remarkably, **3b** did not exhibit appreciable activity on normal cells. In comparison, RHPS4, used here as a positive control, turned out to be more active on all tested tumor cell lines, but it also exhibited a certain degree of toxicity on normal cells. This was an expected result in view of its high affinity for G4, which, coupled with the lack of selectivity for a specific G4 structure, leads to strong, but more indiscriminate cytotoxicity.

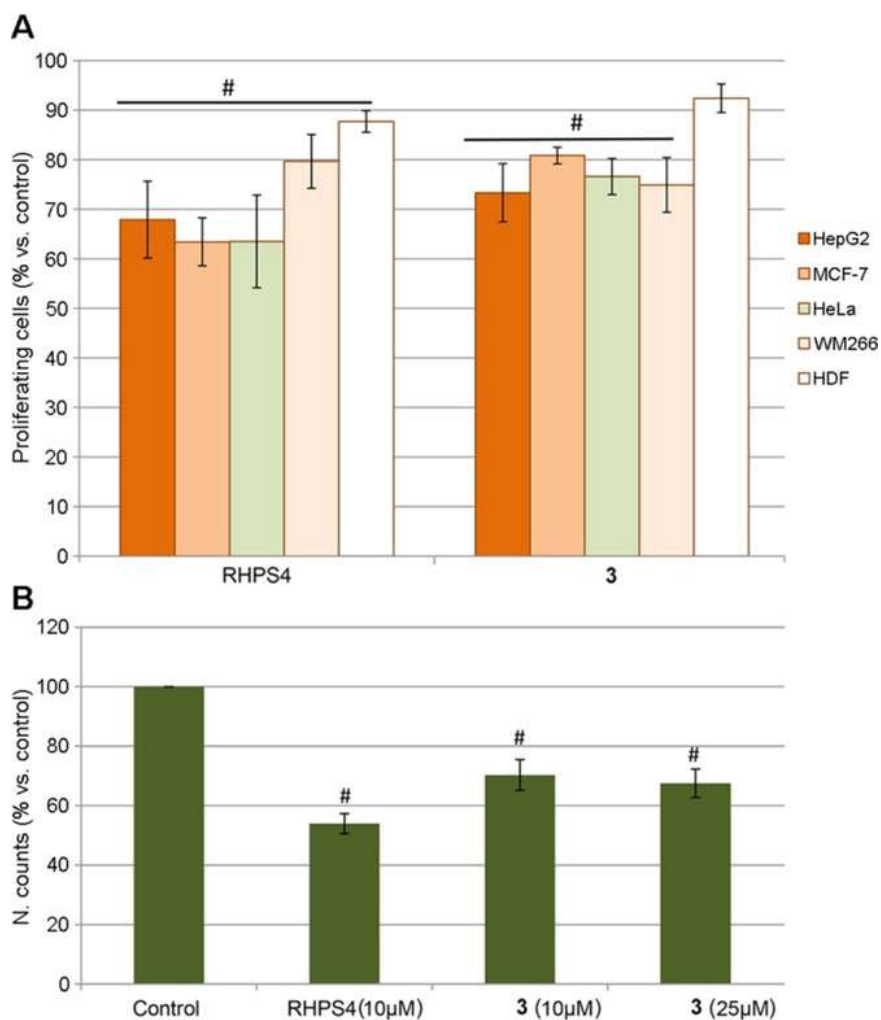


Figure 3.17. Cytotoxic effect of RHPS4 and **3b** on A) HepG2, MCF-7, HeLa, WM266, and HDF cell lines treated with 10 µM ligand concentration for 24 h, and B) Jurkat cells treated with the reported concentrations for 24 h. The results are presented as the percentage of proliferating cells with respect to control (vehicle-treated cells). All data are mean values \pm SEM of at least three independent experiments performed in triplicate ($^{\#}p < 0.05$).

To determine whether **3b** induces apoptosis, Jurkat cells were treated with the compound (and with RHPS4 for comparison) and, after 16 h, apoptosis was evaluated with annexin V-FITC/PI double staining by flow cytometry analysis (Capasso et al., 2014). Results of this experiment (Figure 3.18) showed that the cells treated with **3b** exhibited 30% early apoptotic cells with respect to the control; in contrast, for RHPS4, the entire population of cells was found to be in the late apoptotic state. To determine the effect of **3b** on *BCL2* gene transcription, quantitative analysis of mRNA was carried out by means of real-time PCR. Upon treatment of Jurkat cells with 25 µM of compound **3b** for 24 h, a significant decrease (~24%) in *BCL2* mRNA

was observed (Figure 3.18). In parallel, to confirm that the observed decrease in mRNA levels leads to a decrease in Bcl-2 protein, western blot analysis was performed by using an anti-Bcl-2 antibody. These experiments showed that treatment with **3b** results in a substantial decrease (~40%) in the level of Bcl-2 protein, in agreement with results obtained by qPCR. Similar results were obtained in both analyses with RHPS4 (Figure 3.18).

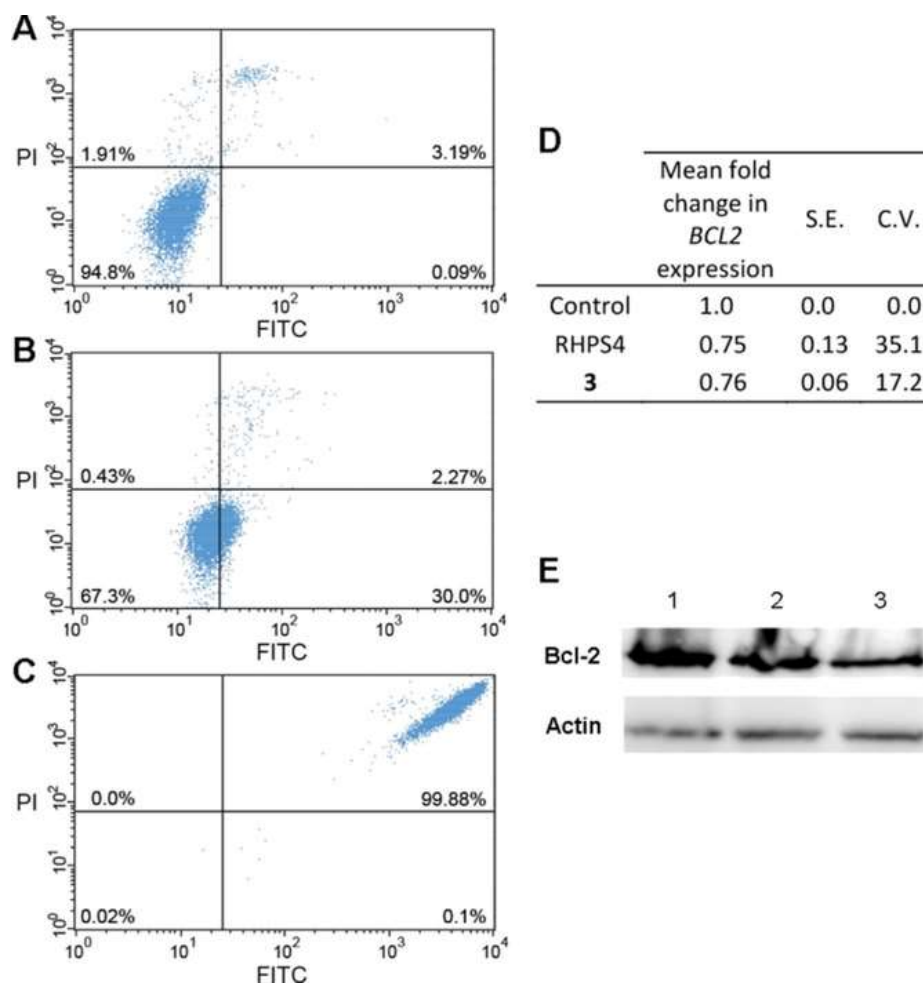


Figure 3.18. Apoptosis analysis with annexin V-FITC/PI double staining on Jurkat cells treated with A) vehicle, B) 3b and C) RHPS4 (25 μ M). For all panels: upper left quadrants, necrotic cells; upper right, advanced apoptotic cells; lower left, viable cells; lower right, early apoptotic cells. These results are representative of three independent experiments. D) qPCR analysis of *BCL2* mRNA levels in Jurkat cells. Results are presented as fold change in gene expression relative to the reference gene (*GAPDH*). Data were analyzed using the $2^{-\Delta\Delta C_t}$ method (Livak and Schmittgen, 2001). S.E.=standard error. C.V.=percent coefficient of variation of $2^{-\Delta\Delta C_t}$. E) Western blot analysis of lysates from Jurkat cells treated with (1) vehicle, (2) RHPS4, and (3) 3b (25 μ M) for 24 h. Specific anti-Bcl-2 antibody was used to detect Bcl-2 protein; actin was used as a control. The blot is representative of three experiments with similar results.

3.2.3 Conclusions

In this study, a series of FP derivatives were evaluated as G4-targeting compounds. Among them, two derivatives were shown to selectively stabilize the G4-forming sequence derived from the promoter region of the *BCL2* gene. The binding of the two selected compounds to the bcl-2 G4 was characterized by means of fluorescence and NMR titration experiments. Moreover, one of such ligands effectively down-regulated the transcription and expression of *BCL2* in Jurkat cells, inducing a significant increase of the percentage of cells in an early apoptotic state. All these experimental results reinforced the idea of inhibition of Bcl-2 expression through the specific targeting of bcl-2 G4 DNA and provided clear evidence that FP derivatives represent a new class of drug-like compounds useful to this aim, thus stimulating further studies aimed at developing more potent derivatives as effective pathway-specific anticancer agents.

3.2.4 Experimental Section

Oligonucleotide synthesis and sample preparation

The DNA sequences were synthesized using standard β -cyanoethylphosphoramidite solid phase chemistry on an ABI 394 DNA/RNA synthesizer (Applied Biosystem) at the 5 μ mol scale. DNA detachment from support and deprotection were performed by treatment with concentrated ammonia aqueous solution at 55 °C for 12 h. The combined filtrates and washings were concentrated under reduced pressure, dissolved in water, and purified by high-performance liquid chromatography (HPLC) on a Nucleogel SAX column (Macherey- Nagel, 1000–8/46), using buffer A consisting of 20 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ aqueous solution (pH 7.0), containing 20% (v/v) CH_3CN , buffer B consisting of 1 M KCl, 20 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ aqueous solution (pH 7.0), containing 20% (v/v) CH_3CN , and a linear gradient from 0% to 100% B for 30 min with a flow rate 1 mL/min. The fractions of the oligomers were collected and successively desalted by Sep-pak cartridges (C-18). The isolated oligomer was proved to be >98% pure by NMR. In particular, the following DNA sequences have been synthesized: the bcl-2 promoter sequence d(GGGCGCGGGAGGAATTGGGCGGG) (bcl-2) (Dai et al., 2006), the *KIT* oncogene promoter sequence d(AGGGAGGGCGCTGGGAGGAGGG) (c-kit1) (Phan et al., 2007b), the modified nuclease hypersensitivity element III1 sequence d(TGAGGGTGGGTAGGGTGGGTAAGG) of the *MYC* promoter (c-myc) (Ambrus et al.,

2005), the modified human telomeric sequence d(TTGGGTTAGGGTTAGGGTTAGGGA) (Tel24) (Luu et al., 2006a), and the self-complementary duplex forming dodecamer d(CGCGAATTCGCG) (ds12). DNAs were prepared in the appropriate buffer (bcl-2: 20 mM potassium phosphate, 40 mM KCl; c-kit1: 20 mM potassium phosphate, 50 mM KCl; c-myc and Tel24: 20 mM KH₂PO₄, 70 mM KCl; ds12: 10 mM Li₃PO₄, 100 mM KCl) at pH 7.0. The concentration of oligonucleotides was determined by UV adsorption measurements at 90 °C using molar extinction coefficient values ϵ ($\lambda = 260$ nm) calculated by the nearest neighbor model (Cantor et al., 1970). Samples were heated at 90 °C for 5 min, and then gradually cooled to room temperature overnight.

Circular dichroism spectroscopy

Circular dichroism (CD) experiments were recorded on a Jasco J-815 spectropolarimeter equipped with a PTC-423S/15 Peltier temperature controller. All the spectra were recorded at 20 °C in the wavelength range of 230–360 nm and averaged over three scans. The scan rate was set to 100 nm/min, with a 1 s response time and 1 nm bandwidth. Buffer baseline was subtracted from each spectrum. For the CD experiments, 10 μ M G4 DNAs and 15 μ M duplex were used. CD spectra of DNA/ligand mixtures were obtained by adding 4 mol equivalents of ligands (stock solutions of ligands were 10 mM in DMSO). CD melting experiments were carried out in the 20–100 °C temperature range at 1 °C/min heating rate, by following changes of CD signal at the wavelengths of the maximum CD intensity. CD melting experiments were recorded in the absence and presence of ligands (4 mol equivalents) added to the folded DNA structures. The melting temperatures (T_m) were determined from curve fit using Origin 7.0 software. ΔT_m values were determined as the difference in the melting temperature of DNA structures with and without ligands. All experiments were performed in triplicate and the values reported are average of three measurements.

Fluorescence Titration experiments

Fluorescence experiments were performed at 20 °C on a FP-8300 spectrofluorimeter (Jasco) equipped with a Peltier temperature controller accessory (Jasco PCT-818). A sealed quartz cuvette with a path length of 1 cm was used. Both excitation and emission slits were set at 5 nm. The titrations were carried out by stepwise addition (5 μ L) of DNA solutions (165 μ M) to a cell containing a fixed concentration (2.0–2.5 μ M) of ligand solution. Compounds **2b** and **3b**

were excited at 377 and 287 nm, respectively, and emission spectra were recorded in the range of 385-600 nm for **2b** and 325-600 nm for **3b**. After each addition of DNA, the solution was stirred and allowed to equilibrate for 5 min. The fraction of bound ligand (α) at each point of the titration was calculated following fluorescence changes at the maximum of intensity, as previously reported (Moraca et al., 2017; Pagano et al., 2015). Titration curves were obtained by plotting α versus the DNA concentration. The equilibrium binding constants (K_b) were estimated from this plot by fitting the resulting curve to an independent and equivalent binding site model (Moraca et al., 2017; Pagano et al., 2015). The experiments were repeated three times, and the obtained results are presented as the mean \pm S.D.

Nuclear Magnetic Resonance

NMR spectra were recorded on a Varian Unity INOVA 700 MHz spectrometer. One-dimensional proton spectra of samples in H₂O were recorded using pulsed-field gradient DPGSE for H₂O suppression. The bcl-2 DNA sample was prepared at a concentration of 0.2 mM in 0.6 mL (H₂O/D₂O 9:1) buffer solution containing 20 mM KH₂PO₄, 40 mM KCl, 0.2 mM EDTA, at pH 7.0. Aliquots of ligand stock solutions in deuterium-DMSO were added directly to the DNA solution inside the NMR tube. The NMR data were processed on iMAC running iNMR software (www.inmr.net).

Cell Lines and Culture Conditions

Human adenocarcinoma (HeLa), human liver cancer (HepG2), breast cancer (MCF-7) cell lines, and human normal dermal fibroblasts (HDF) were grown in DMEM supplemented with 10% fetal bovine serum (FBS), 1% glutamine, 100 U/mL penicillin, and 100 μ g/mL streptomycin (Euroclone, Italy). Human metastatic melanoma cell line (WM266) and human T lymphoblastoid cell line (Jurkat) (ATCC, USA) were grown in RPMI, supplemented with heat-inactivated 10% fetal bovine serum (FBS), 1% glutamine, 100 U/mL penicillin, and 100 μ g/mL streptomycin. The cells were maintained in humidified air containing 5% CO₂ at 37 °C.

Cell Proliferation Assay

WM266, MCF-7 and HepG2 cells were plated at a density of 1×10^4 cells/100 μ L, HeLa and HDF at density of 7×10^3 cells/100 μ L, on 96-well plates in medium 10% FCS. 10 μ M of each ligand, or DMSO as vehicle, were added to the cells. After 24 h incubation at 37 °C, cell

viability was assessed by performing MTT (Sigma Aldrich, Italy) reduction inhibition assay. Jurkat cells were seeded on a 96-well plate at a density of 1×10^4 cells/100 μ L and incubated for 24 h with 10 or 25 μ M of RHPS4 or compound **3b**, respectively. After incubation, 0.4% Trypan Blue was added to the cell suspension and cell numbers were estimated by counting under a light microscope. Cells stained blue were considered non-viable (Di Gaetano et al., 2001). Cytotoxicity experiments were independently performed at least three times. Statistical significance between groups was assessed by Student's t-test. Data are expressed as means standard error (S.E.) *p* values < 0.05 were considered to be statistically significant.

Apoptosis Assay

The apoptosis assay was performed on Jurkat cells seeded at 1×10^6 cells/mL in a 6-well plate. The cells were incubated at 37 °C with ligands at 10 or 25 μ M concentration and apoptosis was analyzed after 16 h by double staining with annexin V/FITC and Propidium iodide (PI) (eBioscience, USA) (Capasso et al., 2014, 2017). The percentage of cell undergoing apoptosis or necrosis was quantified using a flow cytometer equipped with a 488 nm argon laser (Becton Dickinson, USA) by Cell Quest software. All FACS analyses were performed at least 3 times.

Western Blotting

Jurkat cells were analyzed 24 h after treatment with compounds at 25 μ M concentration. Whole cell lysates were obtained by using lysis buffer (50 mM Hepes pH 7.4, 50 mM NaCl, 1% Triton) supplemented with protease inhibitor cocktail (Roche, Italy). Cell lysates were incubated on ice for 30 min and then centrifuged at 13,000 rpm for 30 min to remove cell debris. Protein concentrations were determined by Bradford method using Bio-Rad reagent (BioRad Laboratories, Italy). Proteins (100 μ g) were resolved by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to a PVDF membrane (Bio-Rad). The membrane was probed with the primary antibody (anti-Bcl-2 antibody from Bethyl, USA) o.n. at 4 °C. Proteins were visualized with an enhanced chemiluminescence detection system (Euroclone) and images were acquired with ChemiDoc XRS System (Bio-Rad Laboratories, Italy) and analyzed with the QuantityONE software.

RT-PCR and qPCR.

Total RNA was isolated from treated cells using TRI Reagent (Sigma Aldrich). Reverse transcription was performed using 0.5 µg of total RNA, 200 U of MMLV Reverse Transcriptase RNase H- (Euroclone), dNTPs, random primers, RNase inhibitor. Reaction temperature was set at 42 °C for 60 min. After reverse transcription, qPCR assay was carried out using the following primers: GAPDH forward d(AACGGGAAGCTTGTCATCAATGGAAA), GAPDH reverse d(GCATCAGCAGAGGGGGCAGAG), *BCL-2* forward d(CTGCACCTGACGCCCTTCACC), *BCL-2* reverse d(CACATGACCCCACCGAACTCAAAGA) (Sigma-Genosys, UK). The amplification reactions were run in duplicate. Real-time quantitative PCR (qRT-PCR) amplifications were performed using the SYBR Premix Ex Taq II (Takara, Japan) in Rotor-gene Q (Qiagen, Italy). The qRT-PCR protocol was as follows: 95 °C for 15 min followed by 40 cycles of 95 °C for 15 s, 59 °C for 30 s, and 72 °C for 30 s. Results were expressed as relative fold induction of the target genes relative to the reference gene. Calculations of relative expression levels were performed using the $2^{-\Delta\Delta C_t}$ method (Finetti et al., 2012; Livak and Schmittgen, 2001; Pane et al., 2016), and averaging the values of at least three independent experiments.

3.3 Tailoring a lead-like compound targeting multiple G-quadruplex

3.3.1 Introduction

Improving selectivity and concomitantly reducing toxicity of chemotherapeutic agents is one of the hottest challenges in anticancer research (Liu et al., 2017; Musumeci et al., 2016; Nussbaumer et al., 2011; Wheate et al., 2010). Understanding the molecular mechanisms involved in growth and proliferation of cancer has allowed developing therapies that efficiently act on specific cancer cell processes, not damaging normal cells and thus reducing typical side effects of anticancer drugs (Baudino, 2015; Gao et al., 2015; Huang et al., 2014; Musumeci et al., 2017b; Platella et al., 2017b; Viswanadha Vijaya Padma, 2015). Thus, compounds able to target G4 structures and discriminate duplex DNA can in principle act as effective candidate antitumor drugs by interfering with cancer onset and progression pathways without impairing healthy cells. In addition, by simultaneously targeting multiple G4 structures located in regulatory regions of different oncogenes or at telomeres, these binders could produce multiple inhibitory activities, promoted by a unique general mechanism of action, resulting into valuable synergistic effects.

With a view to expanding the repertoire of available drug-like G4 binders and speeding up the search for true hits, computer-aided virtual screening (VS) can be a valid tool, especially for analyzing large libraries of putative ligands (Ma et al., 2012). In this frame, some of us previously performed a virtual screening of a 6000 compounds library – obtained from a commercially available database – against the tetramolecular, parallel G4 formed by the d(TGGGGT) sequence (Cosconati et al., 2009). Among these compounds, six were found to be good G4 binders (Figure 3.19), as also confirmed by NMR studies. These ligands were subsequently studied in their interaction with different G4s using a variety of biophysical techniques (Pagano et al., 2015). Remarkably, two of these compounds (**2c** and **4c**, Figure 3.19) showed selective G4-stabilizing properties, as well as ability to induce DNA damage response and telomere-dysfunction-induced foci (TIFs) in the μM range. In particular, compound **4c** was shown to effectively stabilize both telomeric and extra-telomeric parallel G4s, also inducing G4 formation in some of the investigated G-rich oligonucleotides in the absence of cations (Pagano et al., 2015).

Encouraged by the promising activity of compound **4c**, here we identified and analyzed a small focused library of its structural analogs, featured by different pendant groups on the N-atom of the oxazine ring, with the aim of developing more potent and selective ligands

efficiently targeting multiple G4 structures. First, the selected molecules have been tested *vs.* topologically different G4s by means of the G4-CPG assay, an affinity chromatography-based method we have recently developed for the rapid screening of putative G4 ligands (Platella et al., 2018). Then, the G4-binding properties of the most promising analogs of this library – according to the results of the G4-CPG assay – have been investigated in detail employing both biophysical and biological tools.

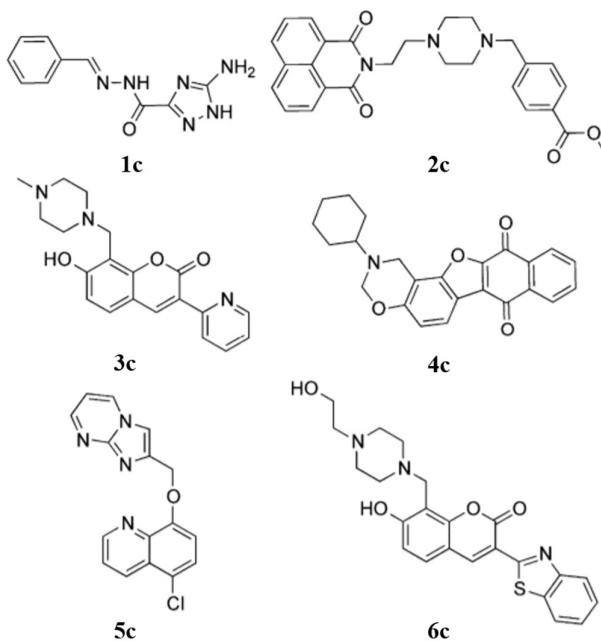


Figure 3.19. Chemical structures of compound 1c-6c (in order to avoid confusion, compounds 1-6 in Paper III here appear as 1c-6c) previously selected from a commercially available database by structure-based virtual screening.

3.3.2 Results and discussion

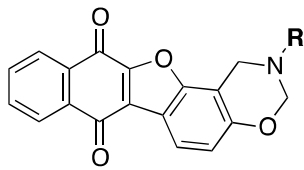
Selection of compound 4c analogs

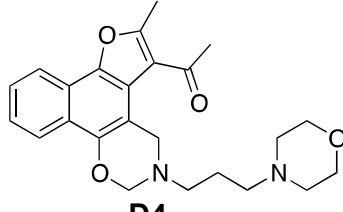
A convenient strategy for the discovery of G4-selective ligands with enhanced activity/toxicity ratio can be to start from a known lead compound featured by a suitable core, and then functionalize it with a set of different decorations. In this way, the interactions with the target G4s ensured by the core are preserved, and improved affinity (and/or selectivity) can be achieved by exploiting additional interactions (*i.e.*, electrostatic, stacking, and/or hydrogen bonding interactions) realized by the decorations. Following this strategy, using **4c** as the lead

compound we here aimed at identifying a set of its analogs featured by the furobenzoxazine naphthoquinone core and bearing different pendant groups on the oxazine N-atom.

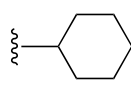
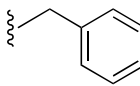

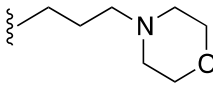
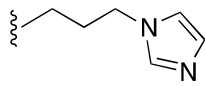
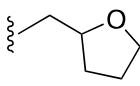
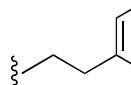
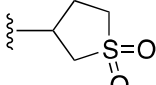
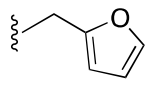
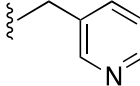
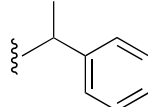
Given the relative synthetic accessibility of pentacyclic furobenzoxazine naphthoquinones, we first searched for analogs of **4c** available in commercial molecular databases. This approach has the advantage of quickly providing a library of structural analogs of the lead compound and being cost-effective. Therefore, the Dice similarity coefficient was computed between **4c** and the compounds present in the ZINC database collection of commercially available compounds (<https://zinc15.docking.org>) setting the similarity threshold to 70%. This search resulted in 40 compounds including, as desired, a number of furobenzoxazine-containing naphthoquinone derivatives, as well as species exhibiting different scaffolds. Ten furobenzoxazine naphthoquinone derivatives obtained from the filtered database were thus selected (**S4-1–S4-10**, Table 3.5), along with one furonaphthoxazine analog (**D4**, Table 3.5), included in this study as a representative example of a different scaffold.

Table 3.6. Chemical structures of the here investigated compounds.





D4

Compound	R	Compound	R
4c		S4-6	
S4-1		S4-7	
S4-2		S4-8	
S4-3		S4-9	
S4-4		S4-10	
S4-5			

G4-CPG assay and selection of the model G4- and duplex-forming oligonucleotide sequences

We recently realized a functionalized CPG support, designed for the High Throughput Screening of G4 ligands by affinity chromatography, named G4-CPG (Platella et al., 2018; Platella and Al.). This support has been conceived to allow the on-line synthesis of the G4-forming oligonucleotides, as well as the successive binding assays with the potential ligands, providing quick, reliable results. Indeed, in contrast to commercially available supports previously used for affinity chromatography of oligonucleotides, (http://www.glenresearch.com/Technical/TB_OAS.pdf; Musumeci et al., 2014, 2017a) the new derivatized CPG has low-to-null unspecific interactions with G4 ligands. Thus the binding assays on the G4-CPG provide clean and reproducible data, well reflecting the affinity trend of the tested molecules *vs.* G4 targets in solution (Platella et al., 2018).

In brief, the G4-CPG assay is based on affinity chromatography screenings, realized on a suitably modified CPG support functionalized with the G-rich oligonucleotide of interest, reproducing the target G4-forming DNA tract. Being bound to the CPG support *via* a long and flexible linker, it can fold in its preferred conformation under various solution conditions and retain high-affinity ligands, thus allowing their easy identification. Indeed, using a fluorescent ligand differently responding to the interaction with various G4 topologies, (Zuffo et al., 2017) we already proved that the G4-forming oligonucleotides linked to the CPG support preserve their native conformation as in solution (Platella et al., 2018).

Thus, aiming at expanding the use of the G4-CPG assay to a wide variety of topologically different G4 structures, we here functionalized the CPG support with a set of biologically relevant G4-forming oligonucleotides. In detail, we selected the following DNA sequences: *a*) the 26-mer d[(TTAGGG)₄TT] (Tel26) as a well-studied model of the 3'-overhang of the human telomeric DNA (Dai et al., 2007); *b*) the 33-mer d(TGGGGAGGGTGGGGAGGGTGGGGAAGGTGGGGA) (c-myc) from the nuclease hypersensitive element in the human *MYC* promoter, containing six G-rich tracts of unequal length and characterized by a remarkable polymorphism (Simonsson et al., 2000; Sun and Hurley, 2009), of which typically only shorter and modified variants have been used as models (Mathad et al., 2011); *c*) the 22-mer d(AGGGAGGGCGCTGGGAGGAGGG) (c-kit1) - of which both the NMR (Phan et al., 2007a) and crystal structures (Wei et al., 2012, 2015) have been solved revealing a unique topology - and the 21-mer d(CGGGCGGGCGCGAGGGAGGGG) (c-kit2) - which, depending on the K⁺ concentration,

adopts two distinct parallel-stranded G4 scaffolds in solution (Kuryavyi et al., 2010) - both from the human *KIT* promoter; *d*) the 24-mer d(AAGGGGAGGGGCTGGGAGGGCCCCGGA) (hTERT1) from the G-rich region in the *hTERT* core promoter, which represents the main regulatory element of the *hTERT* gene (Chaires et al., 2014; Palumbo et al., 2009). Furthermore, in order to probe the ability of the studied compounds to discriminate G4 vs duplex DNA, the CPG support has been also functionalized with a 27-mer here named ds27 (Musumeci et al., 2017a; Platella et al., 2018). Consisting of two Dickerson 12-mer tracts, d(CGCGAATTCGCG), *i.e.* one of the best characterized models of B-DNA duplex, (Drew et al., 1981) connected by a TTT loop, ds27 can in fact fold into a stable hairpin duplex and has been thus selected as a representative model for a duplex DNA tract.

Experimental screenings by the G4-CPG assay

The binding assays on the functionalized CPG supports have been performed following previously reported procedures (Musumeci et al., 2014, 2017a; Platella et al., 2018), with minimal modifications as described in the Experimental section in this Chapter. All the oligonucleotide-functionalized supports have been first tested in their ability to bind a set of known G4 ligands with different affinity for G4s. In particular, thiazole orange (TO) and resveratrol have been used as representative molecules, being respectively a strong and a very weak G4 binder (Monchaud et al., 2008; Platella et al., 2017a). The obtained data, showing high affinity (bound ligand > 97%) for TO and very low affinity (bound ligand < 4%) for resveratrol on all the tested supports, verified the efficacy of the functionalized CPG supports.

A stock solution of each ligand has been prepared by dissolving a weighed amount of the solid compound in pure DMSO; only compound **S4-1** proved to be not completely soluble in DMSO at the concentration required for the binding assays and was therefore discarded. All the other compounds as well as compound **4c** proved to be well soluble in the washing/releasing solutions used in the binding assays (see Experimental Section of this Chapter) and at the concentration chosen for the binding experiments. After solubility checks, we verified the absence of unspecific binding on the solid support by incubating the tested compounds with the non-functionalized CPG (here named nude CPG, Figure 3.20). Low-to-null unspecific interactions with the nude solid support have been observed (Table 3.7), thus further proving that nude CPG is essentially inert in these assays, as demonstrated for known G4 ligands, (Platella et al., 2018) and confirming the general reliability of the method.

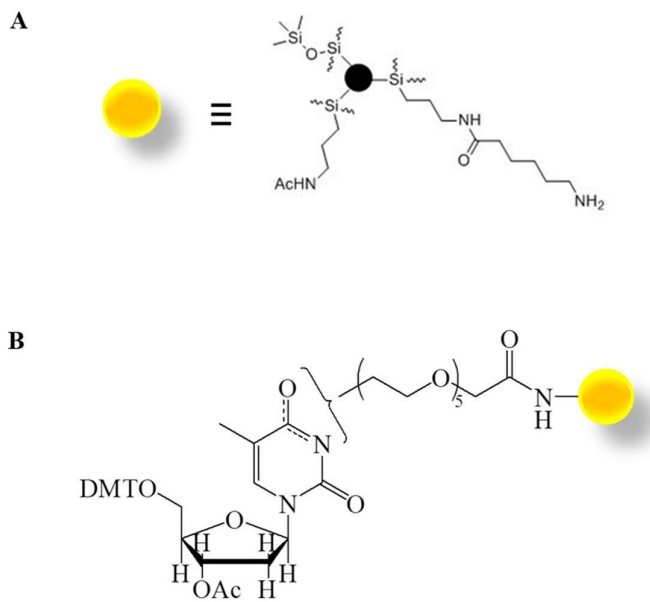


Figure 3.20. Chemical structures of: (A) commercially available Long Chain AlkylAmine-CPG (LCAA-CPG) and (B) LCAA-CPG functionalized with 5'-O-DMT, 3'-O-acetyl-thymidine through a hexaethylene glycol spacer (nude CPG).

Finally, all the new putative ligands have been tested on the G4s- and hairpin duplex-functionalized CPG supports. The results of the binding assays, summarized in Table 3.7, showed that all the analyzed ligands had good affinity for Tel26 immobilized on the solid support. However, no marked improvement was found with respect to **4c**, still the best binder in the series when tested on this G4-forming sequence, with the only exception of **S4-4**, showing comparable results as **4c**. In contrast, when tested on c-myc, compounds **S4-2**, **S4-4**, **S4-7**, **S4-8** and **S4-10** displayed higher affinity than **4c**. As far as c-kit1 is concerned, almost all the tested compounds proved to be stronger binders than compound **4c**, with the sole exception of **S4-3**. In the case of c-kit2, **S4-2**, **S4-4**, and **S4-10** showed higher binding abilities than **4c**. Notably, also in this case compound **S4-3** was the weakest binder of the series. Finally, for hTERT1 all the tested analogs had comparable or higher affinity than **4c**.

Noteworthy, all the examined compounds were found to be effective G4 ligands, some of which even better than lead compound **4c**. Among all the tested compounds, **D4** was the worst G4 ligand, evidencing that its polycyclic core, different from the other tested molecules, is not suitable for G4 targeting. On the other hand, compound **S4-4** was the strongest G4 binder of the series, able to tightly interact with all the investigated G4s. Overall, although the here screened compounds are able to bind all the investigated G4s, thus behaving as multi-target

ligands, the highest percentages of binding have been found for c-myc, emerging as the best target of the series.

In order to evaluate the G4s vs. duplex DNA selectivity, all the compounds have been also tested in their interaction with ds27, *i.e.* the hairpin duplex-forming oligonucleotide. Overall, most of the ligands proved to effectively discriminate G4- vs. duplex-forming oligonucleotides. In particular, **S4-4** and **S4-5** showed a significantly lower affinity for ds27 than **4c**. Remarkably, **S4-5** gave null affinity for duplex DNA, thus emerging as the most promising analog in terms of G4 vs. duplex selectivity on the basis of this binding assay.

Table 3.7. Binding assay results for the investigated compounds on nude and functionalized CPG supports.

Compound	Bound ligand (%) ^a						
	Nude CPG	CPG-Tel26	CPG-c-myc	CPG-c-kit1	CPG-c-kit2	CPG-hTERT1	CPG-ds27
4c	0	86	86	70	82	74	18
S4-2	0	80	94	85	91	82	46
S4-3	0	73	84	55	67	73	16
S4-4	0	86	93	90	92	91	2
S4-5	0	80	84	78	80	82	0
S4-6	10	80	87	76	83	78	12
S4-7	11	80	89	77	84	92	33
S4-8	0	78	90	88	83	87	20
S4-9	5	75	73	72	68	76	16
S4-10	0	79	94	90	88	89	28
D4	7	76	74	75	70	78	22

^aBound ligand is calculated as a difference from the unbound ligand, recovered with 50 mM KCl/10% DMSO/10% ethanol washing solution, and expressed as % of the amount initially loaded on the support. Errors associated with the % are within $\pm 2\%$.

Biological activity studies

Following the binding assays, the G4 ligands have been subjected to biological analyses aimed at identifying derivatives with potent and selective DNA damage response (DDR) and antitumoral activity. Thus, BJ-EHLT cells – a human foreskin-derived fibroblast cell line, expressing the human telomerase reverse transcriptase (hTERT) and SV40 early region – were treated for 24 h with 0.5 μ M of each compound and the DNA damage was evaluated. To appreciate possible improvements in terms of biological activity of the analogs over lead compound **4c**, we used a time and dose that had been previously optimized for **4c**, able to induce a slight DNA damage (Pagano et al., 2015). Immunofluorescence (IF) analysis showed relevant

amounts of the phosphorylated form of histone H2AX (γ H2AX), a hallmark of DNA double-strand breaks,(Thiriet and Hayes, 2005) evidencing that all the tested ligands – with the sole exception of **D4**, **S4-9**, and **S4-10** – induced similar DNA damage extent as the lead compound (Figure 3.21A). In parallel, we also evaluated whether the induced DNA damage – or at least part of it – was located at the telomeres. To address this issue, we performed co-staining IF experiments using antibodies against γ H2AX and TRF1 – an effective marker for interphase telomeres – and the co-localization spots, reported as Telomere Induced Foci (TIF), (Takai et al., 2003) were evaluated (Figure 3.21B).

Notably, quantitative analysis confirmed and reinforced the previous observations, evidencing that all the compounds (with the only exception of **D4**, **S4-9**, and **S4-10**) induced a percentage of TIF positive cells (cells with at least 4 co-localization spots) and average number of TIFs per cell similar to lead compound **4c** (Figure 3.21C). These data are consistent with the comparable affinities of compound **4c** and its analogs towards the telomeric model sequences found using the G4-CPG assay.

Next, to test the antitumor efficacy of compound **4c** analogs, the cell colony-forming ability of the human cervical cancer cells, HeLa, untreated or treated with the different compounds, was evaluated. To set up the optimal dose of the drug, we first performed a dose-response experiment by using three different concentrations (0.5, 1 and 2 μ M) of **4c**; the 2 μ M concentration proved to be the minimal effective dose and was thus selected for testing all the other derivatives (Figure 3.22A). Interestingly, cell survival experiments evidenced that derivatives **S4-4**, **S4-6**, **S4-8**, and **S4-10**, despite their high affinity for G4 structures *in vitro* (Table 3.7), were less effective than **4c** in affecting tumor cell viability (Figures 3.22B and C). Compounds **D4** and **S4-3** also showed an antiproliferative activity lower than compound **4c** (Figures 3.22B and C), which could be related, in this case, to an overall G4-affinity lower than compound **4c** (Table 3.7).

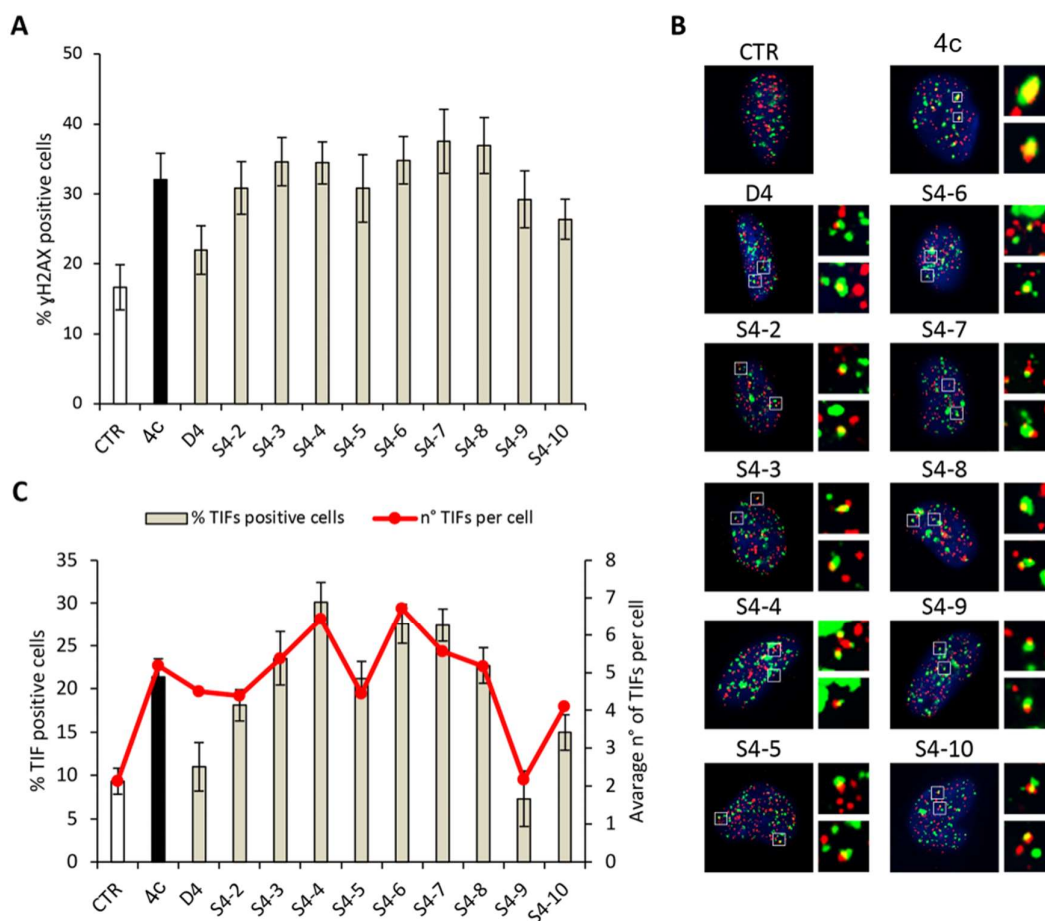


Figure 3.21. Compound 4c analogs induce telomeric DNA damage. BJ-EHLT fibroblasts were untreated (CTR, white bar) or treated for 24 h with 4c (black bar) and the indicated analogs (light-grey bars) at 0.5 μ M concentration. Cells were processed for immunofluorescence (IF) using antibodies against γ H2AX and TRF1 to visualize the DNA damage and telomeres, respectively. (A) Percentages of γ H2AX-positive cells. (B) Representative merged images of IF of untreated and treated BJ-EHLT cells; γ H2AX spots in green, TRF1 spots in red and nuclei in blue. Enlarged views of Telomere Induced Foci (TIFs) are reported on the right panels of each picture. The images were acquired with a Leica Deconvolution microscope (magnification 63x). (C) Quantitative analysis of TIFs. The graph represents the percentages of TIF-positive cells (bars) and the mean number of TIFs for cell (red line) in the indicated samples. Cells with at least four γ H2AX/TRF1 foci were scored as TIF positive. Histograms show the mean values \pm S.D. of three independent experiments.

Conversely, the potent cytotoxic activity observed for derivatives **S4-2**, **S4-7** and **S4-9** (Figures 3.22B and C) might be attributed to G4-independent, off-target effects since these ligands have shown either a good affinity also for duplex DNA (**S4-2** and **S4-7**) or, in the case of **S4-9**, an affinity towards G4 structures generally lower than other analogs (Table 3.7). In summary, ligand **S4-5** emerged as the most interesting ligand of the series. Indeed, its high affinity for G4 - despite an antitumor activity only slightly superior to lead compound **4c** -

associated with null affinity for duplex DNA (Table 3.7) suggests a promisingly improved target specificity compared to the starting compound.

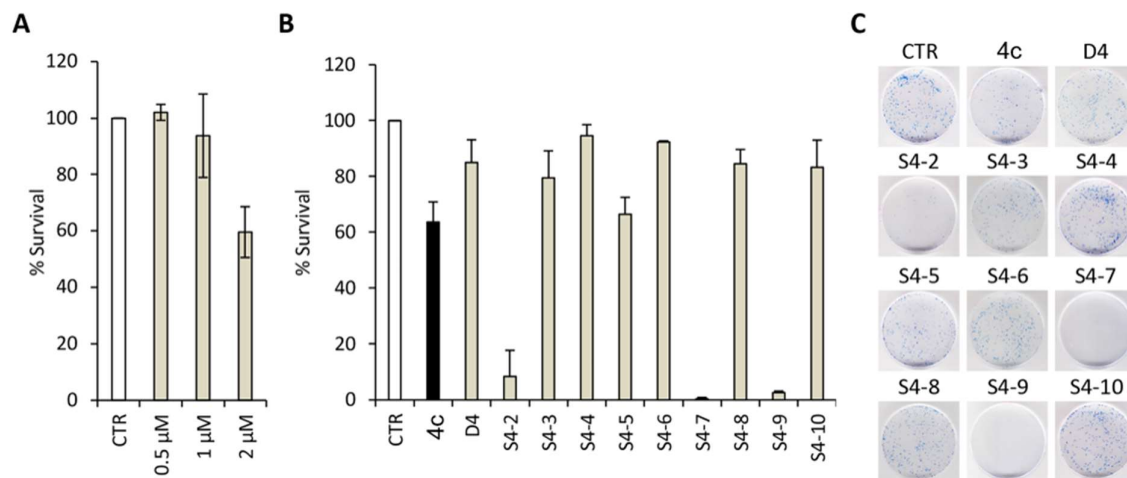


Figure 3.22. Anti-tumor efficacy of compound 4c analogs. (A) Clonogenic activity of human cervical cancer cells, HeLa, untreated (CTR, white bar) or treated with 4c at the indicated doses. Surviving fractions were calculated as the ratio of absolute survival of the treated sample/absolute survival of the untreated sample. (B) Clonogenic activity of HeLa cells, untreated (CTR, white bar), treated with 4c (black bar) or the indicated analogues (light-grey bars) at 2 μM dose. Surviving fractions were calculated as reported in A. (C) Representative images of the clonogenic assay described in B. Histograms show the mean values ±S.D. of three independent experiments.

Next, in order to evaluate the selectivity of **S4-5** in cell, the effects of this compound have been tested also on normal cells. Briefly, human immortalized BJ fibroblasts (BJ-hTERT) have been treated with **4c** or **S4-5** at 1 and 2 μM concentrations and, after 6 days, the number of viable cells has been evaluated. Notably, these experiments clearly evidenced that the cytotoxic effects of ligand **S4-5** on normal cells are definitely lower than those produced by compound **4c** (Figure 3.23A). Moreover, IF analysis of γH2AX evidenced that **S4-5**, in contrast to **4c**, was unable to induce an appreciable DNA damage in normal cells (Figures 3.23B and C). Overall, these data unambiguously indicated that **S4-5** has a much higher selectivity in killing cancer cells than compound **4c**, making it a promising candidate drug.

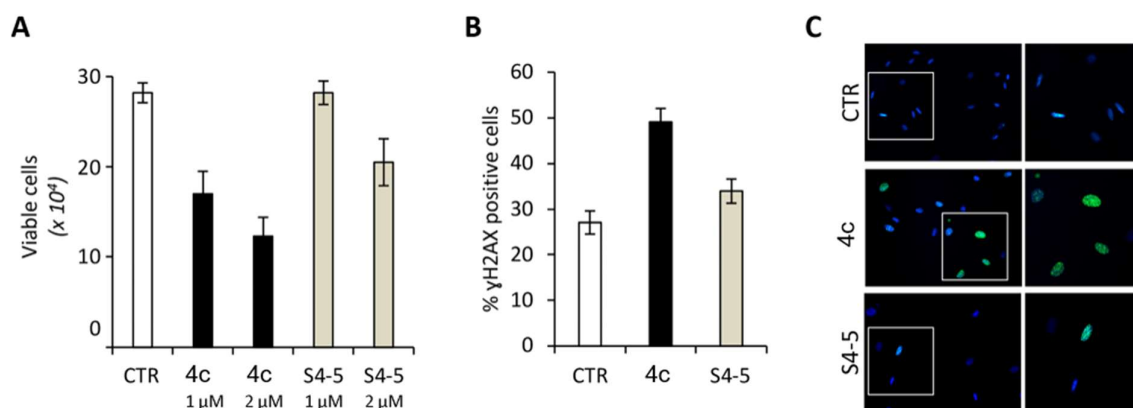


Figure 3.23. S4-5 analog does not affect normal cells. (A) Human immortalized fibroblasts (BJ-hTERT) were treated with compound 4c or derivative S4-5 at the doses of 1 and 2 μ M for 6 days. Viable cell number was determined by the Trypan Blue exclusion test. (B) BJ-hTERT were treated with 4c or S4-5 at the 1 μ M dose for 24 h and processed for IF using antibodies against γ H2AX. The histogram represents the percentages of γ H2AX-positive cells. (C) Representative merged images of IF: γ H2AX spots in green and nuclei in blue. Histograms show the mean values \pm S.D. of three independent experiments.

CD and CD-melting experiments

Based on the above discussed results, we investigated by CD experiments in solution the ability of **S4-5** to interact with a human telomeric G4 (Tel26), an extra-telomeric G4 (c-myc), and a duplex structure (ds27), in comparison with **4c**. In particular, among the extra-telomeric G4s here studied, we chose c-myc since **S4-5**, as well as the other here tested ligands, has shown a slight preference for this G4 structure compared to the others.

CD-monitored titrations have been performed by adding increasing amounts of the tested ligands to samples of these oligonucleotides (2 μ M) in 20 mM KCl - 5 mM KH_2PO_4 - 10% buffer (pH 7). As expected in these buffer conditions, Tel26 folded into a hybrid II-type G4 featured by a maximum at 290 nm and a shoulder at 270 nm (Karsisiotis et al., 2011), c-myc formed a parallel G4 with a maximum at around 263 nm (Mathad et al., 2011; Sun and Hurley, 2009), while ds27 showed the characteristic positive band at 280 nm accompanied by a minimum at 251 nm, typical of a B-DNA duplex structure (Figure 3.24) (Kypr et al., 2009). In the case of Tel26, an increase of the intensity of the 290 nm band along with the reduction of the shoulder at 270 nm have been observed upon adding increasing amounts of **S4-5** in solution (Figure 3.24A), thus suggesting that some interaction took place. In the titration experiments with c-myc, the addition of **S4-5** produced only weak changes, with a slight reduction of the 263 nm band and an increase of the 288 nm band (Figure 3.24B).

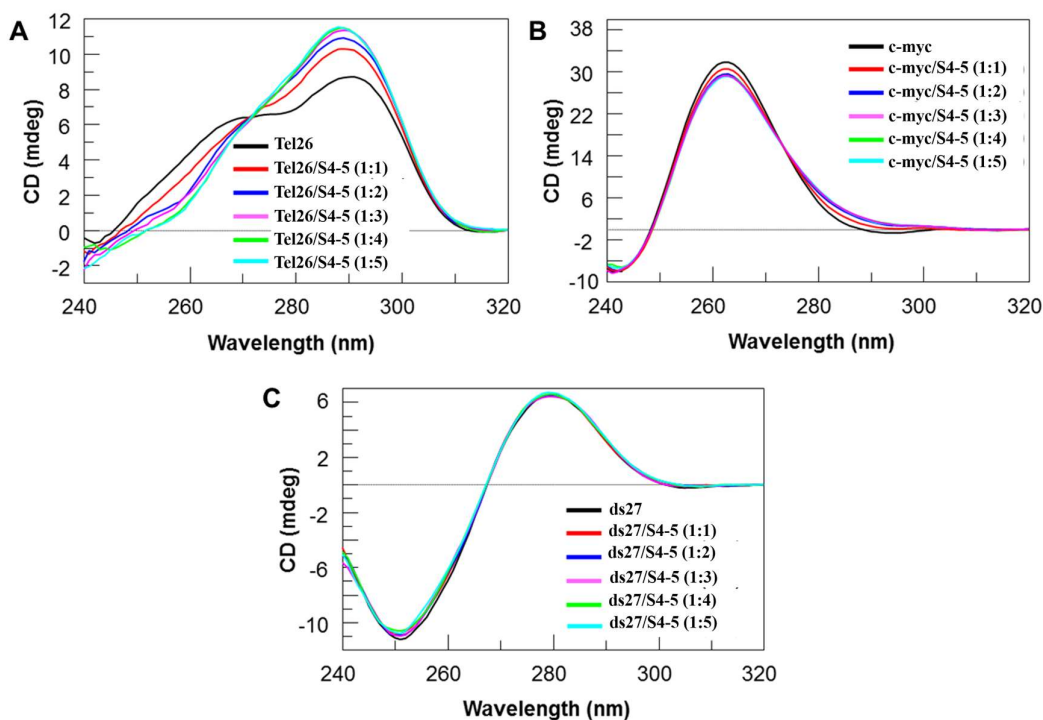


Fig. 3.24. Effect of S4-5 on the overall conformation of the DNA structures. CD spectra of 2 μ M solutions of Tel26 (A), c-myc (B) and ds27 (C) in 20 mM KCl - 5 mM KH_2PO_4 - 10% DMSO buffer (pH 7) in the absence and presence of increasing amount of S4-5 (up to 5 mol equivalents).

In contrast, in the titration of ds27 no detectable change in the CD profile of the duplex has been observed, suggesting that this structure was essentially unaffected by the addition of even a large excess of ligand (Figure 3.24C).

A similar behavior has been found in the titrations of both the G4s and the duplex with **4c** (Figure 3.25). In all cases, up to 5 equivalents of ligand have been added to the oligonucleotide solutions, as the best compromise between solubility of the analyzed ligands and saturation of the oligonucleotide CD signals. No induced CD signal has been detected for all the investigated systems in the range 320-800 nm (Figure 3.25).

CD-melting experiments have been performed for all the oligonucleotide/ligand mixtures to evaluate if stabilizing or destabilizing effects on the G4 and hairpin duplex structures were obtained upon incubation with **4c** and **S4-5**. CD melting curves of Tel26, c-myc and ds27 in the absence or presence of each ligand (at 5:1 ligand/DNA ratio) have been recorded by following the CD signal changes at the wavelength of their intensity maximum (290, 263, and 251 nm for Tel26, c-myc and ds27, respectively).

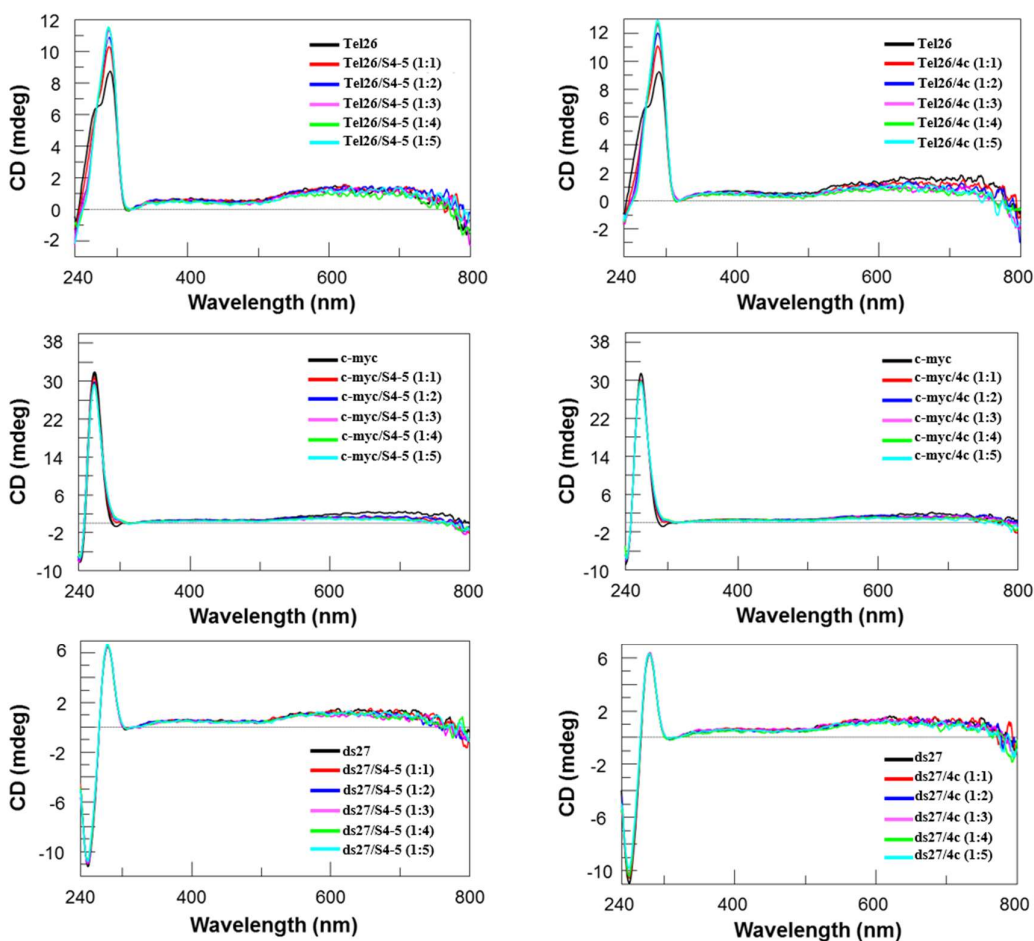


Figure 3.25. CD titrations of Tel26, c-myc and ds27 (top, middle and bottom, respectively) with increasing amounts of S4-5 (left panels) and 4c (right panels).

The results of the CD melting experiments showed that compounds **S4-5** and **4c** did not significantly affect the stability of the Tel26 and ds27 structures (Table 3.8), even if for Tel26 the drop of CD signal between 20 and 90 °C was much higher in the presence of the ligands than in their absence. Conversely, both ligands appreciably stabilized the c-myc G4 structure ($\Delta T_m > 8$ °C).

Overall, these results confirmed the ability of the here investigated pentacyclic scaffold to significantly stabilize parallel G4s, not affecting the thermal stability of other G4 conformations and duplex DNA (Pagano et al., 2015).

Table 3.8. Melting temperatures (T_m) of tel26, cmyc and ds27 in the absence or presence of ligands (5 equiv) measured by CD melting experiments.

Ligand	T_m (°C) (± 1)		
	Tel26	c-myc	ds27
No ligand		82	75
4	49	>90	76
S4-5	50	>90	76

NMR experiments

NMR titration experiments were performed to get information on the binding mode of compounds **4c** and **S4-5** to the G4 structures formed by the human telomeric and *MYC* promoter DNA sequences (Pagano et al., 2012). The modified telomeric sequence d[TTGGG(TTAGGG)3A] (Tel24) and the modified c-myc sequence d(TGAGGGTGGGTAGGGTGGGTAA) (myc22) were used for this study, since they give higher quality NMR spectra compared to the wild-type sequences, being characterized by a single conformation in solution (Ambrus et al., 2005; Luu et al., 2006a). The ^1H -NMR spectra of Tel24 and myc22 well matched the ones reported in the literature, characterized by 12 well-resolved imino protons peaks, corresponding to the 12 guanines involved in the three G-tetrad planes (Figures 3.26A-B) (Ambrus et al., 2005; Luu et al., 2006a).

The investigated oligonucleotides were then titrated with **4c** and **S4-5**, in parallel experiments, up to a 4:1 ligand/G4 ratio. Upon addition of increasing amounts of ligands to the G4 solutions, considerable proton resonance changes were observed in both imino and aromatic proton regions of the spectra (Figures 3.26 and 3.27), thus suggesting that both **4c** and **S4-5** may not have a unique binding site. The relative line broadening observed for all the imino resonances and most of the aromatic ones also suggests that the ligands could explore different poses in the same G4 binding pocket, in fast exchange with each other. On the other hand, some distinct aromatic protons of both Tel24 and myc22 G4s were only weakly affected by the ligand. As far as **4c** is concerned, for Tel24 (Figure 3.26A) this is the case of residues T6-T7-A8, forming the double-chain reversal loop of the G4; while for myc22 (Figure 3.26B) this is the case of A6, stacking on top of the 5' G-tetrad, and of residues T10, T14-A15, and T19 (numbering according to (Ambrus et al., 2005), which form the three double-chain reversal loops of the G4 structure. Regarding **S4-5**, the NMR titrations showed that in the interaction with Tel24 G4 (Figure 3.27A) the ligand affected to a lesser extent the aromatic protons of A8,

which is in the double-chain reversal loop, and of residues T18-T19-A20, which form a lateral loop close to the external G-tetrad; while in the interaction with myc22 (Figure 3.27B), the least affected ones were the aromatic protons of A6, which stacks on top of the 5' G-tetrad, and of residues T14-A15 and T19, which form the second and third double-chain reversal loop of G4, respectively.

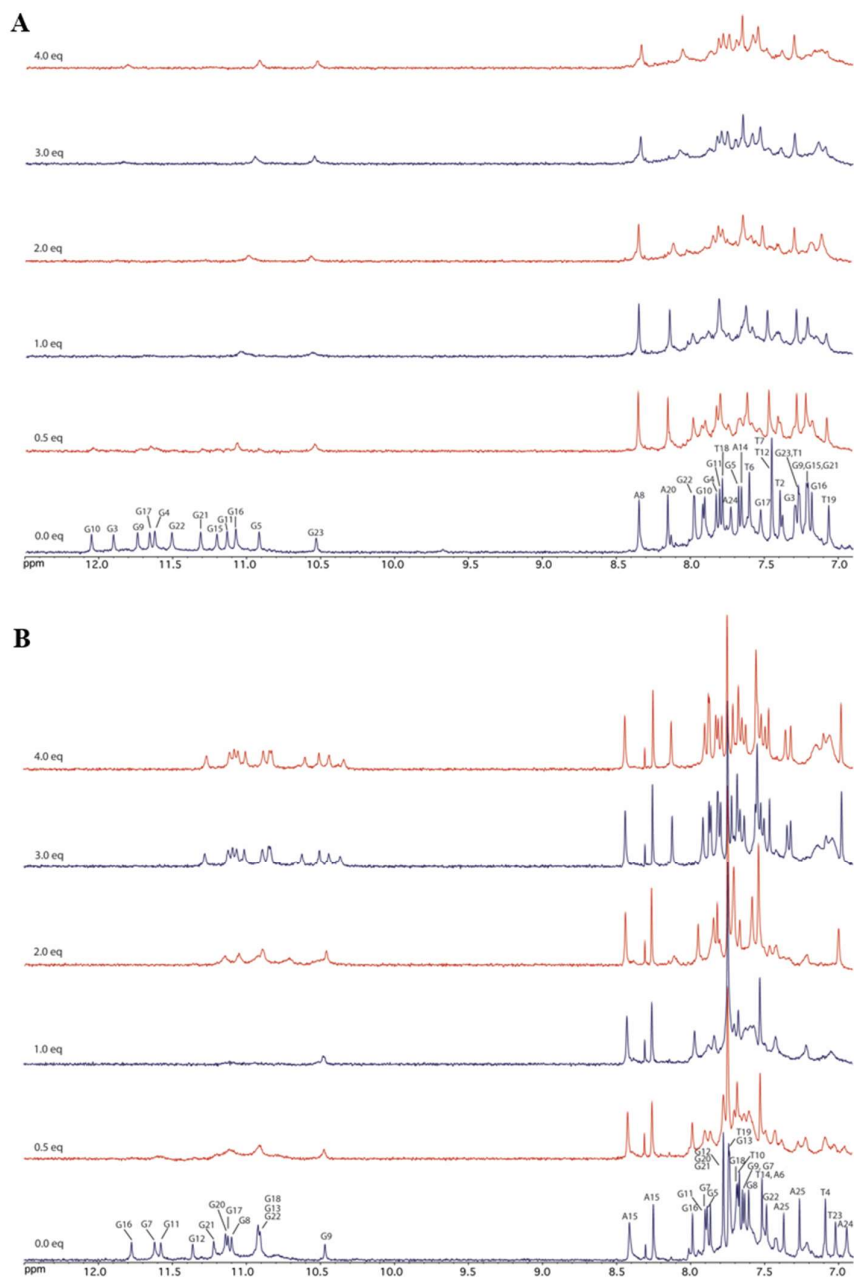


Figure 3.26. Imino and aromatic regions of the ^1H NMR spectra of (A) Tel24 and (B) myc22 G4s titrated with 4c.

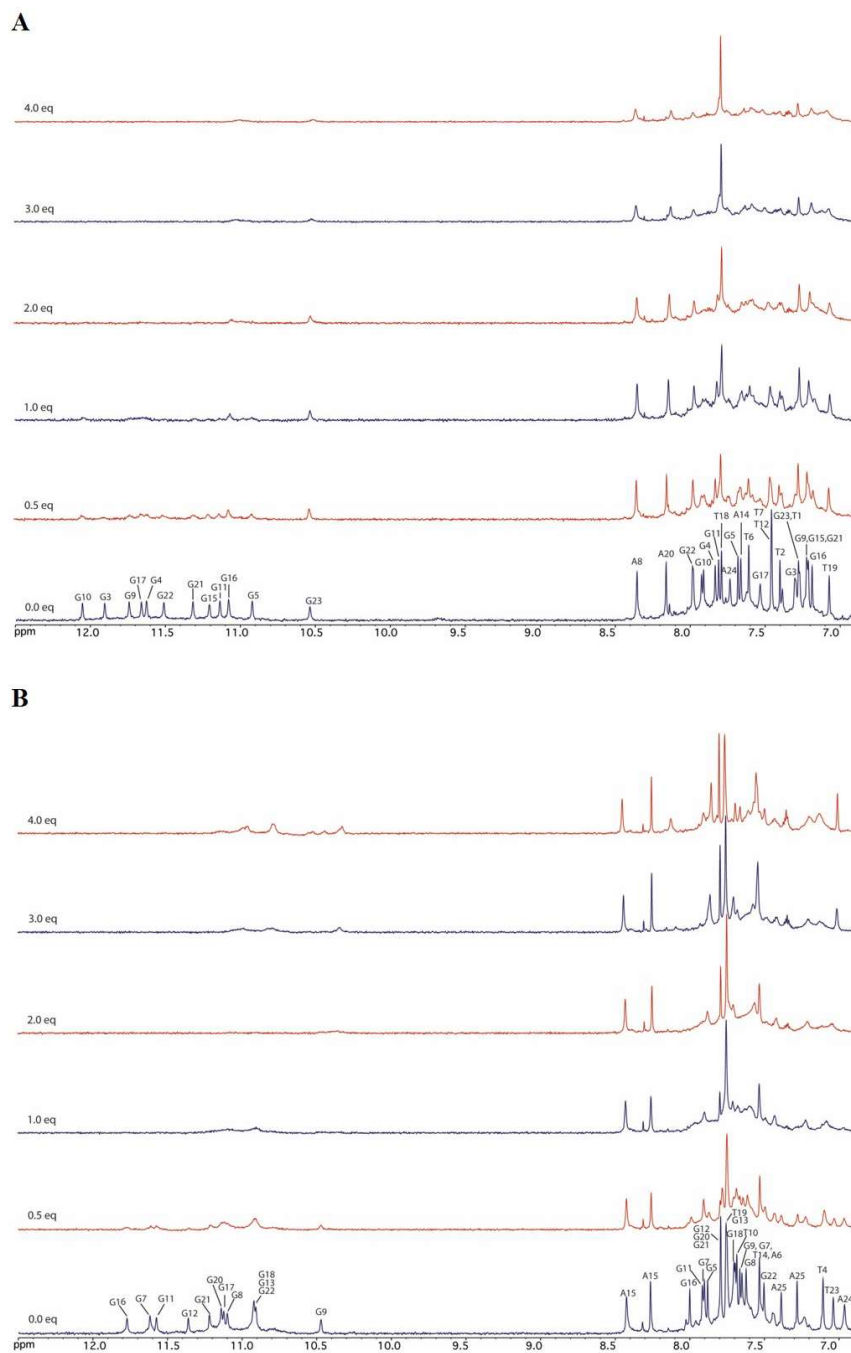


Figure 3.27. Imino and aromatic regions of the ^1H NMR spectra of (A) Tel24 and (B) myc22 G4s titrated with S4-5.

Overall, these findings suggest that both **4c** and **S4-5** may bind these G4s mainly through interaction with those grooves that are not made inaccessible by the presence of loops, generally affecting also guanine residues involved in the three G-tetrad layers. However, our data do not

allow completely excluding ligand stacking on external G-tetrads, while they clearly indicate that loops are generally very little or not affected at all by ligand binding.

NMR experiments helped us to further investigate the selectivity of **4c** and **S4-5** for G4 over duplex DNA. The two ligands were tested on a model duplex DNA formed by the so-called Dickerson sequence d(CGCGAATTCGCG) (ds12). The $^1\text{H-NMR}$ spectrum of ds12 matched the one reported in the literature (Figures 3.28A-B) (Jaroszewski et al., 1996). Results of ds12 titrations with **4c** and **S4-5** clearly showed that no significant variation of the chemical shift values of ds12 was observed upon addition of **S4-5** (Figure 3.28B), up to a 4:1 ligand/DNA ratio, thus confirming the absence of binding with the investigated duplex DNA. On the other hand, **4c** caused a small perturbation of the ds12 spectrum, in terms of both intensity decrease and line broadening of all the $^1\text{H-NMR}$ signals (Figure 3.28B), thus suggesting the possibility that **4c** could somehow interact in a weak and unspecific way with duplex DNA. These results are in full agreement with those obtained from the G4-CPG binding assays.

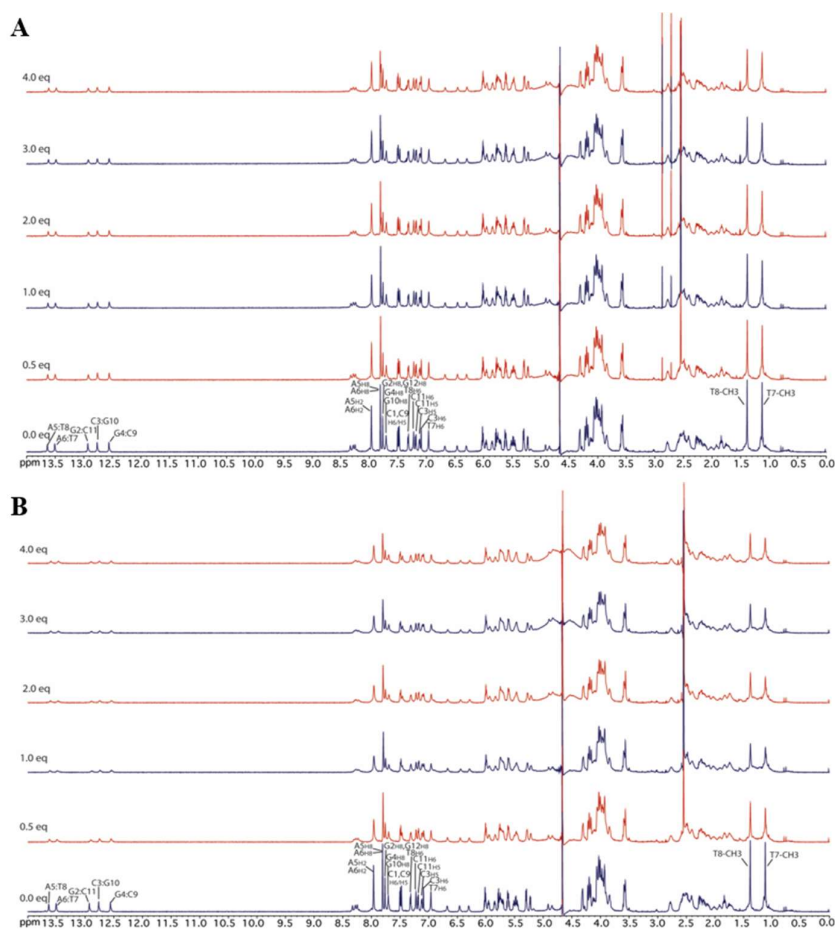


Figure 3.28. $^1\text{H-NMR}$ spectra of the ds12 duplex titrated with (A) **S4-5** and (B) **4c**.

Microscale thermophoresis (MST) experiments

To obtain quantitative data about the affinity of **S4-5** for the human telomeric and *MYC* promoter G4s, MST measurements were performed. MST is a fast and easy tool to characterize small molecule-nucleic acid interactions in solution (Entzian and Schubert, 2016). Basically, it records the thermophoretic movement of a fluorescently-labeled target molecule under microscopic temperature gradients. This molecular motion strongly depends on changes in size, charge, and hydration shell. Since the binding of a ligand to the investigated molecule changes at least one of these parameters, it also alters the thermophoretic behavior of the target. This effect can be used to evaluate equilibrium constants, such as the dissociation constant K_d . To this purpose, serial dilutions of **S4-5** were prepared, mixed with a constant concentration of Cy5-labeled oligonucleotides (Tel26 or myc22), loaded into capillaries and analyzed by MST. Experiments were run by using 8% final DMSO, in order to avoid aggregation effects observed in the absence of DMSO. The results of these experiments (Figure 3.29) showed that compound **S4-5** was able to bind both Tel26 and myc22 G4s. The calculated equilibrium dissociation constants indicated for this ligand ca. twice higher affinity to the parallel *MYC* promoter G4 [$K_d = 13 (\pm 2) \mu\text{M}$] than to the hybrid human telomeric G4 [$K_d = 26 (\pm 4) \mu\text{M}$].

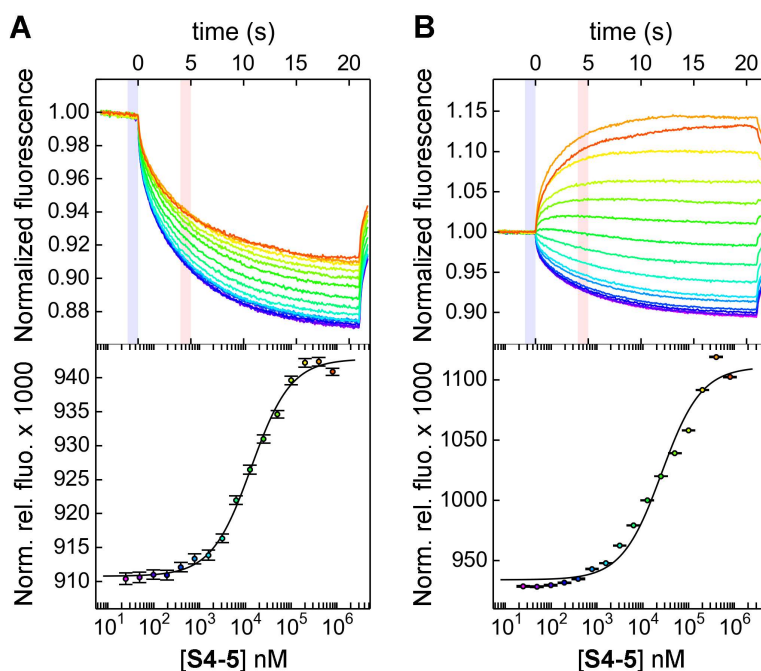


Figure 3.29. Interaction of **S4-5** with (A) myc22 and (B) Tel26 G4s studied using MST experiments. (Top) Time traces recorded by incubating increasing concentrations of **S4-5** with the labeled G4s and (bottom) the corresponding binding curves.

Molecular docking.

To gain major insights into the binding interactions established by **S4-5** with the target G4 DNA structures, we decided to provide a theoretical model for its interaction with human telomeric and *MYC* promoter G4-forming sequences. In particular, the ligand was docked into the solution structures of c-myc G4 (myc22, PDB ID: 1XAV) (Ambrus et al., 2005) and into the hybrid-type G4 structure formed in K⁺ solution by the 26-mer oligonucleotide from the human telomere (Tel26, PDB ID: 2JPZ) (Dai et al., 2007). In this study, the software Autodock 4.2 (AD4.2) (Cosconati et al., 2010; Morris et al., 1998) was employed, which had already been used for our previous virtual screening campaign to identify **4c** as a G4 ligand (Cosconati et al., 2009).

Analysis of the results obtained when **S4-5** was docked into the myc22 structure revealed that the ligand adopts a well-clustered binding in which the compound is inserted into the G4 groove, unoccupied by the loop, where the ligand finds favorable van der Waals interactions with G21, G22, and T23 residues (Figure 3.30). This interaction is further reinforced by ionic interactions between the protonated nitrogen of the fused oxazine ring and the backbone phosphate oxygens of A24 residue. Moreover, the pendant benzyl substituent is well positioned to form a parallel displaced π - π interaction with G9 guanine ring. Interestingly, both the (*R*)- and the (*S*)-isomer of **S4-5** were docked but no enantiodiscriminating binding was detected. Rather, the methyl substituent on the benzylic carbon seems to optimally orient the phenyl ring to form the aforementioned stacking interactions. Indeed, when comparing the obtained theoretical model with the experimental results of the NMR titrations, a good consistency is recorded, which gives further confidence in the viability of the obtained docking results.

If a well-defined binding pose was detected when docking **S4-5** into the myc22 solution structure, some difficulties were encountered when docking the ligand into the solution structure of the hybrid form of the human telomere DNA G4 structure. More in detail, three possible conformations were suggested by the docking software in which the ligand always occupies the available groove between the first and fourth strand of the G4 structure (Figure 3.31).

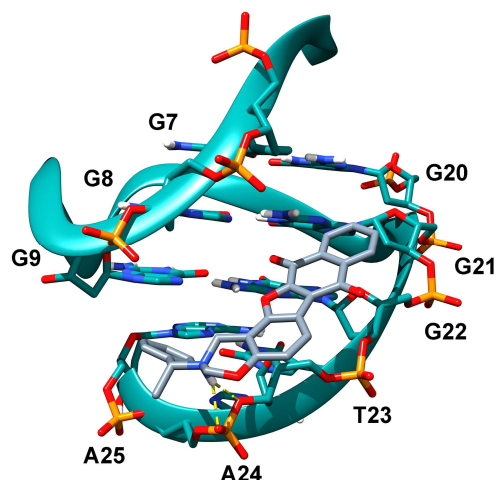


Figure 3.30. Binding mode of S4-5 when docked into the c-myc G4 DNA solution structure. The ligand is represented as grey sticks while the DNA as dark cyan ribbons and sticks. H-bonds are represented as dashed yellow lines. DNA residues were numbered according to Ambrus, A. *et al.*(Ambrus et al., 2005).

In these three conformations, the ligand adopts the same orientation with respect to the G4 but at different levels of the groove, one towards the 3' end (herein referred to as pose A), one in the middle (herein referred to as pose B), and one towards the 5' end (herein referred to as pose C). In pose A, the ligand establishes a H-bond, through its furan oxygen, with the NH₂ group of G24, and an ionic interaction, through its protonated oxazine nitrogen, with the T25 phosphate group. In pose B, a clear H-bond is formed between the ligand furan oxygen and the G23 NH₂ group. In pose C, one of the two ligand quinone oxygens forms a H-bond with the NH₂ group of the G4 and a charged reinforced H-bond through its protonated oxazine nitrogen and the sugar O4' atom of G22. Moreover, the phenyl ring of the ligand benzyl moiety forms parallel displaced π - π interactions with A3 and G22 aromatic rings. Regardless of their positions in all the three conformations (A, B, and C) the ligand forms van der Waals contacts with the groove atoms of the G4 structure. Moreover, as observed for the myc22 structure, also for Tel26 no enantiodiscriminating binding was detected. From the above-described docking results, we cannot give a preference for one of the obtained solutions; rather, it is possible to speculate that the ligand is able to slide into the G4 groove adopting almost isoenergetic binding conformations. This might corroborate the results of the NMR titration experiments, which were not helpful in indicating a specific binding region.

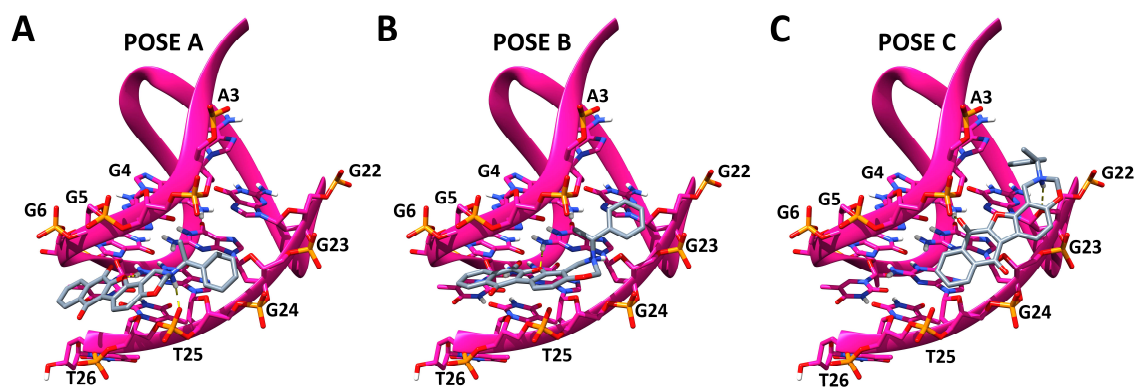


Figure 3.31. Alternative binding poses predicted for S4-5 when docked into the tel26 solution structure. The ligand is represented as grey sticks, the DNA as magenta ribbons and sticks. H-bonds are represented by dashed yellow lines.

3.3.3 Conclusions

In the last decade, increasing efforts in the G4 research field have been devoted to the search of specific G4-binders, able to recognize G4s with high affinity, fully discriminating G4 vs. duplex DNA. Considering the wide distribution and variety of G4 structures within the human genome and their recognized roles in several cancer-related processes, these G4 binders could result into effective multi-target ligands, and expectedly into useful anticancer drugs with almost null side effects. To reach this ambitious goal, an integrated approach is required, involving the massive production of large libraries of new putative G4 ligands, coupled with High Throughput Screening methodologies for the fast and reliable analysis of the selected compounds. In this perspective – which is basically inspired to combinatorial approaches for drug discovery – detailed biophysical characterization studies are carried out only after the biological assays, thus devoted only to the best performing G4-ligands, which, following iterative optimization steps, may finally result into promising candidate drugs.

In this work, we have analyzed a focused library of commercially available analogs of a well-characterized G4-binder, here used as a hit compound, differing for the nature of the pendant groups attached on the N-atom of the oxazine ring. The main idea was to explore the effects of a single diversity element inserted on a common planar scaffold, which *per se* acts as a suitable structural motif for G4 recognition. From a methodological point of view, we also aimed at providing a useful working model, here developed as a proof-of-concept, in which our

affinity chromatography-based G4-CPG binding assays coupled with biological screenings allowed selecting the most promising component of a focused library of putative G4 binders.

In detail, the selected structural analogs of **4c** have been screened by our G4-CPG binding assay to identify the most interesting ligands in terms of affinity towards G4 structures of different topology, as well as ability to discriminate G4 *vs.* duplex DNA. Biological screenings on human cancer cells provided a quite varied picture for these compounds, notwithstanding their high structural similarity, with evidence of marked DNA damage and antiproliferative activity for a restricted set of them. Upon analysis of all the available data, one compound, *i.e.* **S4-5**, has been advanced to successive studies for its highly increased selectivity and reduced toxicity compared with the parent compound **4c**. The combined use of several biophysical techniques – CD, NMR and MST – allowed evaluating qualitative and quantitative aspects of the binding of **S4-5** towards the target G4 structures, indicating a preference for parallel G4s. Notably, NMR data integrated with molecular modeling studies gave a molecular insight into the interaction of the new identified ligand *vs.* G4s of different conformations, essentially suggesting for **S4-5** a G4 groove binder ability. Taken together, our results highlighted also the major role played by the pendant groups of selected G4 ligands in fine-tuning their binding preferences and modes to peculiar G4 structures.

Considering that targeting the grooves of G4s could be a successful, even though still poorly explored, approach for the specific recognition of these structures, and that very few G4-groove binders have been thus far investigated, these results are of great relevance to develop novel effective candidate anticancer drugs. Studies to extend this integrated approach to other classes of putative G4-binders are currently in progress in our laboratories.

3.3.4 Experimental section

Chemistry

All commercial reagents and solvents were purchased from Sigma-Aldrich or Link Technologies, unless otherwise stated. The selected compounds were purchased from Life Chemicals Europe (Germany) and used without further purification. Compound supplier codes: F1094-0196 (**4c**), F1217-0039 (**S4-1**), F1094-0208 (**S4-2**), F1094-0205 (**S4-3**), F1094-0204 (**S4-4**), F1094-0201 (**S4-5**), F1094-0200 (**S4-6**), F1094-0209 (**S4-7**), F1061-0006 (**S4-8**), F1094-0199 (**S4-9**), F1094-0203 (**S4-10**), F1094-0190 (**D4**). All the compounds were analyzed by ¹H NMR for purity control.

G4-CPG assay

Long Chain AlkylAmine-CPG (LCAA-CPG) 1000 Å (purchased from Link Technologies, Bellshill, UK) was functionalized with 3'-O-acetyl-5'-O-(4,4'-dimethoxytrityl)thymidine through a hexaethylene glycol spacer as previously described (Figure 3.20) (Pagano et al., 2015; Platella and Al.). 4.4 mg of this nucleoside-functionalized support (100 nmol) were treated with 3% trichloroacetic acid (TCA) in CH₂Cl₂ (5 min, r.t.) to remove the 4,4'-dimethoxytrityl (DMT) protecting group. After exhaustive washings with CH₂Cl₂, CH₃OH, and again CH₂Cl₂, the support was taken to dryness and then used as a control (nude CPG, Figure 3.20). Standard phosphoramidite chemistry on an automated Applied Biosystem 394 DNA/RNA synthesizer was used to obtain the oligonucleotide-functionalized CPG supports. In particular, using a 1 μmol-scale, “DMT-ON” protocol, the following oligonucleotides were assembled on the CPG supports:

d[(TTAGGG)₄TT] (Tel26),
d(TGGGGAGGGTGGGGAGGGTGGGGAAGGTGGGGA) (c-myc),
d(AGGGAGGGCGCTGGGAGGAGGG) (c-kit1), d(CGGGCGGGCGCGAGGGAGGGG) (c-kit2),
d(AAGGGGAGGGGCTGGGAGGGCCCGGA) (hTERT1), and
d(CGCGAATTCGCGTTTCGCGAATTCGCG) (ds27).

The coupling efficiency of each monomer was monitored by spectrophotometric measurements of the DMT cation, released from the support by acidic treatment with 3% TCA in CH₂Cl₂ before the subsequent coupling step. Considering the number of couplings and the average yield per cycle of 99.8%, 99.7%, 99.9%, 99.7%, 99.9% and 99.2%, respectively for Tel26, c-myc, c-kit1, c-kit2, hTERT1 and ds27, the overall yield was determined to be 95%, 91%, 98%, 94%, 98%, and 80%. Stock solutions of each tested compound were prepared by dissolving a weighed amount of the solid compound in pure DMSO. A known volume of the stock solution was then withdrawn and diluted with the proper buffer, so to obtain a 60 μM ligand solution. All the binding assays were carried out in triplicate, always using freshly prepared 60 μM compound solutions. The binding assay protocol consists in leaving a weighed amount of the functionalized CPG supports (corresponding to 100 nmol of oligonucleotide) in contact with 300 μL of the 60 μM ligand solution in a polypropylene column equipped with a polytetrafluoroethylene frit, a stopcock and a cap (Musumeci et al., 2017a; Platella et al., 2018). After incubation on a vibrating shaker for 4 min, defined volumes of the washing solution (50 mM KCl/10% DMSO/10% CH₃CH₂OH in H₂O) or the releasing solution (2.5 M CaCl₂/15% DMSO in H₂O or pure DMSO) were flown

through the functionalized CPG supports and all the eluted fractions were separately analyzed by UV measurements (Platella et al., 2018). The amount of bound ligand was calculated: i) by subtracting the amount of ligand eluted upon treatment with the washing solution, derived by direct UV measurements, from the ligand amount initially loaded on the supports, or ii) by direct UV measurements of the bound ligand released after treating the supports with the releasing solution (data not shown). The results obtained by the two methods were always in good agreement. After the treatment with the releasing solution inducing G4s and hairpin duplex denaturation, the supports were resuspended in the washing solution and then subjected again to the annealing procedure by taking them at 75 °C for 5 min and then slowly cooling at room temperature. The UV measurements were performed on a JASCO V-550 UV-vis spectrophotometer equipped with a Peltier Thermostat JASCO ETC-505T. The UV quantification of the ligands was determined by measuring the absorbance relative to the λ_{\max} characteristic of each ligand and referring it to the corresponding calibration curves. A quartz cuvette with a path length of 1 cm was used.

Biological experiments

Cells and culture condition. Human fibroblasts (BJ) and human cervical cancer cells (HeLa) were obtained as previously reported. (Salvati et al., 2010) BJ-hTERT cells were obtained infecting primary BJ cells with a retrovirus carrying hTERT (Addgene plasmid #1773); BJ-EHLT derived from the transformation of BJ fibroblasts with hTERT and SV40 early region (BJ-EHLT). BJ-hTERT, BJ-EHLT and HeLa were grown in Dulbecco Modified Eagle Medium (D-MEM, Invitrogen Carlsbad, CA, USA) supplemented with 10% Fetal Bovine Serum (FBS), 2 mM L-glutamin and antibiotics at 37°C in a 5% CO₂-95% air atmosphere.

Immunofluorescence. Cells were fixed in 2% formaldehyde and permeabilized in 0.25% Triton X-100 in phosphate buffered saline (PBS) for 5 min at r.t. For immune-labeling, cells were incubated with primary antibody for 2 h at r.t., washed twice in PBS and finally incubated with the secondary antibodies for 1 h. The following primary antibodies were used: Mouse mAb anti- γ H2AX (Millipore, Billerica, MA, USA) and Rabbit pAb anti-TRF1 N19 (Santa Cruz Biotechnologies, Santa Cruz, CA, USA). The following secondary antibodies were used: Anti-Mouse IgG (H+L), F(ab')₂ Fragment (Alexa Fluor 488 Conjugate) (Cell Signaling) and Anti-rabbit IgG (H+L), F(ab')₂ Fragment (Alexa Fluor 555 Conjugate) (Cell Signaling). Nuclei were stained with 4',6-diamidino-2-phenylindole (DAPI, Sigma). Fluorescence signals were

recorded by using a Leica DMIRE2 microscope equipped with a Leica DFC 350FX camera and elaborated by Leica FW4000 deconvolution software (Leica, Solms, Germany). For quantitative analysis of γ H2AX positivity, 300 cells on triplicate slices were scored. For TIF analysis, a single plane was analyzed, and 30 γ H2AX-positive cells were scored. Cells with at least four co-localizations (γ H2AX/TRF1) were considered as TIF-positive.

Clonogenic assay. Human cervical cancer cells, HeLa, were seeded in 60 mm-Petri dishes at the clonogenic density of 500 cells/plate in DMEM medium with 10% FBS. After 24 h, cells were treated with compound **4** or its analogues at 2 μ M concentration. After 10 days, the cells were stained with 2% methylene blue in 50% ethanol and the number of colonies was counted. Surviving fractions were calculated as the ratio of absolute survival of the treated sample/absolute survival of the untreated sample.

Cell viability. BJ-hTERT fibroblasts were seeded in 60-mm Petri dishes at a density of 5×10^4 cells/plate. After 24 h from plating, cells were treated with compound **4c** or **S4-5** at the doses of 1 and 2 μ M. Cell viability (Trypan Blue dye exclusion) was determined after 6 days from treatment.

CD experiments

CD spectra were recorded in a quartz cuvette with a path length of 1 cm on a Jasco J-715 spectropolarimeter equipped with a Peltier-type temperature control system (model PTC-348WI). The spectra were registered at 20 °C in the range 240-800 nm with 2 s response, 200 nm/min scanning speed, 2.0 nm bandwidth, and corrected by subtraction of the background scan with buffer. All the spectra were averaged over 3 scans. The oligonucleotides d[TTAGGG]₄TT] (Tel26), d(TGGGGAGGGTGGGGAGGGTGGGGAAGGTGGGGA) (c-myc) and d(CGCGAATTCGCGTTTCGCGAATTCGCG) (ds27) were synthesized by standard automated solid phase oligonucleotide synthesis on an Applied Biosystem 394 DNA/RNA synthesizer. After ammonia treatment (55 °C, 12 h) allowing both deprotection and detachment from the solid support, the crude oligonucleotides were purified by HPLC on a SAX analytical column and then dialyzed against water using a Float-A-Lyzer G2 dialysis device (MWCO 0.5-1.0 kDa, three H₂O changes over 24 h). After lyophilization, the oligonucleotides were dissolved in a 20 mM KCl - 5 mM KH₂PO₄ - 10% DMSO buffer (pH 7) to obtain 2 μ M solutions, then annealed by heating to 95 °C for 5 min, followed by slow cooling to room temperature. The ligand stock solutions were 4 mM in DMSO. CD titrations were

obtained by adding increasing amounts of the ligands (max 5 mol equivalents, corresponding to a 10 μ M solution in ligand) to Tel26, cmyc and ds27. After each ligand addition, the system was allowed equilibrating before registering the spectra. For the CD-melting experiments, the ellipticity was recorded at 290, 263 and 251 nm for Tel26, c-myc and ds27, respectively, with a temperature scan rate of 0.5 $^{\circ}$ C/min in the range 20-90 $^{\circ}$ C.

NMR experiments

NMR experiments were performed on a 700 MHz Varian Unity INOVA spectrometer. One dimensional 1 H-NMR spectra of the samples in H₂O were recorded at 25 $^{\circ}$ C using pulsed-field gradient DPGSE for H₂O suppression.(Hwang and Shaka, 1995) Data were processed on iMAC running iNMR software (www.inmr.net). DNA samples were prepared at 0.2-0.4 mM strand concentration in 0.6 mL of H₂O/D₂O (9:1) buffer solution. The following oligonucleotides were used for the NMR experiments: the truncation of human telomeric sequence d[TTGGG(TTAGGG)₃A] (Tel24), (Luu et al., 2006a) the d(TGAGGGTGGGTAGGGTGGGTAA) sequence from the NHE III element of the MYC oncogene (myc22), containing two G to T substitutions (G14/T and G23/T, numbering according to Ambrus, A. *et al.*),(Ambrus et al., 2005) and the self-complementary duplex-forming dodecamer d(CGCGAATTCGCG) (ds12) (Jaroszewski et al., 1996). In order to avoid different G4 folds or conformational heterogeneity in solution, DNA samples were prepared using the appropriate experimental conditions as used for the determination of their 3D structures. Thus, the following buffers containing 10% D₂O were used: 25 mM KH₂PO₄, 70 mM KCl, 0.2 mM EDTA (pH 7.0) for myc22 and Tel24; 20 mM NaH₂PO₄, 200 mM NaCl (pH 7.0) for ds12. The samples were heated at 90 $^{\circ}$ C for 5 min and then slowly cooled to room temperature overnight to achieve the correct folding of the oligonucleotides. Aliquots of **4c** and **S4-5** stock solutions in DMSO-d₆ were directly added to the DNA solution inside the NMR tube; the final DMSO concentration was 14%. A control titration was also performed by adding DMSO-d₆ alone to the DNA solution and subtracted from each titration experiment to remove effects due to DMSO on the chemical shift changes.

Microscale thermophoresis experiments

MST measurements were performed using the Monolith NT.115 (Nanotemper Technologies, Munich, Germany). The Cy5-fluorescently labeled Tel26 and myc22 oligonucleotides (Biomers) were prepared in 5 mM potassium phosphate buffer (pH 7.0)

containing 20 mM KCl supplemented with 0.1% Tween. The concentration of the labeled oligonucleotide was kept constant at 80 nM, while a serial dilution of the ligand (1:2 from 4 mM ligand stock solution in 100% DMSO) in the same buffer used for DNAs was prepared and mixed with the oligonucleotide solution with a volume ratio of 1:1. All the samples, containing 8% DMSO as the final concentration, were loaded into standard capillaries (NanoTemper Technologies). Measurements were performed at 20 °C, using auto-tune LED power and medium MST power. MST data analysis was performed by employing the MO.Affinity Analysis software (v2.3) provided with the instrument. Plots were rendered with GUSSE version 1.2.1 software (<http://biophysics.swmed.edu/MBR/software.html>).

Docking

The latest version of the docking software AD4 (version 4.2), (Morris and Huey, 2009) along with its graphical user interface AutoDockTools (ADT), was employed. The NMR structures used for the docking studies had the following PDB codes: 2JPZ, and 1XAV. (Ambrus et al., 2005; Dai et al., 2007) The DNA G4s were prepared for the docking using the Maestro suite (Schrödinger Release 2017-2: Maestro, Schrödinger, LLC, New York, NY, 2017), which assigns bond orders, add hydrogen atoms, deletes water molecules and generates the appropriate protonation states. The 2D Sketcher tool of Maestro was used to build **S4-5**. Of this ligand, the protonation, tautomeric, and isomeric states were calculated through LigPrep, part of the same suite. The ligand and the G4 DNA structures were converted to the AD4 specific file format (PDBQT) using the python scripts `prepare_ligand4.py` and `prepare_receptor4.py`, part of ADT, applying the standard settings. The docking area was centered on the DNA center of mass. For each G4 structure, a set of grids of $60 \text{ \AA} \times 60 \text{ \AA} \times 60 \text{ \AA}$ with 0.375 \AA spacing was calculated around the docking area for the ligand atom types using AutoGrid4. For each G4, 100 separate docking calculations were performed. Each docking calculation consisted of 10 million energy evaluations using the Lamarckian genetic algorithm local search (GALS) method. The GALS method evaluates a population of possible docking solutions and propagates the most successful individuals from each generation into the subsequent generation of possible solutions. A low-frequency local search according to the method of Solis and Wets is applied to docking trials to ensure that the final solution represents a local minimum. All dockings were performed with a population size of 250, and 300 rounds of Solis and Wets local search were applied with a probability of 0.06. A mutation rate of 0.02 and a crossover rate of

0.8 were used to generate new docking trials for subsequent generations, and the best individual from each generation was propagated over the next generation. The docking results from each of the 100 calculations were clustered on the basis of root-mean square deviation (rmsd) (solutions differing by less than 2.0 Å) between the Cartesian coordinates of the atoms and were ranked on the basis of the free energy of binding (ΔG_{AD4}). Molecular modeling figures were rendered using the UCSF Chimera software.(Pettersen et al., 2004)

Chapter 4

TARGETING G-TRIPLEX

4.1 Discovery of the first G-triplex/G-quadruplex stabilizing compound (*Paper IV*)

The discovery of molecules able to bind and stabilize G3 structures is highly appealing, but their transient state makes challenging their recognition. In order to identify putative G3 binders, a virtual screening approach was employed. The G3 stabilizing properties of the identified compounds were analyzed by CD melting assay. DSC, non-denaturing gel electrophoresis, NMR and molecular modeling studies were performed to investigate the interaction between a selected compound and G-rich DNA structures. Cytotoxic activity of this compound was also evaluated by MTT cell proliferation assay (Amato et al., 2017).

In order to avoid confusion, compounds **1-15** in Paper IV here appear as **1d-15d**.

4.1.1 Introduction

Several spectroscopic studies show that some G-rich sequences fold into intramolecular G4 structures by following a multistep pathway that proceeds through a quite stable G3 intermediate (Cerofolini et al., 2014; Limongelli et al., 2013; Rajendran et al., 2014; Wu et al., 2015; Xi-Miao et al., 2015). In addition, recent studies have revealed that G3s play a role also in the unfolding process of some G4 structures (Xi-Miao et al., 2015). The G3 biological role is still unclear, so the discovery of G3 binders could be a way to stabilize these structures in cell, also to investigate their role.

In this study, a structure-based virtual screening (VS) approach together with a combination of experimental techniques were used to identify new molecular scaffolds able to target a very simple G3 structure, namely the 11-mer-TBA (PDB ID: 2MKM) (Cerofolini et al., 2014). Particularly, 15 small molecules were selected by high-throughput in silico screening of a large number of compounds from a commercially available database. These small organic molecules (Table 4.1) are characterized by hydrogen-bond acceptor/donor groups, (hetero)aromatic and/or

non-aromatic heterocyclic rings, as best hits for the G3 formed by the 11-mer-TBA DNA sequence.

Table 4.1. Vendor codes for the investigated compounds and correspondence in the text (in order to avoid confusion, compounds 1-15 in Paper IV here appear as 1d-15d).

Compound ID	Compound name
MCULE-1440087997	1d
MCULE-2379405833	2d
MCULE-3336930203	3d
MCULE-3364906924	4d
MCULE-1305746880	5d
MCULE-4489228167	6d
MCULE-5140576857	7d
MCULE-5152866419	8d
MCULE-5399977447	9d
MCULE-5519408107	10d
MCULE-5841321496	11d
MCULE-6515894787	12d
MCULE-7496218237	13d
MCULE-9515923402	14d
MCULE-4224006137	15d

The selected hits were evaluated for their binding properties by biophysical methodologies, and one of them, a dihydropyrimidin-4-one derivative (6-(4-fluorophenyl)-2-[(8-methoxy-4-methylquinazolin-2-yl)amino]-3,4-dihydropyrimidin-4-one), herein referred to as compound **1d** (Figure 4.1), turned out to be a promising lead compound able to stabilize the G3 structure. Interestingly, the physicochemical characterization of the binding profile of this ligand, led to the discovery that **1d** is also able to interact and stabilize some G4 structures, thus acting as a dual G3 and G4 stabilizer, with no affinity for the duplex DNA. The subsequent investigation of the biological activity of the selected compound demonstrated that it is endowed with cytotoxic effect on human U2OS cells.

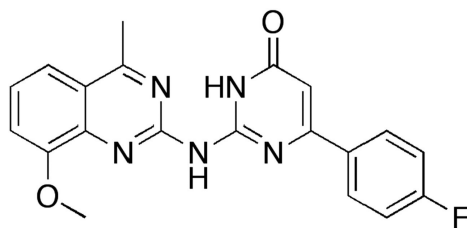


Figure 4.1. Chemical structure of compound 1d.

4.1.2 Results and discussion

Compound selection

There are several studies using VS calculations for the identification of G4-targeting ligands (Cosconati et al., 2009, 2012; Di Leva et al., 2013), however none for the targeting of the G3 DNA structure. To this end, the intramolecular G3 structure formed by the d(GGTTGGTGTGG) sequence (11-mer-TBA) was used as a target model (Figure 4.2). This structure consists of two G-triads linked with two parallel G-strands (G1–G2 and G10–G11) interleaved with an antiparallel G-strand (G5–G6), which are connected together by two side loops (T3–T4, T7–G8–T9) (Cerofolini et al., 2014; Limongelli et al., 2013). Overall, the G3 formed by the 11-mer-TBA sequence maintains a structure that recalls the chair-like, antiparallel folding of TBA also in the absence of the fourth missing strand (Schultze et al., 1994).

Starting from this structure, in collaboration with Prof. Sandro Cosconati (DiSTABiF, University of Campania “Luigi Vanvitelli”, Caserta, Italy), the docking software AD4.2 was used to dock a diversity set of a commercially available molecular database (Morris et al., 2009). This VS process resulted in 15 compounds that were selected and purchased after the visual inspection of the theoretical results.

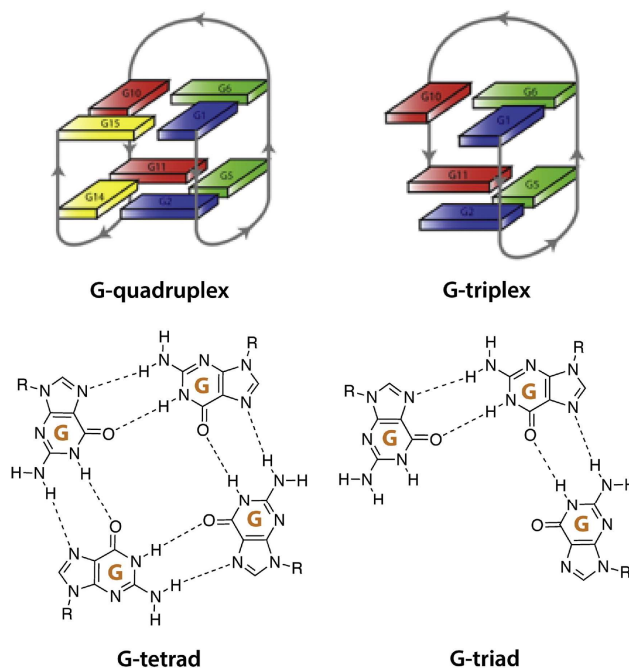


Figure 4.2. Comparison between G4 and G3 structures formed by the 15-mer and 11-mer-TBA sequences.

Circular dichroism studies

In order to identify true hits, the 15 computationally selected small molecules were first investigated by CD spectra and CD-melting analysis (Karsisiotis et al., 2011; Masiero et al., 2010b; Randazzo et al., 2013). CD experiments were performed to examine the potential of compounds **1d–15d** to alter the native folding topology of the G3. The CD spectrum of the structure adopted by the 11-mer-TBA sequence alone showed two positive bands at 289 nm and 253 nm, and a small negative one below 240 nm. These bands are characteristic of a homopolar stacking of the nucleobases, as expected for the antiparallel-stranded G3 folding adopted by the 11-mer-TBA sequence (Cerofolini et al., 2014; Limongelli et al., 2013).

Then, DNA/ligand mixtures were obtained by adding ligands (4 mol equivalents) to the folded 11-mer-TBA G3 structure. No significant variations of CD signal were observed upon addition of any compound (**1d–15d**), clearly suggesting an overall preservation of the G3 architecture in the presence of each ligand. The stabilizing properties of the compounds were evaluated by CD melting experiments measuring the ligand-induced change in the melting temperature (ΔT_m) of G3 structure. CD melting experiments in the absence and presence of each compound (4:1 [compound]/[DNA] ratio) were recorded by following CD changes at the wavelength of maximum of intensity (289 nm) (Table 4.2). These experiments showed that 14 out of 15 compounds do not have the ability to significantly increase the T_m of the 11-mer-TBA structure ($\Delta T_m < 2$ °C), thus proving that they are non-binders or weak binders of G3. On the other hand, these CD experiments allowed for the identification of a compound (**1d**) that significantly increased the thermal stability of the 11-mer-TBA folded form and that could be considered as a promising G3 binder.

Indeed, this compound turned out to have a good stabilizing effect with a ΔT_m of 8.5 °C at the investigated ligand/ DNA ratio (Table 4.2). Therefore, we performed additional CD melting experiments for compound **1d** by using a wider range of ligand concentrations to evaluate the effect of increasing concentrations of ligand on the stability of the G3 structure. Four different [ligand]/ [DNA] ratios were examined (1:1, 2:1, 4:1, 8:1) (Figure 4.3). Under the experimental conditions used, biphasic behavior was observed at low [ligand]/[DNA] ratios ($\leq 2:1$), which prevented the obtaining of accurate melting temperature data from those curves. The first melting transition most likely results from the fraction of the G3 that does not bind **1d** and the second one from the fraction of DNA binding the compound. Overall, a dose-dependent increase of G3 thermal stability was observed up to 4:1 [ligand]/[DNA] ratio. Unfortunately, at

8:1 [ligand]/[DNA] ratio, a decrease of thermal stability was observed. This was probably an artifact generated by the progressive formation of a suspension inside the cuvette during the melting experiment.

Table 4.2 Compound-induced thermal stabilization of the 11-mer-TBA G3 measured by CD melting experiments.

Compound name	ΔT_m (°C) ^a
1d	+8.5
2d	-1.0
3d	0.0
4d	0.0
5d	-1.0
6d	-0.5
7d	-0.5
8d	-0.5
9d	-1.0
10d	0.0
11d	-0.5
12d	-0.5
13d	-0.5
14d	+1.0
15d	0.0

^a ΔT_m represents the difference in melting temperature [$\Delta T_m = T_m$ (DNA + 4 ligand equivalents) - T_m (DNA)]. All experiments were duplicated, and the values reported are average of two measurements.

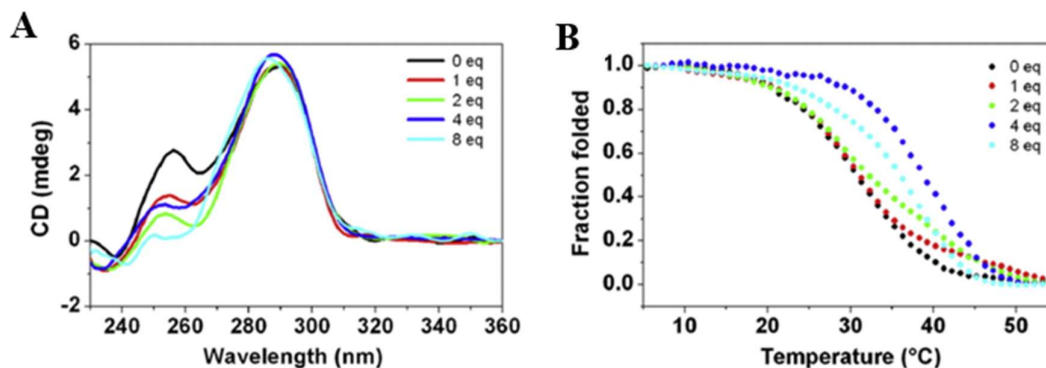


Figure 4.3. (A) CD spectra and (B) CD melting experiments of 11-mer-TBA G3 in the absence and presence of increasing amounts of compound **1d**.

Since G3 is an intermediate in the folding process of G4-forming sequences, we decided to verify the ability of **1d** to bind and stabilize different G4 structures in solution. To this purpose, several diverse G4-forming sequences able to form parallel, antiparallel and hybrid G4 structures were selected for this study. Particularly, two parallel G4-forming sequences from the *KIT* (*c-kit1*) and the *MYC* (*myc22*) promoter regions were used, in addition to TBA and telomeric Tel23 sequences, which are able to fold into antiparallel and hybrid G4 structures, respectively. The structures adopted by each G4 sample were verified by CD spectroscopy. *c-kit1* (Figure 4.4A) and *myc22* (Figure 4.4B) displayed the characteristic CD signature of parallel-stranded G4 topologies (positive and negative bands at 264 nm and around 240 nm, respectively) (Karsisiotis et al., 2011; Masiero et al., 2010b; Randazzo et al., 2013). On the other hand, TBA showed a positive band at 295 nm and a negative band at around 240 nm (Figure 4.4C), characteristics of an antiparallel G4 conformation, while Tel23 showed a positive band at 289 nm with a shoulder at ca. 268 nm, and a weak negative band at around 240 nm, in agreement with the formation of a hybrid G4 structure (Figure 4.4D) (Karsisiotis et al., 2011; Masiero et al., 2010b; Randazzo et al., 2013). These bands were not significantly modified upon addition of **1d** (4 mol equivalents) to any of the structures analyzed, clearly suggesting an overall preservation of their G4 architectures in the presence of ligand. Hence, the stabilizing properties of **1d** on those structures were evaluated by recording CD melting experiments of DNAs in the absence and presence of ligand (4:1 ligand/DNA ratio), following the variations of the intensities of CD signals at the wavelength of 264, 295 and 289 nm for parallel, antiparallel and hybrid G4s, respectively (Figure 4.4A-D).

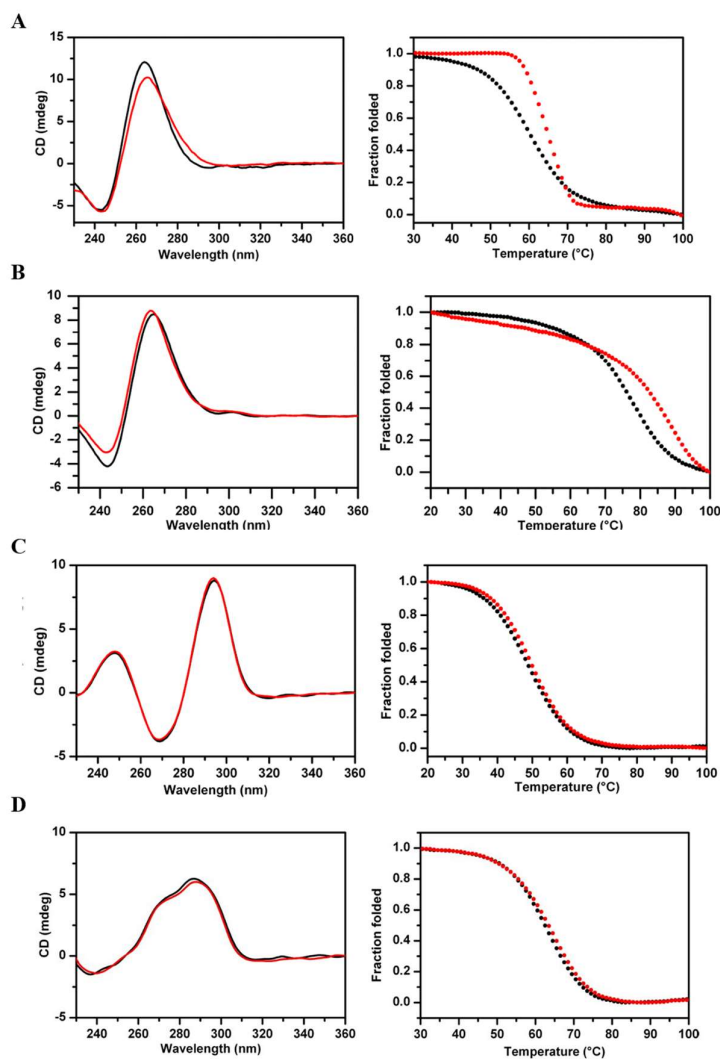


Figure 4.4. CD spectra (left) and CD melting (right) experiments in the absence (black) and presence (red) of compound **1d** (4:1 compound/DNA ratio) of (A) *KIT* G4 (c-kit1), (B) *MYC* G4 (myc22), (C) 15-mer-TBA G4, and (D) human telomeric Tel23 G4.

Results of these experiments, reported in Table 4.3, reveal interesting differences in the binding behavior of **1d** to the different G4 topologies. In fact, whereas **1d** demonstrated no increase in melting temperature (T_m) for the antiparallel and hybrid G4 structures, it exhibited good G4-stabilizing properties toward the parallel G4 structures adopted by myc22 and c-kit1 promoter sequences. Actually, the addition of this ligand exerted a more prominent thermal stabilization of myc22 rather than c-kit1 displaying $\Delta T_m > 15$ °C for myc22 and of 4.5 °C for c-kit1 (Table 4.3).

Table 4.3. Ligand **1d-induced thermal stabilization of G4 and duplex DNAs measured by CD melting experiments.**

Compound	ΔT_m (°C) ^a				
	c-kit	myc22	Tel23	TBA	ds12
1d	4.5	> 15.0	0.0	0.0	0.0

^a ΔT_m represents the difference in melting temperature [$\Delta T_m = T_m(\text{DNA}+4 \text{ ligand equivalents}) - T_m(\text{DNA})$].

We hypothesize that the more exposed terminal G-tetrad surface of the myc22 G4 might be more conducive of favorable stacking interactions rather than in c-kit1 (Amato et al., 2016; Dai et al., 2011; Wei et al., 2015). Overall, these results suggest a selective preference of **1d** for parallel-type G4 folds over antiparallel ones and reveal that the parallel fold of the myc22 G4 is the preferred target of **1d** among the investigated G4s.

Finally, the effect of increasing concentrations of **1d** on the stability of myc22 G4 structure was also investigated by carrying out CD melting experiments (Figure 4.5). An increase of the T_m of this G4 was observed with every addition of ligand up to the 4:1 ratio, indicating that **1d** could bind and stabilize the G4 in a specific manner (the addition of more than four equivalents of ligand does not lead to a further enhancement of G4 stability).

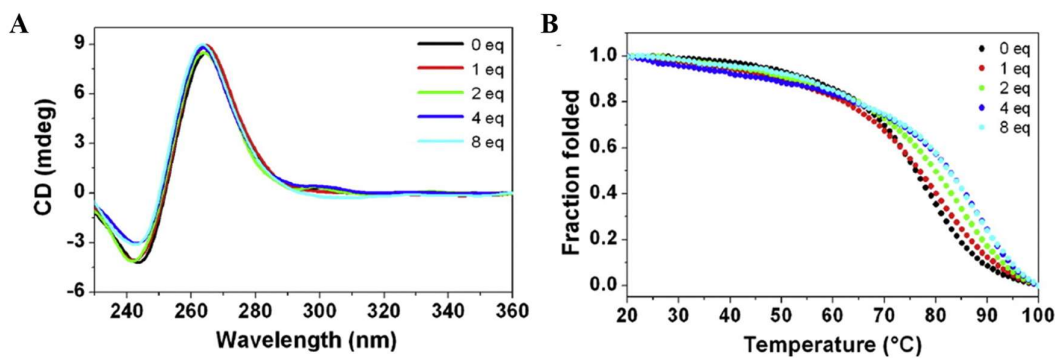


Figure 4.5. (A) CD spectra and (B) CD melting experiments of myc22 in the absence and presence of increasing amounts of compound **1d.**

The selectivity of **1d** toward G3 and G4 structures over duplex DNA, was determined by comparing the above results with those obtained using a duplex DNA (ds12) (Dickerson and Drew, 1981; Drew and Dickerson, 1981), which is formed by a self-complementary 12-nt sequence. In the presence of K^+ , the CD spectrum of ds12 was characterized by a positive band at around 280 nm and a negative one at 250 nm, typical of values observed for duplex DNA

(Figure 4.6). These bands were not modified upon ligand addition. Furthermore, CD melting studies show that **1d** was unable to induce any change in the melting temperature of the duplex structure, clearly indicating no affinity of this compound for the duplex DNA (Figure 4.6).

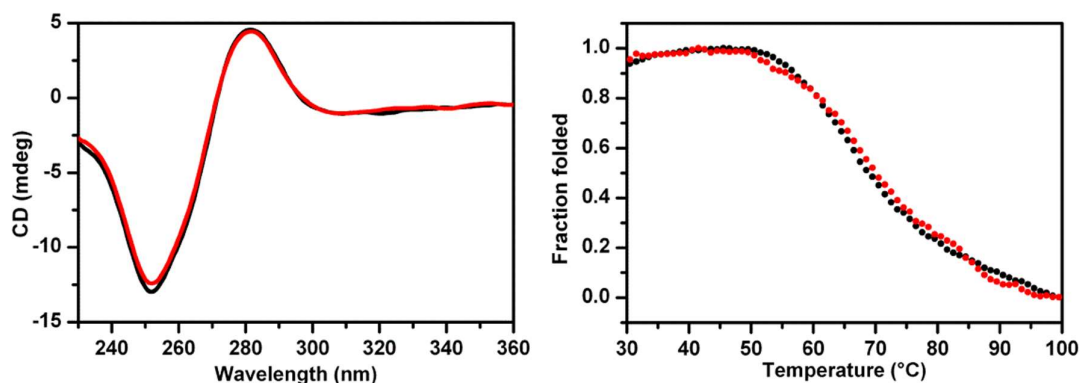


Figure 4.6. CD spectra (left) and CD melting (right) experiments of ds12 in the absence (black) and presence (red) of compound **1d** (4:1 compound/DNA ratio).

Non-denaturing gel electrophoresis

To better clarify what happens to the 11-mer-TBA G3 upon addition of large amounts of compound **1d**, ligand-DNA complexes were also studied by non-denaturing gel electrophoresis (PAGE). As shown in Figure 4.7, the 11-mer-TBA G3 moves essentially as a single band in the gel, thus suggesting the presence of the unimolecular G3 structure as unique conformation in solution and the absence of high-order DNA structures (Petraccone et al., 2012). Compound **1d** addition to the G3 did not affect the G3 mobility until 4:1 [ligand]/[DNA] ratio, unless the appearance of some faint lower mobility bands. Conversely, in the presence of a large excess of **1d** (i.e. 8:1 and 10:1 [ligand]/[DNA] ratio), a continuous dark background of unresolved polymeric structures was observed in the corresponding lanes. Overall, PAGE results are in perfect agreement with the observation from CD studies, suggesting that large amounts of ligand induce DNA aggregation.

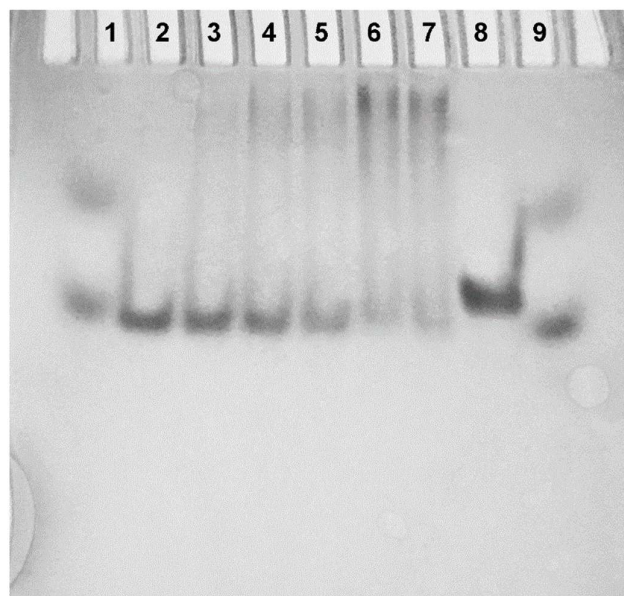


Figure 4.7. Non-denaturing 20% PAGE of 11-mer-TBA G3 in the absence and presence of increasing amounts of compound **1d**. Lanes 1 and 9: mixture of 1:1 Bromophenol blue and Xylene cyanol; lane 2: G3 alone; lane 3: G3 + 1 eq ligand; lane 4: G3 + 2 eq ligand; lane 5: G3 + 4 eq ligand; lane 6: G3 + 8 eq ligand; lane 7: G3 + 10 eq ligand; lane 8: d(TTGGGTTAGGGTTAGGGTTAGGGA) (Tel24).

Differential scanning calorimetry analysis

DSC experiments were carried out to study the thermodynamic stability of the 11-mer-TBA G3 in the absence and presence of **1d**. The obtained DSC melting profiles are shown in Figure 4.8. The 11-mer-TBA G3 in the absence of ligand showed a highly reversible unfolding process, as the heating profiles of subsequent scans are superimposable, thus indicating that the observed unfolding/folding process is an equilibrium process (Limongelli et al., 2013).

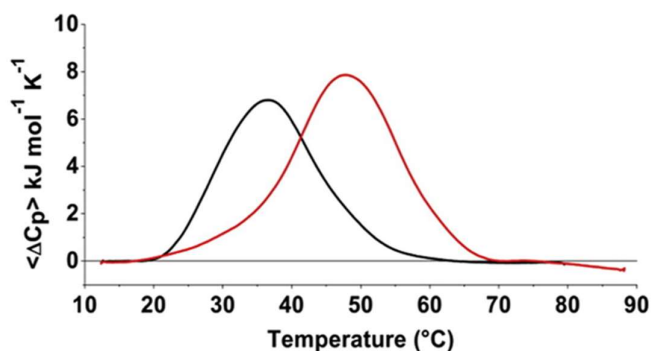


Figure 4.8. DSC profiles for 11-mer-TBA G3 in the absence (black line) and presence (red line) of 4 mol equivalents of compound **1d**.

The DSC curves show a maximum at T_m of 34.5 (± 0.5) °C and a $\Delta H^0(T_m)$ of 130 (± 5) kJmol⁻¹ (very close to the calculated van't Hoff enthalpy) (Table 4.4). These values are in good agreement with what reported in the literature (Limongelli et al., 2013), and with the results obtained by CD reported herein. On the other hand, the 11-mer-TBA G3 in the presence of ligand results in non-reversible DSC profiles. The first DSC curve transition shows a maximum at T_m of 47.5 (± 0.5) °C and a $\Delta H^0(T_m)$ of 170 (± 5) kJ mol⁻¹. The total enthalpy measured in the DSC heating scans, which directly represents the amount of all the interactions existing in the folded DNA structure, suggests the presence of additional stabilizing interactions in the presence of **1d**. Overall, DSC data are indicative of the propensity of **1d** to form stable interactions with G3-forming sequence resulting in the formation of a bound complex whose thermal stability is significantly enhanced compared to the native state.

Table 4.4. Thermodynamic parameters for the unfolding of 11-mer-TBA in the absence and presence of 1d determined by DSC.

	T_m (°C)	$\Delta_{cal}H^0$ (kJ/mol)	$\Delta_{v.H.}H^0$ (kJ/mol)
11-mer-TBA	34.5 \pm 0.5	130 \pm 5	128 \pm 10
11-mer-TBA/ 1d	47.5 \pm 0.5	170 \pm 5	-

Nuclear Magnetic Resonance studies

NMR studies were carried out to obtain structural information about the binding mode of compound **1d**. As stated above, this compound is able to stabilize the 11-mer-TBA G3 structure. In addition, it was also found to stabilize G4 structures in the parallel conformation, thus acting as a dual G3/G4 binder, with no affinity for the duplex DNA.

The spectral regions of the imino and aromatic protons of 11-mer- TBA G3 and myc22 in the absence and presence of increasing amounts of compound **1d** are shown in Figure 4.9. The imino protons of either 11-mer-TBA G3 (Cerofolini et al., 2014; Limongelli et al., 2013) and myc22 (Ambrus et al., 2005; Dai et al., 2011) were previously assigned and these were used to monitor the shift of the signals and thus grasping the mode of interaction. Considering 11-mer-TBA G3, no appreciable shift of any signal was observed until 2:1 [ligand]/[DNA] ratio (Figure 4.9). However, at 2 mol equivalents. of ligand, a severe broadening of all DNA signals occurred, suggesting that i) the binding of **1d** to the G3 is a dynamic process with a fast exchange rate on the NMR time-scale, and ii) probably additional non-specific ligand-DNA interactions occur.

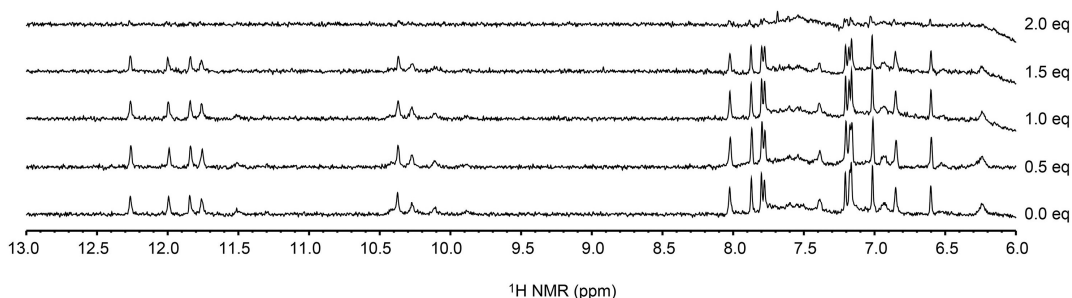


Figure 4.9. Imino and aromatic proton regions of the 11-mer TBA G3 titrated with compound 1d.

Concerning myc22, imino protons became already broad upon addition of just 0.5 mol equivalents of **1d** (Figure 4.10). Major perturbation was observed for the imino protons belonging to the external 5' G-tetrad (G7–G11–G16–G20) and 3' G-tetrad (G9–G13–G18–G22), while the remaining imino protons of the middle tetrad (G8–G12–G17–G21) were less affected. In contrast to G3, aromatic protons of myc22 G4 were poorly affected by the binding. Altogether, these results suggest that end-stacking is the preferred binding mode of the ligand on this molecule.

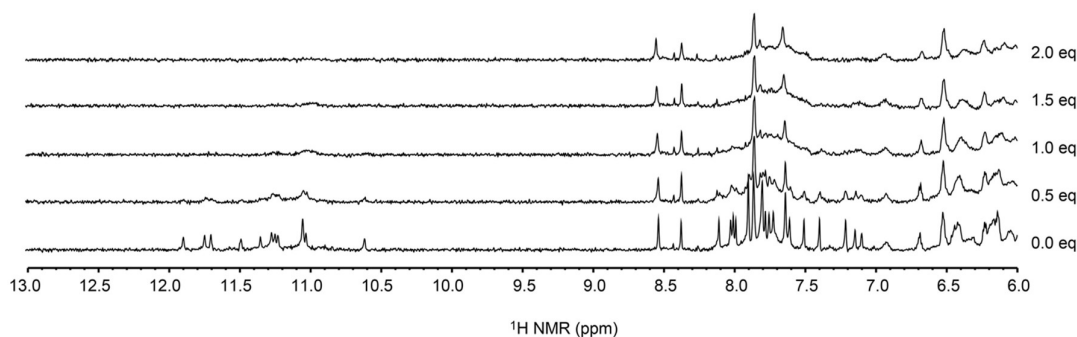


Figure 4.10. Imino and aromatic proton regions of the myc22 G4 titrated with compound 1d.

Molecular Modeling studies

To better characterize the binding of **1d** to the investigated G3 and G4 structures, theoretical studies were attained by performing docking experiments coupled with molecular dynamics simulations. These studies were carried out in collaboration with Prof. Sandro Cosconati (DiSTABiF, University of Campania “Luigi Vanvitelli”, Caserta, Italy). In particular the ligand was docked into the published X-ray structure of *KIT* G4 (PDB ID: 4WO2), into the solution

structures of *MYC* G4 (PDB ID: 1XAV) and into the structure of G3 (PDB ID: 2MKM). These calculations (also attained with the AD4.2 software) suggest that, in the *MYC* G4 structure, **1d** is able to engage stacking interactions at the G4:G8:G13:G17 plane. Moreover, this pose is stabilized by H-bonds with the adjacent A3 residue while the pendant 4-fluorophenyl moiety seems to be solvent exposed (Figure. 4.11A). When docking simulations were attained on the *KIT* G4, two different solutions were suggested in which the ligand is proposed to occupy one of the G4 available grooves. In the first one, (herein referred to as pose 1) the ligand spans the entire groove by maximizing the van der Waals interactions between its hydrophobic side with the DNA and exposing its polar atoms towards the solvent. Conversely, in the second calculated binding pose (herein referred to as pose 2) the ligand occupies the same groove by exposing its polar groups towards the DNA. This allows the ligand to establish a double-bridged H-bond with the NH₂ group of G15 while the pendant hydrophobic p-fluorophenyl group is exposed to the solvent (Figure 4.11B-C). The presence of two possible binding poses was also suggested when **1d** was docked in the G3 solution structure. As happened for the calculations achieved for the *MYC* G4, also the **1d**-G3 recognition seems to be mediated by stacking interactions between the ligand and the DNA G1 residue with the two predicted binding poses differing for the ligand orientation with respect to the DNA structure. In both poses, the ligand seems to be able to establish H-bonds with the G8 residue, (Figure 4.11D-E) while its methoxy substituent is alternatively able to form an additional H-bond with G8 and G10 (pose 1 and 2 in Figure 4.11D-E, respectively). From the data reported so far, a preference for a single binding interaction pattern between the ligand and the target DNAs could not be given.

In this respect, MD simulations allow to further study the viability of the calculated binding poses from a dynamical point of view. Thus, 40 ns explicit solvent MD simulations were performed for the 5 supposed complexes of compound **1d** (one for the *MYC* G4 and two for the *KIT* G4 and G3), starting from the binding poses calculated by the docking software and taking advantage of the Amber 14 program.

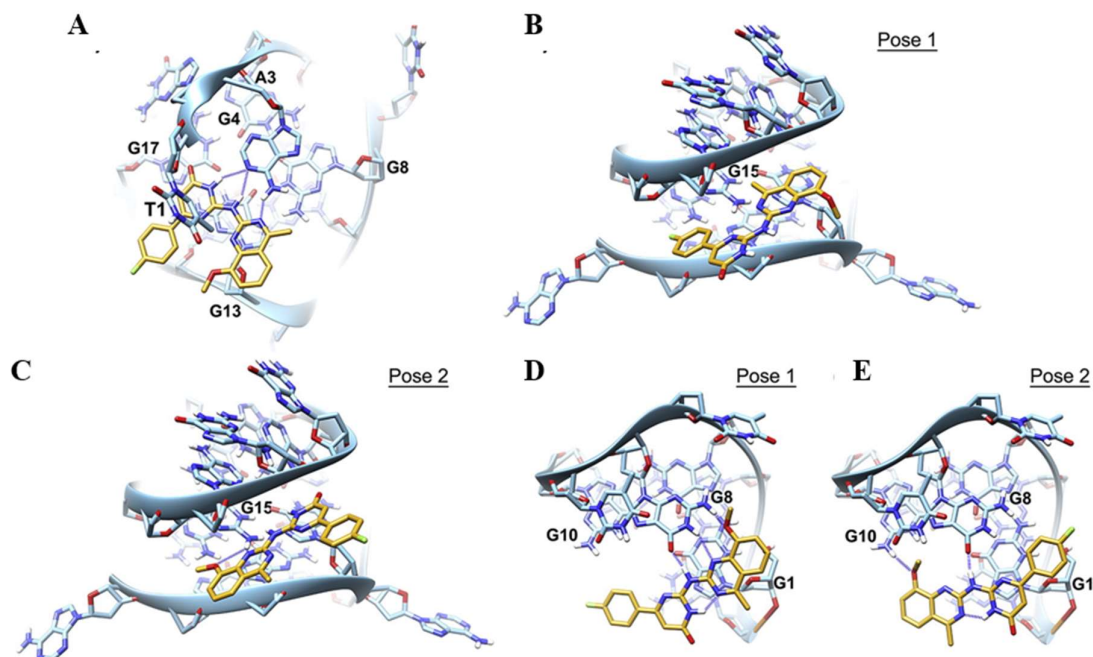


Fig. 4.11. Binding mode of **1d** as calculated by the docking program in the structures of (A) *MYC*, (B and C) *KIT* G4s, and (D and E) G3. DNA is represented as blue sticks and ribbons while the ligand as yellow sticks. H-bond interactions are indicated as dashed blue lines while only interacting residues are labeled for clarity reasons.

In regards to the *1/MYC* complex, plotting of the root mean square deviation (rmsd) of the ligand heavy atoms over the 40 ns production run reveals that the ligand adopts a stable binding conformation after 6 ns which is kept for the remainder of the simulation time (Figure 4.12A). In this binding mode, the ligand is sandwiched between the G4:G8:G13:G17 and T1 while additional H-bonds with the A3 residue are formed (Figure 4.12B). Interestingly, plotting of the rmsd calculated for the DNA backbone atoms over time demonstrates that the starting DNA topology is preserved for the entire simulation time (with an average rmsd value of 2.1 Å).

Similar calculations were very helpful to probe the stability of the two possible binding modes calculated for the ligand in the *KIT* G4. In this complex, MD simulations for pose 1 suggest that the ligand is able to adopt a fairly stable conformation after around 9 ns in which, while still occupying the G4 groove, engages an additional H-bond interaction with the G15 NH₂ and a T-shaped charged transfer interaction with A1 residue. Analysis of the rmsd calculated considering the DNA backbone atoms allowed to infer that also in this case the DNA G4 topology is preserved during the entire simulation with an average value of 1.7 Å. On the

other hand, plotting of the rmsd values over time of the ligand heavy atoms in pose 2 of the *KIT* G4 complex demonstrated that the ligand is unable to find a stable position within the G4 groove, thereby suggesting that pose 1 should give a more viable picture of the recognition between **1d** and *KIT* G4.

Analysis of MD simulations attained starting from the two possible binding mode of the inspected ligand on the G3 structure indicate that both the binding poses are stable during the simulation time. Nevertheless, visualization of the trajectory that the ligand follows during the simulation time reveals that in pose 1 the ligand slightly begins to leave the binding site. To this end, plotting of the distances between the centers of mass calculated for the DNA atoms and the ligand ones outline that the ligand and the DNA move apart during the simulation time(Figure 4.12B). The same analysis attained for pose 2 of the **1d**/G3 complex demonstrates that this interaction is indeed more stable. Further confirmation is also given by the analysis of the rmsd plots calculated considering the DNA backbone atoms that would indicate a certain instability of the G3 topology when the ligand adopts the binding pose 1. In turn, these considerations would allow to give preference to pose 2 in which the ligand remains fairly stable in the G3 site during the MD production run and engages H-bond interactions with the T9 residue and charge-transfer contacts with G1.

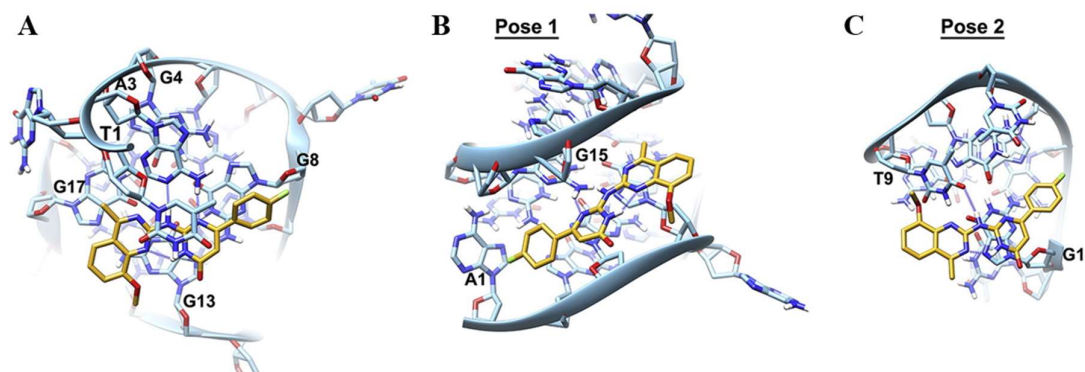


Figure 4.12. Energy-minimized average structures of the complex between **1d** and (A) *MYC*, (pose 1, B) *KYT* G4s, and (pose 2, C) G3. DNA is represented as blue sticks and ribbons while the ligand as yellow sticks. H-bond interactions are indicated as dashed blue lines while only interacting residues are labeled for clarity reasons.

Biological assays

The biological assays were performed in collaboration with Prof. Giovanni Capranico (Department of Pharmacy and Biotechnology, University of Bologna, Italy).

The cytotoxic effect of **1d** was evaluated in human U2OS cells by MTT assay after 1 and 24 h of treatment. Short drug treatments did not show any ability to inhibit cells proliferation at concentrations below 50 μM (Table 4.5). Differently, after 24 h of treatment, the compound demonstrated cell killing activity with $\text{IC}_{50} = 5.9 \mu\text{M}$.

Table 4.5. Cytotoxic activity of compound 1d in human U2OS cells.^a

	$\text{IC}_{50} (\mu\text{M}) - 1 \text{ h}$	$\text{IC}_{50} (\mu\text{M}) - 24 \text{ h}$
1d	> 50	5.9 ± 1.88

^a The concentrations killing 50% of cells (IC_{50}) are means \pm SE of three independent experiments performed in triplicates.

We next asked if **1d** was able to stabilize G4 structures in living cells. Nuclear G4 structures were visualized in human U2OS cells by immunofluorescence microscopy using the BG4 antibody, a known specific antibody against G4s (Biffi et al., 2013). A dose-response analysis was performed treating cells with **1d** for 24 h at concentrations below the IC_{50} (0.2 μM , 0.5 μM , 2 μM). Optical microscopy visualization before the immunostaining, revealed a minimal cell death only at 2 μM concentration. The fluorescence signal generated by incubation with a fluorochrome-labeled secondary antibody was visualized and analyzed. As it is possible to see in Figure 4.13, BG4 nuclear staining in control untreated cells showed a typical dotted signal, indicating that G4 structures were present at discrete chromatin sites in the nucleus, in agreement with previous studies (Amato et al., 2016). The drug treatment determines a slight but significant (p -value < 0.01) increase in the fluorescence intensity in the nucleus at all tested concentrations, demonstrating that **1d** is a stabilizer of G4 structures in human U2OS cells.

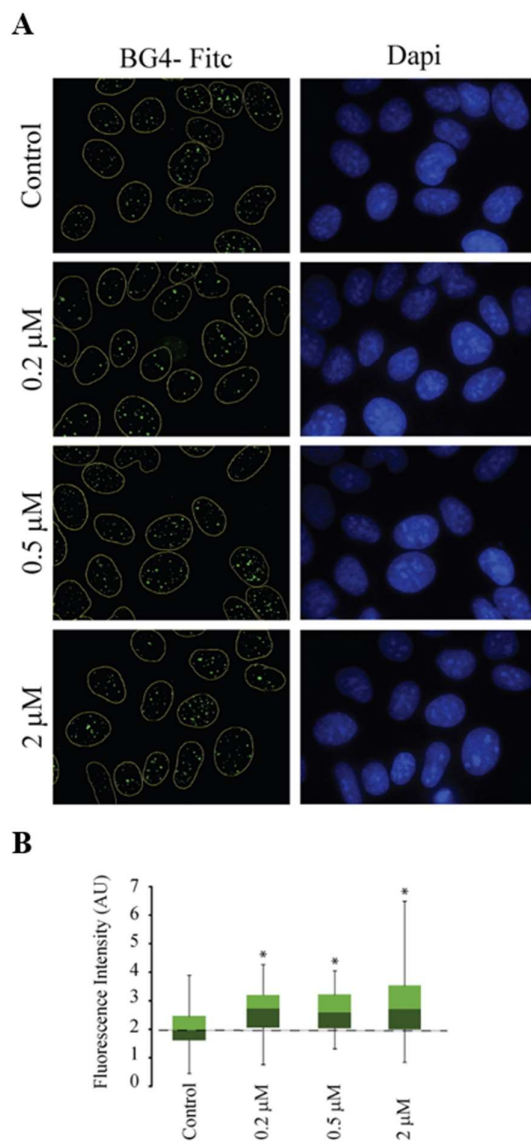


Figure 4.13. (A) Immunofluorescence showing BG4 foci (green) in human osteosarcoma U2OS cell nuclei (blue). Cells were treated with compound 1d for 24 h and then fixed. G4s and DNA were stained with BG4 antibody and DAPI, respectively. (b) Box Plot analysis of G4 signal in the nucleus of control and treated cells. Statistical significance was determined with the Student's *t*-test ($*p < 0.05$).

Taken together the results suggest that **1d** is able to stabilize G4 structures in cultured human cancer cells even if to a lesser extent compared to other molecules developed for being specific G4 stabilizers (Amato et al., 2016). This outcome is also in agreement with the CD-melting data indicating that this molecule is not able to bind and stabilize indiscriminately all the G4 structures.

4.1.3 Conclusions

In this study, a tandem application of VS calculations along with the experimental investigation (CD, NMR, and DSC) was employed to discover a novel G3-targeting ligand as a potential anticancer drug. CD experiments provided evidence that, among the selected chemotypes, compound **1d** was able to interact with and stabilize the G3 structure, which essentially conserved unaltered its native conformation upon interaction. However, the results of CD experiments revealed that the selected compound is actually a dual G3/G4 stabilizer, since it was also able to stabilize the parallel G4 DNA structures, while no duplex stabilizing properties were observed. Molecular modeling calculations were performed to provide a possible molecular basis for the binding of **1d** to the G3 and G4 structures, highlighting the possible structural elements involved in the interaction. The analysis of the biological properties of the selected compound demonstrated that it is endowed with cytotoxic effect on human osteosarcoma cells.

In summary, these findings suggest that the identification of ligands with dual G3/G4 stabilizing properties may represent a new route for the design of G-rich DNA-targeting compounds, thus stimulating further studies aimed at the development of novel anticancer drugs.

4.1.4 Experimental Section

Virtual screening

VS calculations on the G3 structure were performed employing the software Autodock 4.2 (AD4.2) (Cosconati et al., 2010; HUEY et al., 2007; Morris et al., 1998) which is effective in detecting nucleic-acid-binding small molecules (Cosconati et al., 2009, 2012; Di Leva et al., 2013). Thus, AD4.2 was used to dock a diversity set of 58,870 compounds from the commercially available Mcule chemical database (www.mcule.com). Prior to docking calculations, the library was processed according to the ZINC protocol (Irwin and Shoichet, 2005) to generate all the tautomeric and protomeric states for each compound. Given the presence of multiple druggable sites on G3 (PDB ID: 2MKM), the docking search area was chosen to enclose the entire DNA structure. VS results were sorted on the basis of the predicted ligand binding free energies (ΔG_{AD4}). All the solutions with a ΔG_{AD4} greater than -6.00 kcal/mol and a cluster size population lower than 10 out of 100 individuals were discarded. The

remaining solutions were visually inspected to discard all the compounds that were not predicted to establish tight contacts with G3 (i.e., coulombic interactions with the phosphate backbone atoms or stacking interactions with the guanine residues). Subsequently, the remaining solutions were grouped based on their structural similarity to avoid structural redundancy, and the individuals with the lowest ΔG_{AD4} value within each group were finally selected. Finally, these compounds were carefully inspected for good geometries in the predicted binding pose. At the end of the visual inspection, 15 compounds were considered for further experimental investigations.

Oligonucleotide synthesis and sample preparation

The DNA sequences were synthesized using standard β -cyanoethylphosphoramidite solid phase chemistry on an ABI 394 DNA/RNA synthesizer (Applied Biosystem) at the 5- μ mol scale. DNA detachment from support and deprotection were performed by treatment with concentrated ammonia aqueous solution at 55 °C for 12 h. The combined filtrates and washings were concentrated under reduced pressure, dissolved in water, and purified by high-performance liquid chromatography (HPLC) on a Nucleogel SAX column (Macherey- Nagel, 1000-8/46), using buffer A consisting of 20 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ aqueous solution (pH 7.0), containing 20% (v/v) CH_3CN , buffer B consisting of 1 M KCl, 20 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ aqueous solution (pH 7.0), containing 20% (v/v) CH_3CN , and a linear gradient from 0% to 100% B for 30 min with a flowrate 1 mL/min. The fractions of the oligomers were collected and successively desalted by Sep-pak cartridges (C-18). The isolated oligomer was proved to be >98% pure by NMR. In particular, the following oligonucleotides have been used for the experiments: the G3-forming oligonucleotide corresponding to the 3'-end truncation of the thrombin binding aptamer (TBA) sequence d(GGTTGGTGTGG) (11-mer-TBA); two G4-forming sequences from the promoter regions of *KIT* d(AGGGAGGGCGCTGGGAGGAGGG) (c-kit1) and *MYC* d(TGAGGGTGGGTAGGGTGGGTAA) (myc22) oncogenes; and the self-complementary duplex-forming dodecamer d(CGCGAATTCGCG) (ds12). The concentration of the oligonucleotides was determined by UV absorption measurements at 90 °C using appropriate molar extinction coefficient values ϵ ($\lambda=260$ nm) calculated by the nearest neighbor model (Cantor et al., 1970). G3 and G4s were prepared in the appropriate buffer (10 mM potassium phosphate, 70 mM KCl, pH 7.0 for 11-mer-TBA (Cerofolini et al., 2014); 10 mM potassium phosphate, 100 mM KCl, pH 7.0 for c-kit1 (Wei et al., 2015) and 5 mM potassium

phosphate, 20 mM KCl, pH 7.0 for myc22 (Ambrus et al., 2005)) at 10–30 μM oligonucleotide concentration, unless otherwise stated. Samples were heated at 90 $^{\circ}\text{C}$ for 5 min, then gradually cooled to room temperature overnight and stored at 4 $^{\circ}\text{C}$ for 24 h before experiments. Finally, duplex DNA was prepared at 30 μM single strand concentration in 10 mM Li_3PO_4 , 100 mM KCl, pH 7.0 and annealed as reported above.

Circular dichroism spectroscopy

Circular dichroism (CD) experiments were recorded on a Jasco J-815 spectropolarimeter equipped with a PTC-423S/15 Peltier temperature controller. CD spectra were recorded at 5 $^{\circ}\text{C}$ (for G3) and 20 $^{\circ}\text{C}$ (for G4s and ds12) in the wavelength range of 230–360 nm and averaged over 3 scans. The scan rate was set to 100 nm/min, with a 1 s response time, and 1 nm bandwidth. CD spectra were baseline corrected and analyzed using Origin 7.0 software. 10 μM of G4 and 30 μM of G3 and duplex DNA were used. DNA/ligand mixtures were obtained by adding 4 mol equivalents of ligands (stock solutions of ligands were 6 mM in DMSO). CD melting experiments were carried out in the 5–80 $^{\circ}\text{C}$ or 20–105 $^{\circ}\text{C}$ temperature range, at 1 $^{\circ}\text{C}/\text{min}$ heating rate by following changes of CD signal at the wavelengths of the maximum CD intensity. CD melting experiments were recorded in the absence and presence of ligands added to the folded DNA structures. The melting temperatures were determined from curve fit using Origin 7.0 software. ΔT_m values were determined as the difference in melting temperature between the DNA with and without ligands (Pagano et al., 2015). Finally, CD melting experiments were also performed by varying concentration of **1d** added to DNA. In particular, four different [ligand]/[DNA] ratio were examined (1:1, 2:1, 4:1 and 8:1).

Gel electrophoresis

Native gel electrophoresis analysis was carried out on 20% polyacrylamide gel at 5 $^{\circ}\text{C}$, which was run in 1 \times TB (pH 7.5) buffer supplemented with 80 mM KCl. An oligonucleotide concentration of 50 μM was used for each sample. Various amounts (1–10 mol equivalents) of **1d** were incubated with DNA at 5 $^{\circ}\text{C}$ for 1 h before loading. Prior to loading the mixture onto the gel, 1 μL of glycerol solution (60% v/v) was added. The total volume loaded in each well was 10 μL . Gel was imaged by UV-shadowing, using the exposure wavelength of 254 nm (for nucleobase detection).

Differential scanning calorimetry

Differential scanning calorimetry (DSC) measurements were carried out on a nano-DSC (TA Instruments). The 11-mer-TBA DNA sequence was annealed in a buffer solution containing 10 mM potassium phosphate, 70 mM KCl (pH 7.0) and 10% (v/v) DMSO. DNA/**1b** complex was obtained by adding ligand solution (4 mol equivalents) to the folded G3 structure. The 11-mer-TBA concentration was fixed for all DSC experiments to 1.1 mg mL⁻¹. Scans were performed at 0.5 °C/min in the 2–95 °C temperature range. A buffer-buffer scan, under the same conditions, was subtracted from the corresponding buffer-sample endotherm scans and linear-polynomial baselines were drawn for each scan. The corrected thermograms were then normalized per mole of DNA to obtain the corresponding molar heat capacity curves. Model-free enthalpy ($\Delta H^0(T_m)$) for the overall unfolding of DNA structure was estimated by integrating the area under the heat capacity versus temperature curves and represent the average of at least three different heating experiments. T_m values correspond to the maximum of each thermogram peak. DSC curves provided the van't Hoff enthalpy ($\Delta H^0_{v.H.}$), calculated assuming a simple two-state transition (Pagano et al., 2013). The reported errors on thermodynamic parameters are the deviations of the mean from the multiple determinations.

Nuclear Magnetic Resonance experiments

NMR experiments were performed on a 700 MHz Varian UnityINOVA spectrometer. One-dimensional proton spectra of the sample in H₂O were recorded at 5 °C using pulsed-field gradient DPGSE for H₂O suppression. DNA samples were prepared at 0.2–0.5 mM strand concentration in 0.6 mL (H₂O/D₂O 9:1) buffer solution. The solution was either 10 mM KH₂PO₄, 70 mM KCl, 0.2 mM EDTA, pH 7.0 (for 11-mer-TBA) or 25 mM KH₂PO₄, 70 mM KCl, 0.2 mM EDTA, pH 6.8 (for myc22). Aliquots of a stock solution of **1d** in DMSO-d₆ were added directly to the DNA solution inside the NMR tube. The NMR data were processed on an iMAC running iNMR software (www.inmr.net).

Additional docking calculations and Molecular Dynamics (MD) simulations

By following the above-described protocol, compound **1d** was also docked on the G4 target structures from the promoter regions of *KIT* (c-kit1, PDB ID: 4WO2) and *MYC* (myc22, PDB ID: 1XAV) oncogenes. Accurate analysis of the docking results revealed that two possible solutions were achieved for the dockings attained on the G3 and c-kit1 DNA structures, while

a single well-clustered solution was achieved for the dockings attained on the myc22 structure. Thus five [1d]/[DNA] complexes were created, two for the G3 and c-kit1 and one for the myc22 DNA. These binary complexes were further investigated by means of MD simulations carried out with the AMBER 14.0 package software (AMBER 14, University of California, San Francisco). Ligand force-field parameters were derived using the Antechamber program (Wang et al., 2004), and partial charges for the substrates were derived using the AM1-BCC procedure in Antechamber. MD simulations on the aforementioned five complexes were carried out in explicit solvent and periodic boundary conditions (Huppert and Balasubramanian, 2007; Moye et al., 2015), and 22 K⁺ ions were added to the solvent bulk of the G3, *KIT*, and *MYC* complexes, respectively, to maintain neutrality in the systems. First, water shells and counterions were minimized using steepest descent and conjugate gradient algorithms. Then, a minimization of the entire ensemble was performed setting a convergence criterion on the gradient of 0.001 kcal mol⁻¹ Å⁻¹. Equilibration runs were carried out by heating the system to 300 K in 1 ns followed by 1 ns of constant pressure equilibration at 300 K. This was followed by 40 ns MD simulations in the NPT ensemble (constant temperature and pressure). The parm99 Amber force field for nucleic acids, modified using the recently developed parmbsc0 parameters (Cheatham et al., 1999; Cornell et al., 1995; Pérez et al., 2007), was used for the DNA and the counterion, whereas the TIP3P model (Jorgensen et al., 1983) was employed to explicitly represent water molecules. van der Waals and short-range electrostatic interactions were estimated within a 10 Å cutoff, whereas the long-range electrostatic interactions were assessed by using the particle mesh Ewald (PME) method (Essmann et al., 1995) with a 1 Å charge grid spacing interpolated by a fourth-order B-spline, and by setting the direct sum tolerance to 10⁻⁵. Bonds involving hydrogen atoms were constrained by using the SHAKE algorithm (Ryckaert et al., 1977) with a relative geometric tolerance for coordinate resetting of 0.00001 Å. Berendsen's coupling algorithms were employed to maintain constant temperature and pressure with the same scaling factor for both solvent and solutes and with the time constant for heat bath coupling maintained at 1.5 ps (Berendsen et al., 1984). The pressure for the isothermal–isobaric ensemble was regulated by using a pressure relaxation time of 1 ps in Berendsen's algorithm. The simulations of the solvated [1d]/[DNA] models were performed using a constant pressure of 1 atm and a constant temperature of 300 K. Analysis of MD trajectories was attained using the ptraj software (Shao et al., 2007). This software was also used to calculate the average structures of the four complexes that were energy minimized by employing the same geometrical

optimization protocol mentioned above. Visual inspections of the calculated complexes were attained using the MGL Tools (Sanner, 1999) and the UCSF Chimera packages (Pettersen et al., 2004).

Cell cultures and immunofluorescence

Human osteosarcoma U2OS were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Carlo Erba), supplemented with 2 mM L glutamine and 10% FBS (fetal bovine serum). U2OS cells were seeded in 35 mm dishes at a concentration of 100,000 cell/mL. Twenty-four hours after seeding, cells were treated with compound 1d. Cells were then fixed in methanol:acetic acid (3:1), permeabilized with 0.1% triton- \times 100/PBS, and blocking was performed in 2% milk/PBS. Immunofluorescence was performed using standard methods with BG4, anti-FLAG (Cell Signaling Technology), and anti-rabbit Alexa 488-conjugated (Invitrogen) antibodies.

Nuclei were stained with DAPI (Sigma-Aldrich), and coverslips were mounted with Mowiol 4-88 (Sigma-Aldrich). Fluorescence signal was determined using ImageJ software with the following formula: corrected total cell fluorescence (CTCF) = integrated density - (area of selected cell \times mean fluorescence of background readings). Significance was determined by parametric tests "Student's t-test". Plasmid for BG4 antibody purification was kindly obtained by S. Balasubramanian. BG4 protein expression was induced by the autoinduction method as described by Studier (Studier, 2005). BG4 was purified using silica-based resin (Protino Ni-IDA, Macherey Nagel) pre-charged with Ni²⁺ ions, eluted with 250 mM Imidazole/ PBS pH 8.0. The eluted antibody was concentrated and purified from imidazole using Amicon Ultra-15 centrifugal filter units (Millipore).

MTT cell proliferation assay

U2OS cells were seeded in 24-wells plate at a concentration of 20,000 cell/mL. Twenty-four hours after seeding, the compound was added to the cells. After 1 h or 24 h of treatment, the compound was removed, and cells were further cultured in complete drug-free medium for 2 days. Then, MTT (Sigma-Aldrich) was added to each well and incubated for 1 h at 37 °C. Next, the medium was removed and 300 μ L of DMSO (Sigma-Aldrich) were added to cells and incubated for 1 h at room temperature. Finally, 100 μ L of the solution were transferred to a 96-well plate, and the absorbance at 595 nm was measured using a multiplate reader (Perkin

Elmer). The linear regression parameters were determined to calculate the IC_{50} (GraphPad Prism 4.0, Graph Pad Software Inc.).

Chapter 5

TARGETING I-MOTIF-FORMING DNA

In this Chapter, some of well-known G4 ligands have been tested by biophysical methodologies to evaluate if they are also able to interact with the i-motif structure. The data showed that the investigated compounds are actually able to interact with both DNA *in vitro*, thus acting de facto as multi-target-directed agents. Interestingly, while all the compounds stabilize the G4, some of them significantly reduce the stability of the i-motif. The present study highlights the importance, when studying G4-targeting compounds, of evaluating also their behavior toward the i-motif-forming counterpart (Paper V) (Pagano et al., 2018).

5.1 Common G-quadruplex binding agents found to interact with i-motif-forming DNA (Paper V)

5.1.1 Introduction

As mentioned in Chapter 1, i-motif structures that occur throughout the genome may be involved in important biological functions. Moreover, i-motif-targeting molecules have been shown to modulate biological functions (Amato et al., 2014b). For example, stabilization of the i-motif forming sequence in the human telomere was found to inhibit telomerase activity and interfere with telomere biology (Chen et al., 2012; Li et al., 2006); stabilization of i-motif forming sequence in the promoter region of *BCL2* was found to cause an increase in gene expression (Kang et al., 2014; Kendrick et al., 2014) and an i-motif interacting compound was found to downregulate *PDGFR-b* promoter activity (Brown et al., 2017). As it appears that formation of i-motif and/or G4 structures could incite different biological outcomes, it is important to understand the potential structures a compound is able to interact with. In contrast to the hundreds of G4 binding ligands (Amato et al., 2014a, 2018; Di Leva et al., 2013; Li et al., 2013; Pagano et al., 2007, 2010, 2015), there are comparatively very few i-motif binding compounds reported in the literature (Day et al., 2014). Some ligands which were described to bind G4 have also been found to bind i-motif (Fedoroff et al., 2000; Wright et al.; Xu et al.,

2016), so we decided to assess and compare the capability to interact with i-motif-forming DNA of several known bioactive G4 binding agents: Berberine (Franceschin et al., 2006), BRACO-19 (Gowan et al., 2002), Mitoxantrone (Huang et al., 2007), Phen-DC3 (De Cian et al., 2007a), Pyridostatin (Rodriguez et al., 2008), and RHPS4 (Izbicka et al., 1999) (Figure 5.1).

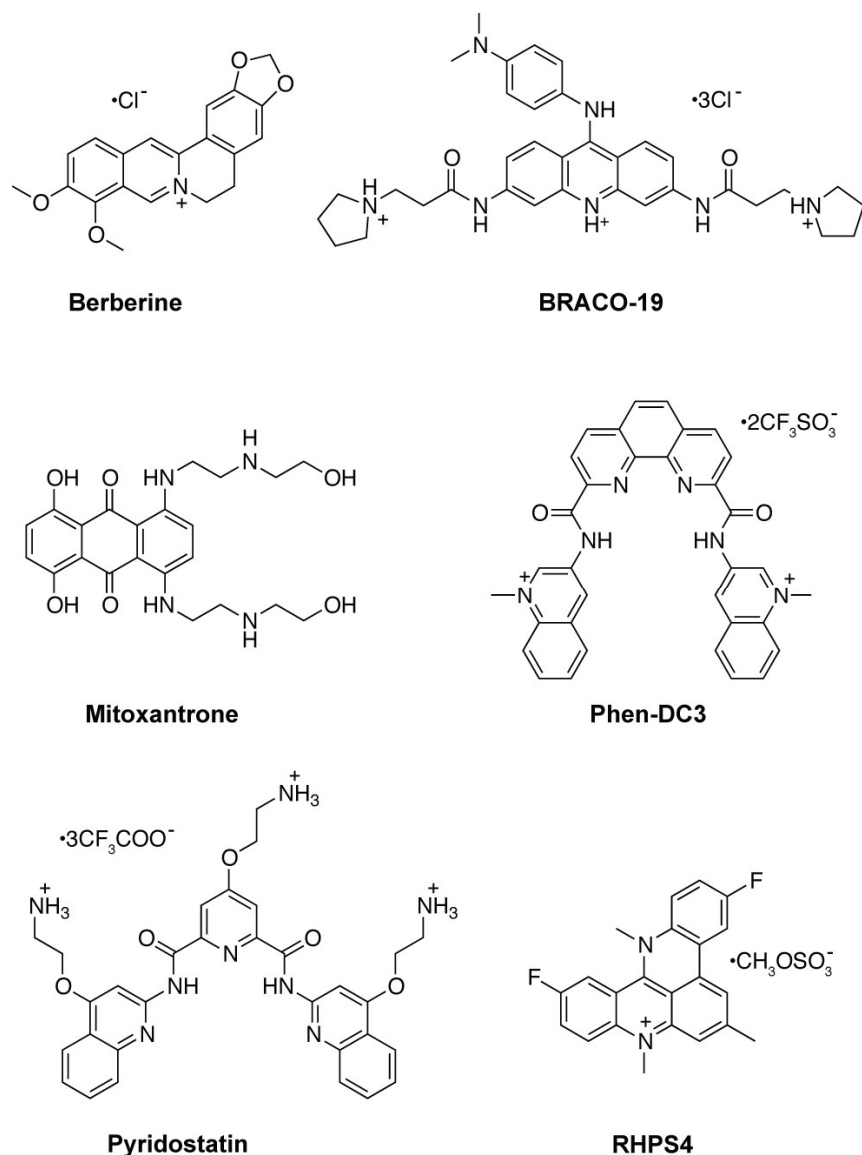


Figure 5.1. Chemical structures of the investigated compounds.

The interaction of these compounds with G4- and i-motif-forming sequences were investigated *in vitro* in different experimental conditions employing several biophysical methodologies (Pagano et al., 2012) (Figure 5.2). The data unequivocally demonstrate that,

even if in different ways, actually these molecules interact with both DNA, thus acting *de facto* as multi-target-directed compounds.

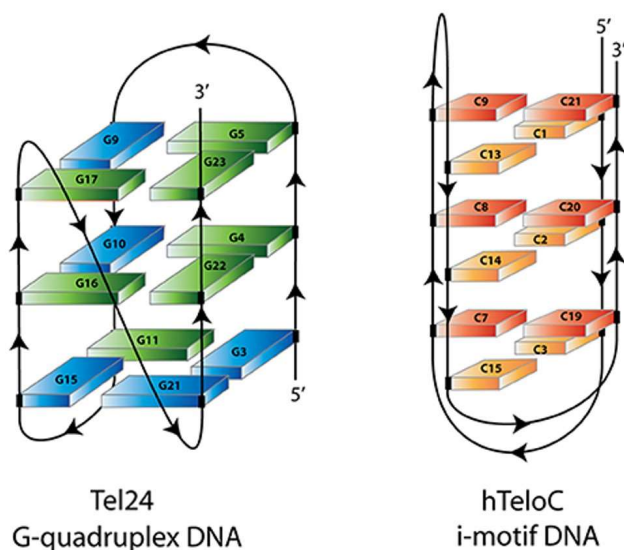


Figure 5.2. Schematic representation of Tel24 G4 and hTeloC i-motif structures. Blue and green solids represent guanines in *syn* and *anti* glycosidic conformation. Orange and red solids represent hemiprotonated C-C⁺ base pairs.

5.1.2 Results

Most of the investigations reported in the literature dealing with the determination of i-motif structures in solution have been accomplished in sodium buffer and under acidic conditions, generally at pH values down to 4.3 (Gallego et al., 1997; Gehring et al., 1993; Malliavin et al., 2003). This is because the cytidine pK_a is about 4.2, and the use of low pH values guarantees to obtain stable hemi-protonated C-C⁺ pairs. However, these conditions are far away from physiological. This may also have consequences on the study of the interaction between these target molecules and potential ligands, which may be differently protonated with respect to their state under physiological pH. Therefore, in order to find experimental conditions as close to physiological as possible, the behavior of the telomeric i-motif-forming sequence (hTeloC) in 10mM sodium buffer at different pH values was investigated by 1D ¹H-NMR and CD spectroscopies. In particular, the pH range 4.3–7.0 was explored. NMR and CD spectra clearly indicated that the hTeloC sequence turned out to be folded into an i-motif structure only

at pH values lower than 5.7, while above this value the i-motif structure is in equilibrium with the unfolded species. Therefore, we decided to perform our studies at the two boundaries pH values, namely 4.3 and 5.7. Since the main aim of this investigation was to evaluate the interaction of some known G4 ligands with an i-motif-forming DNA, experiments on the human telomeric G4 (Tel24) were performed in parallel for comparison.

Circular dichroism studies

The structures adopted by hTeloC and Tel24 were first investigated by circular dichroism (CD) spectroscopy in the absence of ligands. At both pH 4.3 and 5.7, hTeloC showed almost superimposable CD spectra having a positive band at 288 nm and a negative one at around 260 nm (Figure 5.3A). These bands are characteristic of an i-motif folding topology (Guo et al., 2008), Tel24 also displayed almost superimposable CD spectra at both pH values. These spectra were characterized by two positive bands at around 290 and 270 nm, and a negative one at around 240 nm (Figure 5.3B). These bands are perfectly superimposable to those observed for the same molecule at pH 7.0, thus indicating the presence of the expected hybrid [3+1] G4 structure (hybrid-1) as the major conformation under acidic conditions (Gray et al., 2014; Karsisiotis et al., 2011).

CD experiments were also performed to examine the potential of the six investigated compounds to alter the native folding topology of the two investigated DNA structures both at pH 4.3 and 5.7. DNA/ligand mixtures were obtained by adding 5 mol equivalents of compound to the folded G4 and i-motif structures so as to have an excess with respect to potential binding sites.

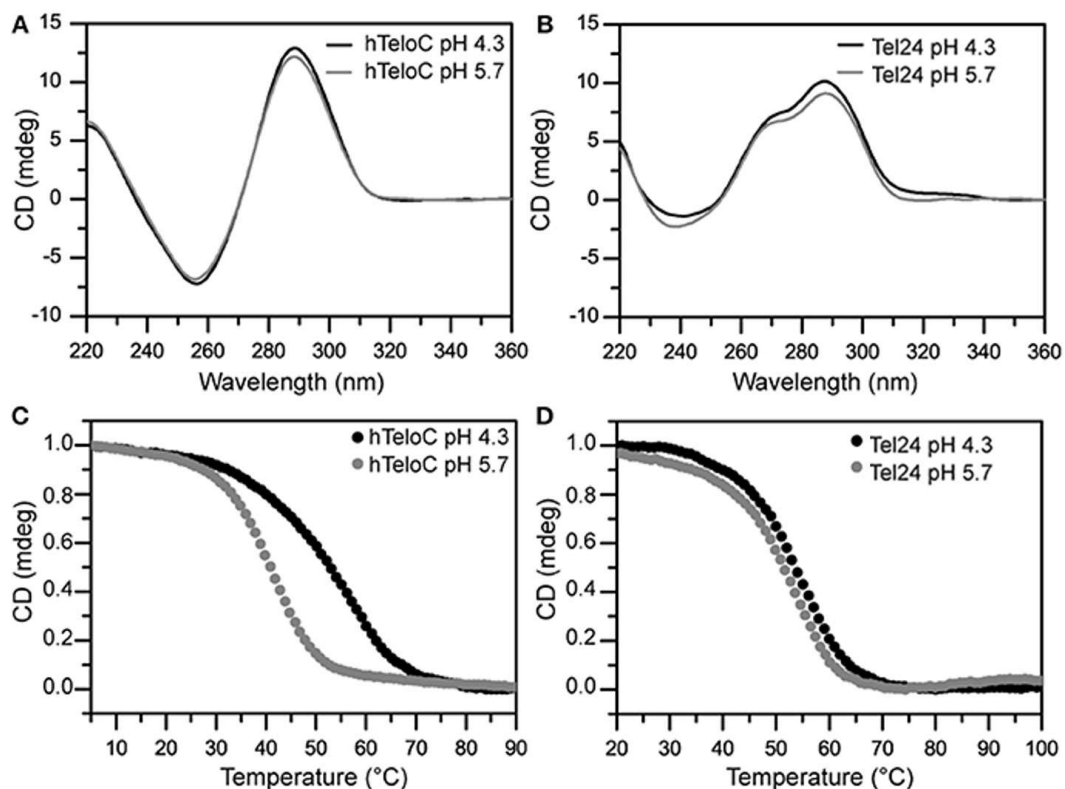


Figure 5.3. CD spectra of hTeloC (A) and Tel24 (B) at pH 4.3 (black) and 5.7 (gray). CD melting of hTeloC (C) and Tel24 (D) at pH 4.3 (black) and 5.7 (gray).

In the case of Tel24, regardless of the pH, Berberine, BRACO-19, Phen-DC3 and RHPS4 induced a significant change in the CD spectrum of the G4 structure (Figure 5.4A-B). In particular, the loss of the band at 270 nm followed by an intensity's increase of the band at 290 nm suggested a conformational change of the G4 topology from the hybrid to the antiparallel conformation (Masiero et al., 2010b; Randazzo et al., 2013). Conversely, Mitoxantrone and Pyridostatin did not produce any measurable conformational change of the G4 structure, even if a decrease of the band at 290 nm is observed upon their addition.

As for the i-motif structure, at pH 4.3, Mitoxantrone, Phen-DC3, and RHPS4 induced a significant hypochromic shift of the positive band at 288 nm, with more marked effects observed for Mitoxantrone and Phen-DC3. Conversely, Berberine, BRACO-19, and Pyridostatin caused a hyperchromic shift of the band at 288 nm at this pH (Figure 5.5A).

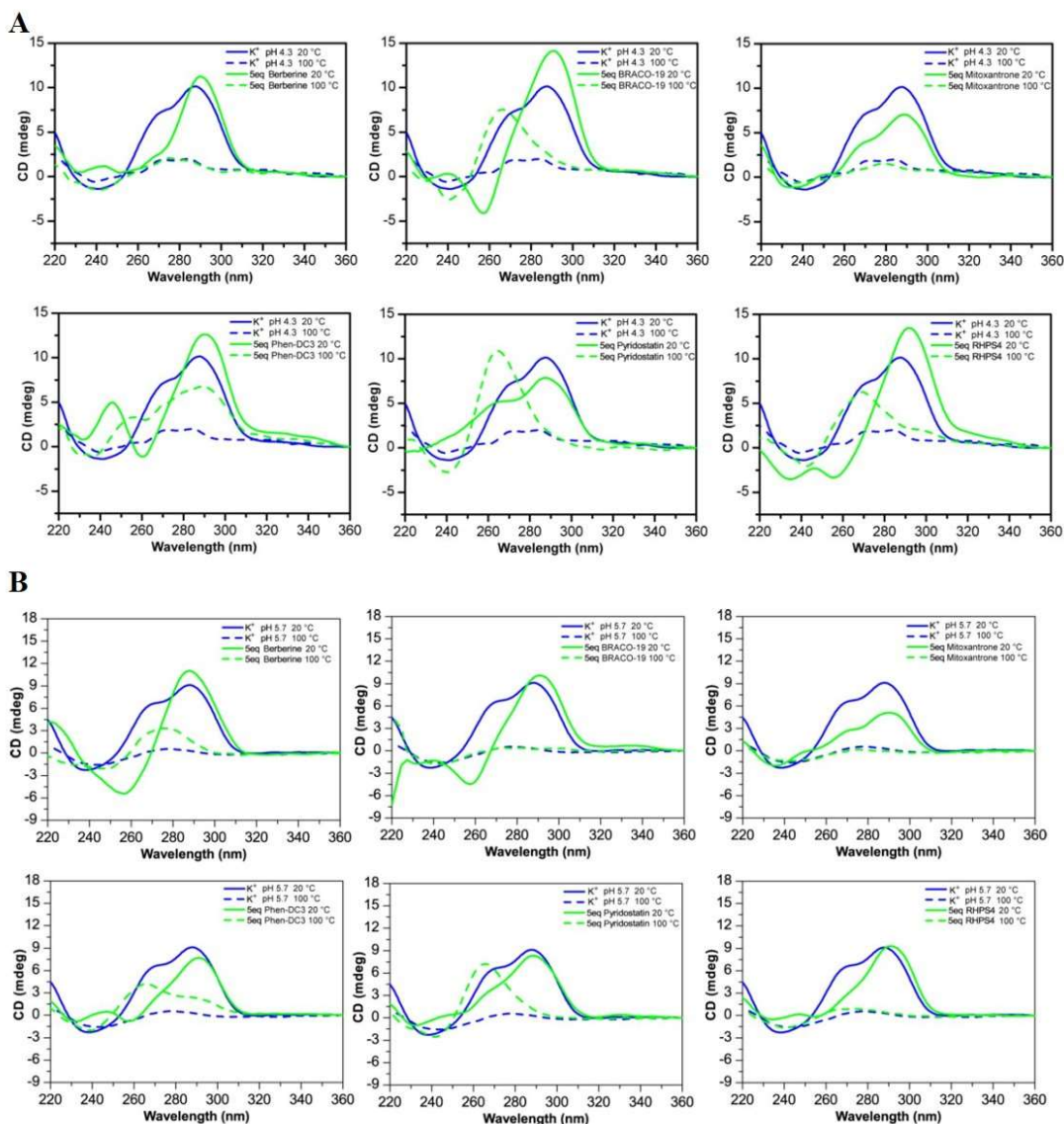


Figure 5.4. (A) CD spectra of Tel24 (10 μ M) in 10 mM KH_2PO_4 at pH 4.3 and (B) CD spectra of Tel24 (10 μ M) in 10 mM KH_2PO_4 at pH 5.7 with different ligands.

On the other hand, at pH 5.7 all the compounds have been shown to induce a hypochromic effect of the band at 288 nm that turned out to be particularly marked in the case of BRACO-19, Mitoxantrone, and Phen-DC3 (Figure 5.5B). These data suggest that some interaction takes place and that, in some cases, the molecules seem to induce the unfolding of the structure.

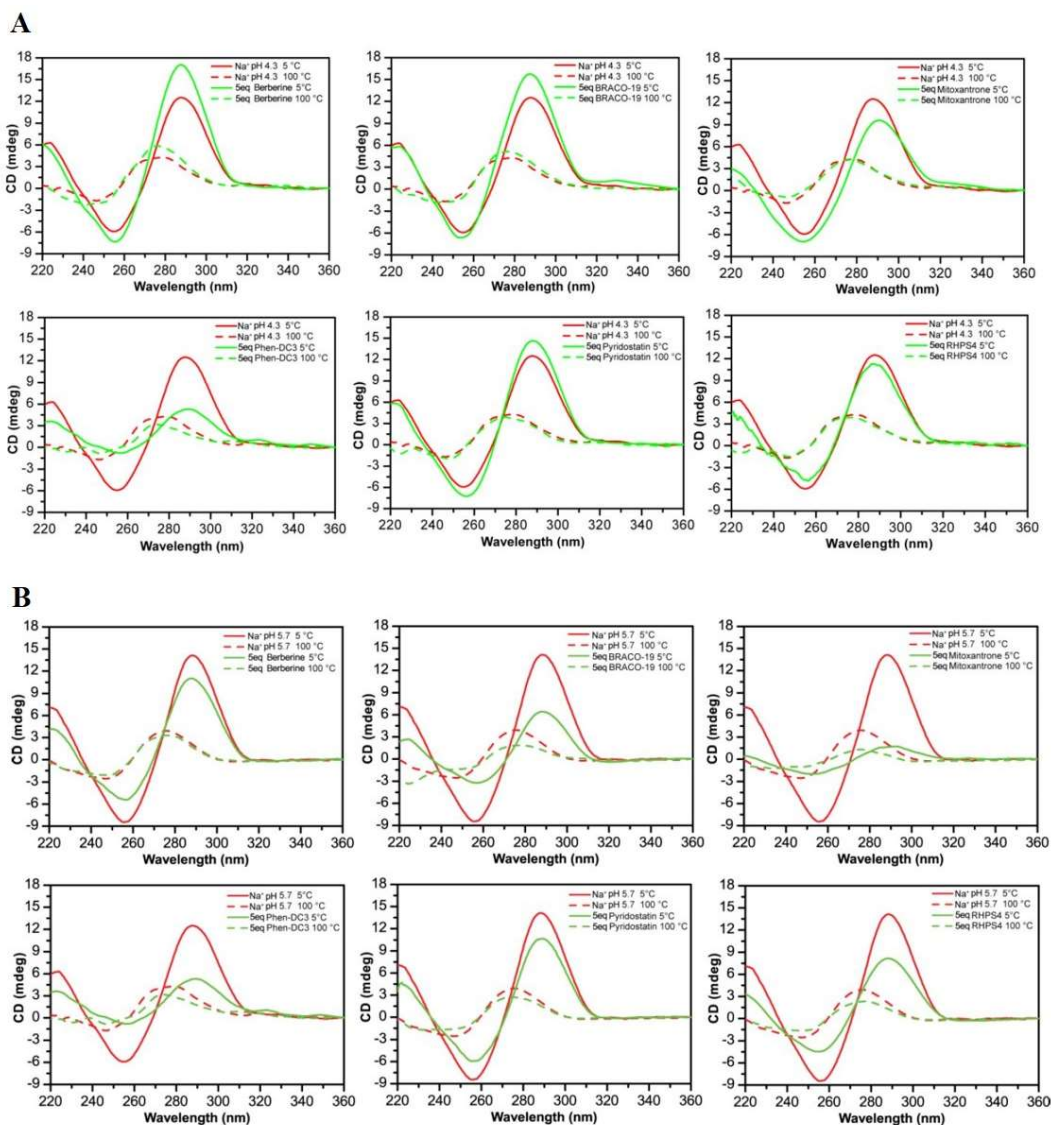


Figure 5.5. (A) CD spectra of hTeloC (10 μ M) in 10 mM NaH_2PO_4 at pH 4.3, and (B) hTeloC (15 μ M) in 10 mM NaH_2PO_4 at pH 5.7 with different ligands.

CD-melting experiments were employed to evaluate the thermal stability of i-motif and G4 structures adopted by hTeloC and Tel24 sequences, respectively, under the two experimental conditions used (pH 4.3 and 5.7). The melting temperatures (T_m) of the i-motif structure were found to be 52.8 and 40.6 $^\circ\text{C}$ at pH 4.3 and 5.7, respectively (Figure 5.6A). The lower thermal stability of the i-motif structure observed at pH 5.7 can be ascribed to the lower extent of protonation of the cytosines. On the other hand, only a very small variation in T_m values was observed for the G4 structure between pH 4.3 and 5.7 (53.5 and 51.1 $^\circ\text{C}$, respectively, Figure 5.6B).

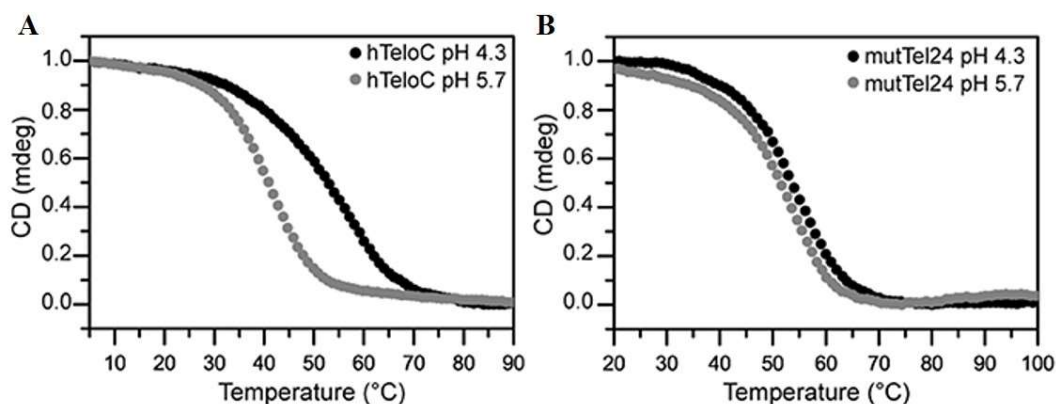


Figure 5.6. CD melting of hTeloC (A) and Tel24 (B) at pH 4.3 (black) and 5.7 (gray).

Then, the effect of the investigated compounds on the stability of the DNA secondary structures was evaluated by measuring the ligand-induced change in the melting temperature (T_m) of G4 and i-motif at pH 4.3 and 5.7. CD-melting curves of DNA in the absence and presence of each compound were obtained by following the variations of the intensity of CD signal at the wavelengths of 290 and 288 nm for Tel24 and hTeloC, respectively (Figure 5.4-5.5). Very intriguingly, results of these experiments (Table 5.1) clearly indicated that the compounds exert a different effect on i-motif compared to G4. As expected, all the tested compounds were able to thermally stabilize the G4 structure at both pH values, even if to a different extent.

Table 5.1. Ligand-induced thermal stabilization of hTeloC and Tel24 DNA measured by CD melting experiments.

	ΔT_m (°C) ^a			
	hTeloC		Tel24	
	pH 4.3	pH 5.7	pH 4.3	pH 5.7
Berberine	-2.5	-0.8	+13.4	+12.4
BRACO-19	-13.4	-9.2	+12.4	+8.9
Mitoxantrone	-4.8	-9.9	+7.7	+4.5
Phen-DC3	-13.4	-6.8	ND	+14.4
Pyridostatin	-2.8	+0.8	+12.9	+8.8
RHPS4	-1.0	-0.3	+22.0	+20.6

^a $\Delta T_m = T_m$ (DNA+ligand) - T_m (DNA). All experiments were performed in duplicate, and T_m values are reported as the mean. Errors were ± 0.5 °C. ND, not determined. Errors were ± 0.5 °C. ND, not determined.

On the contrary, regardless of pH, Berberine, Pyridostatin, and RHPS4 did not show a remarkable influence on the i-motif thermal stability, while BRACO-19, Mitoxantrone, and Phen-DC3 significantly decreased it (Figure 5.7A and B). These results are in agreement with those obtained from the CD spectra analysis.

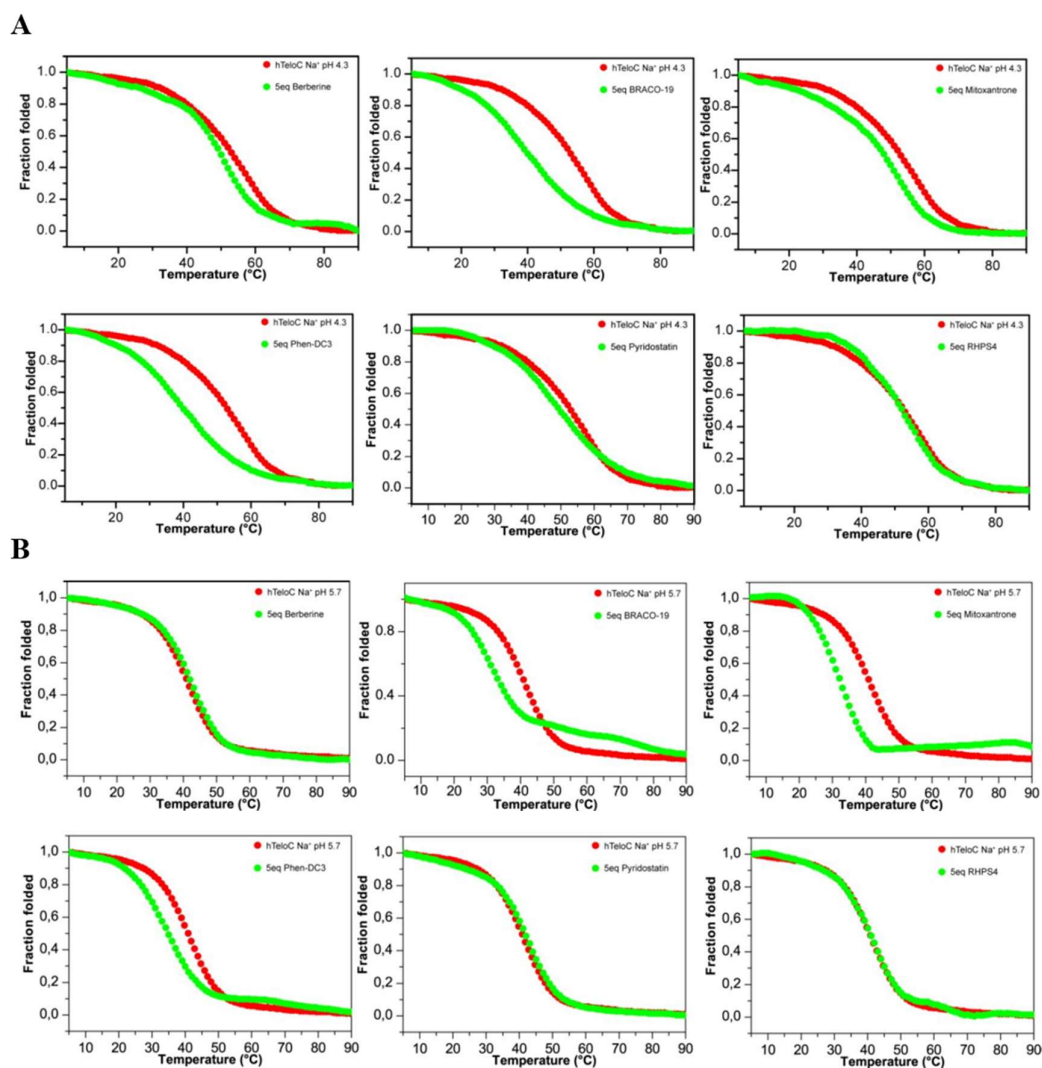


Figure 5.7. CD melting of hTeloC (10 μ M) in 10 mM NaH₂PO₄ at pH 4.3 (A), and hTeloC (15 μ M) in 10 mM NaH₂PO₄ at pH 5.7 (B) with different ligands at 288 nm.

UV melting studies

The effect of compounds 1-6 on the stability of the G4 and i-motif structures was also investigated by UV-melting experiments. As for CD-melting studies, the ligand-induced

changes in the melting temperature (T_m) of the two DNA structures were obtained by recording UV-melting experiments in the absence and presence of each compound at both pH 4.3 and 5.7. UV-melting curves were acquired by following the change in UV signal intensity at 295 nm for both Tel24 and hTeloC. Results of these experiments (Table 5.2) are consistent with CD melting ones, and denote, once again, a different behavior for the investigated compounds toward the i-motif and G4 DNA structures.

Table 5.2. Ligand-induced thermal stabilization of hTeloC and Tel24 DNA measured by UV melting experiments

	ΔT_m (°C) ^a			
	hTeloC		Tel24	
	pH 4.3	pH 5.7	pH 4.3	pH 5.7
Berberine	+1.2	+0.9	+11.7	+15.3
BRACO-19	-17.6	-6.5	ND	ND
Mitoxantrone	-16.6	-1.0	+4.1	+0.6
Phen-DC3	-17.1	-6.3	ND	+5.3
Pyridostatin	-1.1	+0.2	+9.3	+5.5
RHPS4	+0.7	-0.1	ND	ND

^a $\Delta T_m = T_m$ (DNA+ligand) - T_m (DNA). All experiments were performed in duplicate, and T_m values are reported as the mean. Errors were ± 0.5 °C. ND, not determined. Errors were ± 0.5 °C. ND, not determined.

Nuclear Magnetic Resonance Studies

NMR spectroscopy was employed in order to obtain structural information about the DNA interaction of the six compounds. Also in this case, 1D ¹H-NMR spectra were recorded at pH 4.3 and 5.7. Under the experimental conditions used, the Tel24 sequence forms a single G4 conformation characterized by 12 well-resolved imino proton peaks, corresponding to the 12 guanines involved in the three G-tetrad planes (Luu et al., 2006b). On the other hand, hTeloC folds in an i-motif structure characterized by 3 well-resolved imino proton peaks that correspond to the 6 intercalated C-C⁺ pairs (Phan et al., 2000). The imino and aromatic proton regions of Tel24 and hTeloC in the absence and presence of 5 mol equivalents of each compound are shown in Figures 5.8-5.11. Regardless of the pH, both imino and aromatic proton signals of Tel24 turned out to be significantly affected by the addition of the ligands (Figure 5.8-5.9).

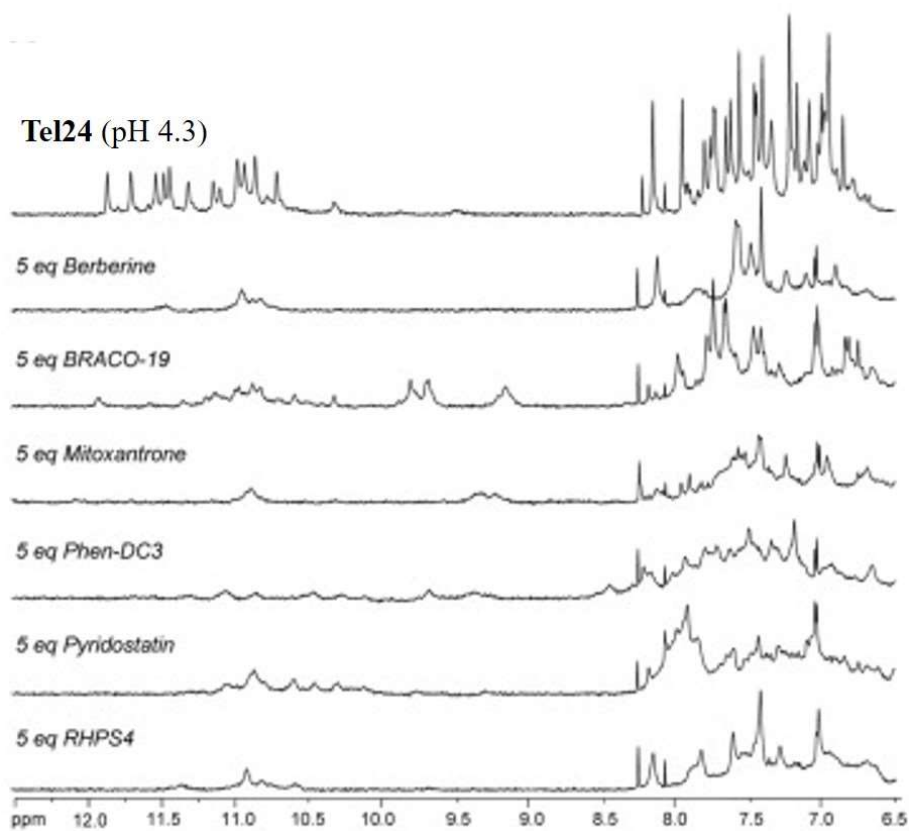


Figure 5.8. ¹H-NMR spectra of Tel24 in 10 mM KH₂PO₄ at pH 4.3, before and after the addition of 5 mol equivalents of each ligand.

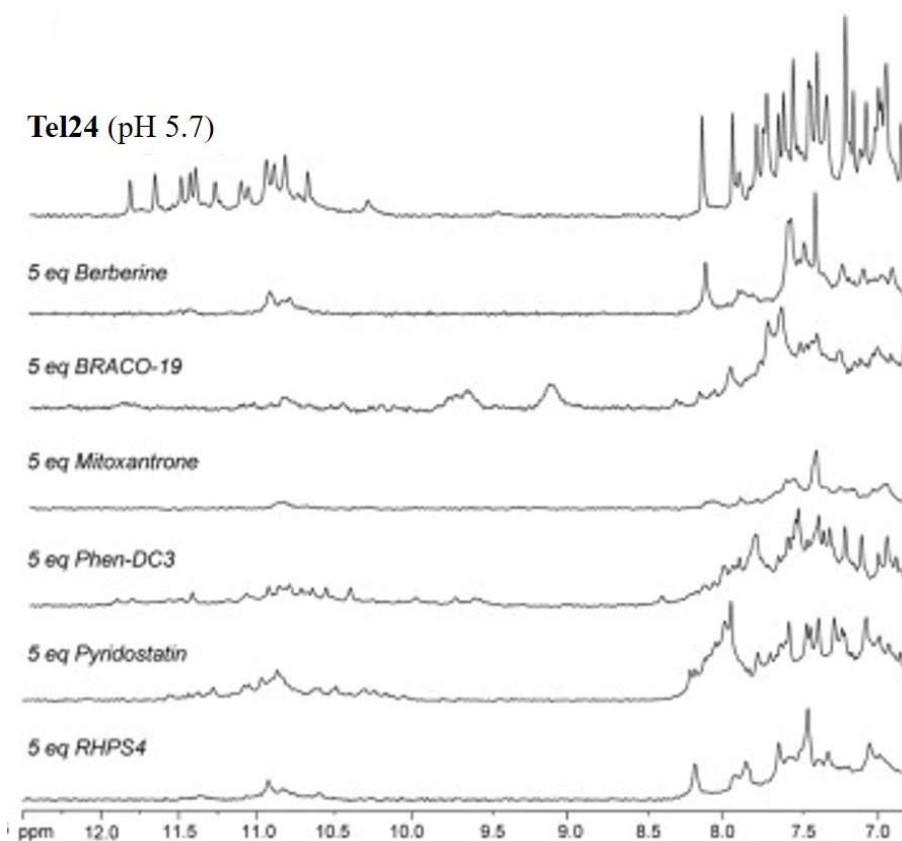


Figure 5.9. ^1H -NMR spectra of Tel24 in 10 mM KH_2PO_4 at pH 5.7, before and after the addition of 5 mol equivalents of each ligand.

On the other hand, addition of compounds to the i-motif structure led to different results. At pH 4.3, the main changes were observed for Berberine and RHPS4 in both aromatic and imino regions, while little changes could be observed for BRACO-19 and Mitoxantrone. Instead, Phen-DC3 caused a general decrease of signal intensities (Figure 5.10). Very little changes could be observed in the spectrum of i-motif upon addition of Pyridostatin. At pH 5.7, changes were observed for Berberine, BRACO-19, Mitoxantrone, and RHPS4 (Figure 5.11), with BRACO-19 affecting the most the NMR spectrum of the i-motif structure. Interestingly, as for the experiment at pH 4.3, Phen-DC3 and Pyridostatin caused a general decrease of the signal intensities (Figure 5.10).

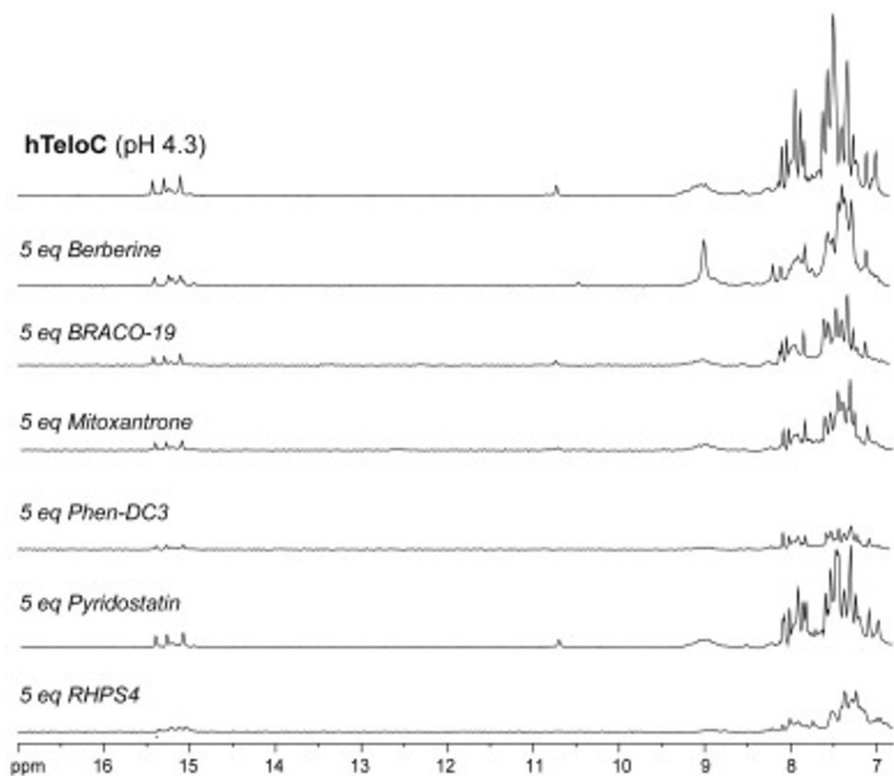


Figure 5.10. ¹H-NMR spectra of hTeloC in 10 mM NaH₂PO₄ at pH 4.3, before and after the addition of 5 mol equivalents of each ligand.

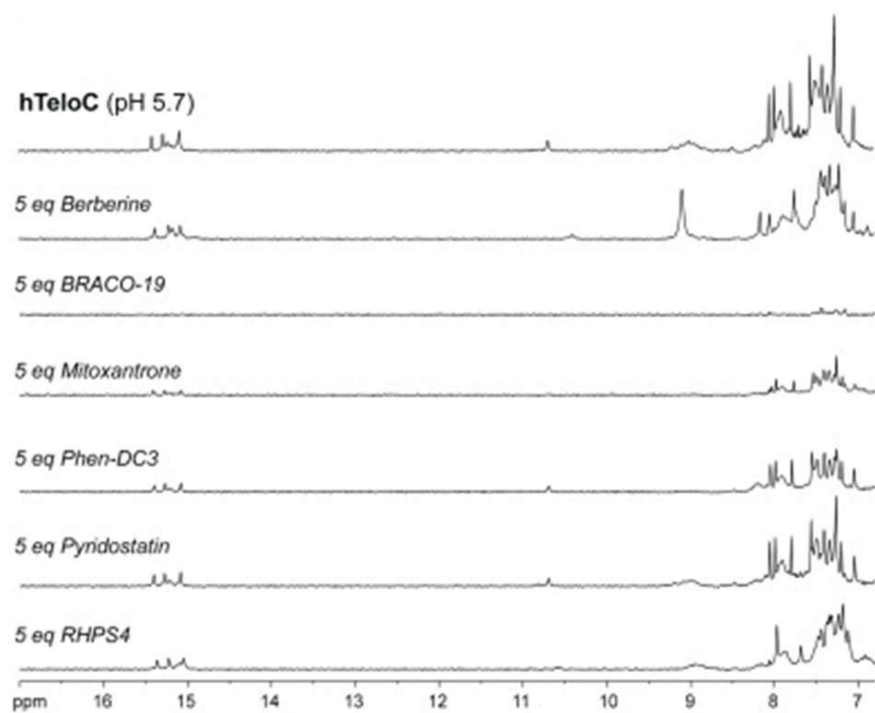


Figure 5.11. ¹H-NMR spectra of hTeloC in 10 mM NaH₂PO₄ at pH 5.7, before and after the addition of 5 mol equivalents of each ligand.

FRET studies

Dual labeled (FAM/TAMRA) human telomeric sequences G4- F21T and i-motif-F24T, which are able to form G4 and i-motif structures, respectively, were used. Experiments were performed only at pH 5.7, because the fluorescence of the FAM is not stable at pH 4.3. In order to verify that the selected compounds did not interfere in the FAM emission spectrum, the fluorescence spectrum of each compound was recorded by exciting at 492 nm and collecting its emission spectrum between 500 and 650 nm. Unfortunately, the emission spectrum of RHPS4 was found to overlap with FAM, and, therefore, it was not used in these experiments. FRET-melting assays were then performed to further investigate the compound-induced effects on the i-motif and G4 thermal stabilities. In agreement with CD- and UV-melting experiments, all the tested compounds induced a thermal stabilization of G4-F21T. Conversely, in the case of i-motif-F24T, FRET melting data did not agree at all with CD and UV melting results. In fact, all ligands, except Berberine, showed a significant thermal stabilization of the i-motif structure. FRET spectra of FAM/TAMRA-modified oligonucleotides in the absence and presence of each compound (at 1:5 DNA/ligand ratio) were also analyzed (Figure 5.12). The results show that, in both cases, BRACO-19, Mitoxantrone, Phen-DC3, and Pyridostatin caused a significant decrease of the band intensity at 580 nm, thus suggesting that actually they could interact with the probes.

Fluorescent Intercalator Displacement (FID) Assay

FID assays are long established and have been used extensively to determine relative binding affinities for ligands with duplex DNA (Boger et al., 2000). They have recently been adapted and validated for use with both G4 and i-motif structures (Monchaud et al., 2006; Sheng et al., 2017). The assay relies on a light-up fluorescent probe, in this case thiazole orange (TO), which binds to the structure of interest and can be competitively displaced by candidate ligands thus enabling the determination of their relative binding affinity to the structure under examination. Here, aliquots of the examined ligands were titrated in triplicate against both of the G4 and i-motif structures and the concentrations at which 50% displacement (DC_{50}) was achieved were calculated from dose-response curves fitted to this data. Unfortunately, as with the FRET, RHPS4 was excluded from analysis due to the fluorescence profile of the ligand overlapping with the assay parameters. The results of the FID assay showed that all the examined ligands bound to both the G4 and the i-motif DNA. Unsurprisingly, they also all showed a slightly

higher affinity for the G4 formed by Tel24FID over the i-motif formed by hTeloCFID (Table 5.3). Nevertheless, all ligands tested were found to displace TO at both pH 4.3 and 5.7, indicative of an interaction with the i-motif-forming DNA.

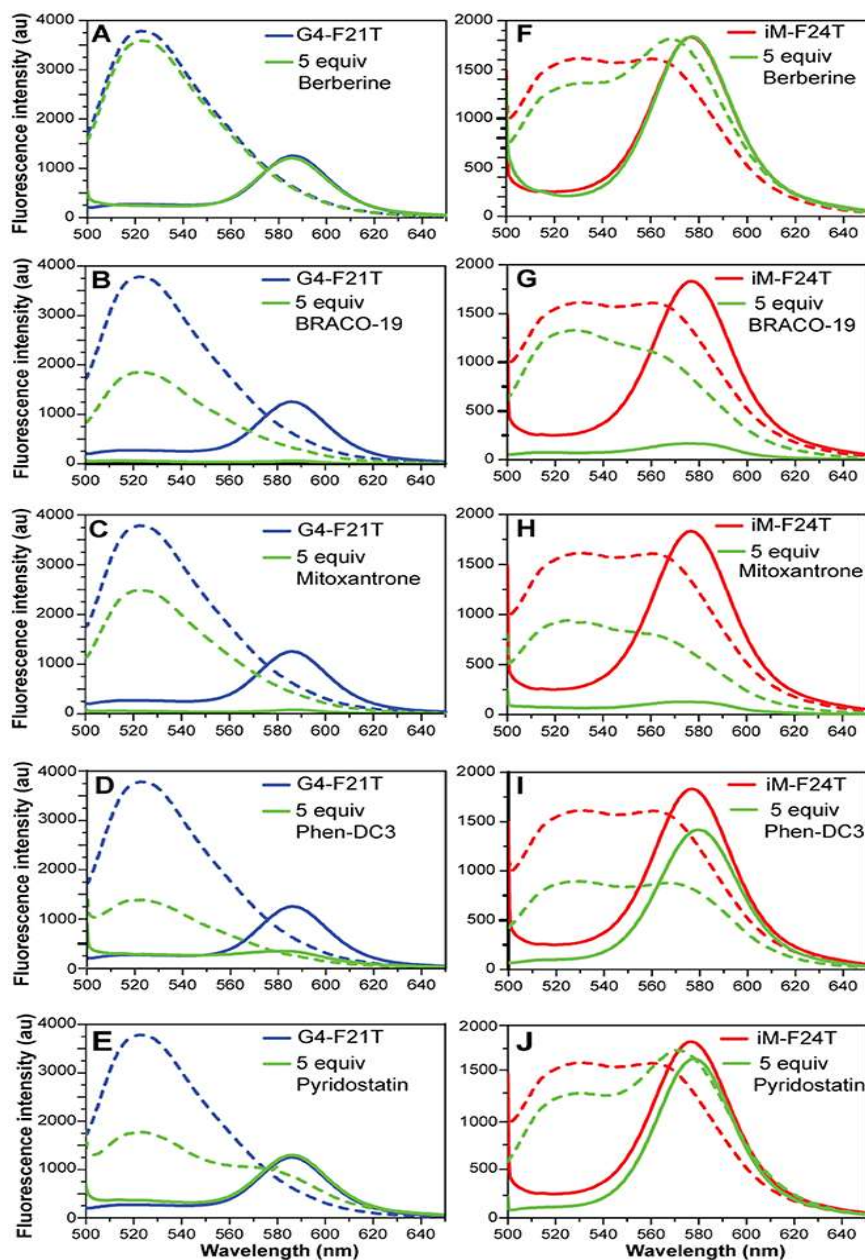


Figure 5.12. (A–E) Fluorescence spectra of G4-F21T alone (in 10mM KH_2PO_4 buffer) at 5 °C (blue line) and at 90 °C (dashed blue line), and after the addition of 5 equivalents of each ligand at 5 °C (green line) and at 90 °C (dashed green line). (F–J) Fluorescence spectra of i-motif-F24T alone (in 10mM NaH_2PO_4 buffer) at 5°C (red line) and at 90 °C (red dashed line), and after the addition of 5 equivalents of each ligand at 5 °C (green line) and at 90 °C (dashed green line). All the experiments were performed at pH 5.7.

Table 5.3. Ligand DC₅₀ values for hTeloCFID and Tel24FID determined using the FID assay*.

	hTeloCFID				Tel24FID			
	pH 4.3		pH 5.7		pH 4.3		pH 5.7	
	DC ₅₀ (μ M)	SE (μ M)	DC ₅₀ (μ M)	SE (μ M)	DC ₅₀ (μ M)	SE (μ M)	DC ₅₀ (μ M)	SE (μ M)
Berberine	30.38	1.58E-02	1.46	5.88E-03	3.32	2.63E-02	1.26	7.22E-03
BRACO-19	0.66	1.61E-03	0.87	8.08E-04	0.26	5.14E-03	0.50	1.95E-03
Mitoxantrone	0.70	1.61E-03	1.34	6.99E-03	0.54	3.03E-04	0.95	6.07E-03
Phen-DC3	0.97	1.32E-02	0.95	3.79E-03	0.26	1.58E-03	0.39	1.42E-03
Pyridostatin	9.09	6.18E-03	18.02	5.89E-02	3.15	2.33E-03	9.42	3.90E-02

*All experiments were performed in triplicate and DC₅₀ values are reported as the 50% displacement value calculated from fitted dose response curves. Standard errors are calculated using R-square values from the statistics on the data fit.

5.1.3 Discussion

DNA has a well-known propensity to adopt various alternative non-canonical conformations in vitro, including G4 and i-motif. Recent investigations have demonstrated the formation of such structures in regulatory regions of the human genome, including gene promoters and telomeres, also providing evidences for the key role that G4s and i-motifs can play in several biological pathways (Kang et al., 2014; Maizels, 2015; Salvati et al., 2014). Human telomeric DNA consists of a 2–20 kb double-stranded region composed by (TTAGGG)/(CCCTAA) repeats, and of a single-stranded 3'-end G-rich sequence. The G-rich strand can adopt G4 conformations, while the opposite C-rich strand can fold into the i-motif structure. Of the two structures, telomeric G4 is by far the most studied. This disparity is mainly due to acidic pH required for the protonation of cytosine, since the parallel duplexes, the basic component of i-motif, are stabilized by hemiprotonated C–C⁺ base pairs. In this investigation, the behavior of the human telomeric i-motif-forming sequence (hTeloC) was studied in the pH range of 4.3–7.0 by 1D ¹H-NMR and CD spectroscopies. Results clearly indicate that under the experimental conditions used, the i-motif structure is still well preserved at pH 5.7, while, just above this pH value, it turns out to be in equilibrium with the random coil. Moreover, the lower stability of the i-motif structure observed at pH 5.7 well reflects the lower extent of cytosine protonation compared to pH 4.3, thus confirming, once again, that i-motif structures are very sensitive to pH (Kovanda et al., 2015). On the other hand, results of NMR and CD experiments on the G-

rich telomeric sequence clearly showed that it retains the hybrid [3+1] G4 structure as major conformation at both pH 4.3 and 5.7, and that its thermal stability is basically not affected by the pH. Once the working conditions were established, the interaction of the well-known G4 ligands Berberine, BRACO-19, Mitoxantrone, Phen-DC3, Pyridostatin, and RHPS4, with the i-motif-forming sequence was explored in comparison with the G4, by using a combination of spectroscopic techniques. Experiments were performed at the two boundary pH values. Results of the different experiments undoubtedly showed that all the investigated compounds actually are able to interact, even if in a different way, with both G4- and i-motif-forming DNA. Fluorescent intercalator displacement (FID) assay clearly showed that the thiazole orange probe, which binds to the investigated structures, is competitively displaced by the tested compounds (RHPS4 could not be used). The consequent determination of their relative affinity for the DNA under examination reveals that, regardless of the pH (4.3 or 5.7), the compounds exhibit only a slightly higher affinity for Tel24 over hTeloC. Ligand-induced effects on both G4 and i-motif structures were examined by means of UV, CD, and NMR. As far as Tel24 is concerned, results of CD experiments show that four (Berberine, BRACO-19, Phen-DC3, and RHPS4) out of the six ligands induce a conformational change from the hybrid [3+1] to an antiparallel structure. All the ligands also indiscriminately affect the NMR spectrum of Tel24 (Figure 5.7). Results of CD- and UV-melting experiments agree in indicating that the six compounds are able to stabilize the G4 at both pH values. Interestingly, the two ligands not inducing the G4 conformational change (Mitoxantrone and Pyridostatin) seem also to be the less effective in terms of thermal stabilization under the experimental conditions used. Moreover, some of the investigated ligands induced a higher increase of G4 melting temperature (T_m) at pH 4.3 rather than 5.7 (Table 5.1), that may be ascribed in part to the different protonation states of such molecules at the different pHs. This was, for example, the case of BRACO-19 and Pyridostatin. For these molecules the protonation state was theoretically determined in the 3.5–7.5 pH range, by using the pK_a prediction program (ChemAxon, www.chemaxon.com) based on the calculation of partial charge of atoms in the molecule. Concerning the i-motif, some differences were observed between pH 4.3 and 5.7. In particular, the addition of compounds to the i-motif caused major effects at pH 5.7, in which the i-motif structure is less stable per se. Interestingly, CD- and UV-melting results clearly indicate that the molecules which were found to mainly decrease the CD and NMR signals (namely BRACO-19, Mitoxantrone, and Phen-DC3) were also those that significantly

decreased i-motif thermal stability at both pH values. Overall, these results show that tested compounds are able to interact with the telomeric i-motif-forming DNA. However, three of them (Berberine, Pyridostatin, and RHPS4) do not have relevant effects on the thermal stability of i-motif, while the others (BRACO-19, Mitoxantrone, and Phen-DC3) are able to destabilize it. We speculate that these compounds could reasonably make some non-specific interactions with the single-stranded (unfolded) C-rich DNA, resulting in a shift of folded-unfolded equilibrium toward the unfolded form, especially during the melting experiment, which in turn results in a decrease of T_m . The FRET methodology, used here to further characterize the interaction of G4 and i-motif DNA with the investigated ligands, deserves a separate discussion. With respect of other spectroscopic techniques (such as UV, CD, and NMR), FRET has a higher sensitivity and it can explore a large range of ligand concentrations (Monchaud et al., 2006; Sheng et al., 2017). Additionally, it turns out to be the main methodology when the UV absorbance of a ligand overlaps with that of the DNA (Guédin et al., 2010). However, some artifacts may occur when compounds are inherently fluorescent and/or interact with the fluorescent probes rather than the DNA itself (De Cian et al., 2007b), and this is what probably happened in this case. Indeed, FRET melting results did not agree at all with both CD and UV melting data, especially in the case of i-motif-F24T. To understand the reasons for this different behavior, FRET spectra of labeled DNA (G4-F21T and i-motif-F24T) in the absence and presence of compounds were analyzed. Typically, when DNA is folded, the two dyes are in close proximity so FAM fluorescence peak at 522 nm (upon excitation at 492 nm) is quenched and its energy is transferred to TAMRA, which then emits light at 580 nm. On the other hand, FAM's fluorescence is no longer quenched when sufficient spatial separation of the two dyes occurs (for example upon unfolding of the DNA structure), therefore its fluorescence signal at 522 nm is observable. In principle, compounds that are able to interact with the fluorophores may affect the emission properties of the probes and decrease the intensity of the bands at 580 nm (if the DNA is structured) or at 522 nm (if the DNA is unstructured). Interestingly, four compounds (BRACO-19, Mitoxantrone, Phen-DC3, and Pyridostatin) caused a significant decrease of the band intensity at 580 nm (for both G4-F21T and i-motif-F24T), clearly suggesting that they actually interact with the fluorophores. Therefore, the different stabilizing effects observed for the i-motif structure across the different spectroscopic techniques could be ascribed to ligand interaction with the FRET probes. This hypothesis is further corroborated by the fact that Berberine, which did not cause any observable change in the FRET spectrum,

showed no variation in the DNA's thermal stability by FRET, in agreement with CD and UV experiments. For the same reasons, this suggests that the thermal stabilizations of G4- F21T measured by FRET are not accurate, being also potentially affected by the ligands' interaction with the probes. Therefore, a clear message came out from a careful examination of FRET data, meaning that false-positive responses can be obtained due to ligands ability to bind end-labeling DNA probes. This could occur particularly when the investigated compounds have an extended aromatic core for which π - π stacking interactions with the large aromatic surface of the probes could be favored. This is especially the case of G4-interacting molecules. Overall, this study emphasizes the need of using a combination of techniques when examining DNA targeting ligands, in order to avoid an inaccurate evaluation of their binding/stabilizing properties.

5.1.4 Conclusions

In this study, a combination of spectroscopic techniques have allowed to demonstrate that some of well-known bioactive G4 ligands (Berberine, BRACO-19, Mitoxantrone, Phen-DC3, Pyridostatin, and RHPS4) are able to interact with an i-motif forming DNA. Two human telomeric DNA sequences able to form i-motif and G4 structures were studied and the experiments performed at two different pH values. The experimental results showed that all the investigated G4 ligands were also able to interact with the telomeric i-motif-forming DNA. Very interestingly, BRACO-19, Mitoxantrone, and Phen-DC3 have been shown to destabilize the i-motif structure.

The i-motif formation is very sensitive to the environment conditions and the majority of i-motif forming sequences are less stable than G4s under physiological conditions. Interestingly, recent investigations have suggested that G4 and i-motif structures may have opposing functions in the control of oncogene transcription: while G4 formation and its ligand-induced stabilization generally inhibits gene expression, stabilization of i-motif seems to have transcription activating capabilities (Kang et al., 2014; Kendrick et al., 2014). Therefore, a molecule that is able to stabilize a G4 structure and to destabilize an i-motif structure may exert a synergistic effect on the inhibition of transcription. These are the cases of BRACO-19, Mitoxantrone, and Phen-DC3, whose biological activity may be ascribed to both mechanisms.

Overall, the present study highlights the necessity, when studying G4-targeting compounds, of evaluating also their effects on the i-motif counterparts, especially if one is looking for a "specific" drug.

5.1.5 Experimental Section

Oligonucleotide Synthesis and Sample Preparation

An ABI 394 DNA/RNA synthesizer (Applied Biosystem) was employed to prepare DNA sequences at 5- μ mol scale by using standard β -cyanoethylphosphoramidite solid phase chemistry. The subsequent detachment of DNA from support and its deprotection were carried out by means of an aqueous solution of concentrated ammonia at 55 °C for 12 h. The filtrates and the washings were combined and concentrated under reduced pressure, solubilized in water, and then purified by high-performance liquid chromatography (HPLC) equipped with a Nucleogel SAX column (Macherey-Nagel, 1000-8/46). Two buffers were employed for the purification: buffer A, consisting of a 20mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ aqueous solution (pH 7.0) and containing 20% (v/v) CH_3CN , and buffer B, consisting of 1M KCl, 20mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ aqueous solution (pH 7.0), containing 20% (v/v) CH_3CN , combined in a 30 min linear gradient going from 0 to 100% B with a flow rate of 1 mL/min. The purified fractions of the oligomers were then desalted by using C-18 cartridges (Sep-pak). The purity of the isolated oligomer was evaluated by NMR and it turned out to be higher than 98%. In particular, the following oligonucleotides were employed for the experiments: $d[\text{CCCT}(\text{AACCCT})_3]$ (hTeloC) and $d[(\text{TT}(\text{GGGT})_3\text{GGGA})]$ (Tel24). The oligonucleotide concentrations were established by measuring the UV absorption at 90 °C taking into account the molar extinction coefficient values + (l = 260 nm) determined by the nearest neighbor model (Cantor et al., 1970). hTeloC and Tel24 were dissolved in 10mM sodium phosphate buffer and 10mM potassium phosphate buffer, respectively, at different pH values before establishing the experimental conditions to be used (pH 4.3 and pH 5.7). DNA samples were heated at 90 °C for 5min, and then gradually cooled to room temperature overnight.

Circular Dichroism Spectroscopy

Circular dichroism (CD) experiments were recorded on a Jasco J-815 spectropolarimeter equipped with a PTC-423S/15 Peltier temperature controller. Each spectrum was recorded in the 220–360 nm wavelength range, averaged over three scans and subtracted of the buffer baseline. The scan rate was set to 100 nm/min, with a 1 s response time, and 1 nm bandwidth. Spectra were analyzed using Origin 7.0 software. CD experiments (spectra and melting) were performed using 10–15 μ M oligonucleotide concentration, in the absence and presence of 5 molar equivalents of ligands (10mM in DMSO). CD melting were performed at 1 °C/min

heating rate in the 5–90 and 20–100 °C temperature range for hTeloC and Tel24, respectively. Changes of CD signal were followed at the wavelengths of the maximum CD intensity, 288 and 290 nm for hTeloC and Tel24, respectively. The melting temperatures (T_m) were mathematically calculated by using the curve fitting function in Origin 7.0 software. T_m values represent the difference between the melting temperature of the DNA with and without ligands.

UV-Melting

A JASCO V-730 UV-visible spectrophotometer equipped with a Peltier temperature controller was employed to perform the UV thermal denaturation experiments. The oligonucleotide concentrations were 10 μ M for both hTeloC and Tel24 DNA in the appropriate buffer, as indicated above. Experiments were performed by following changes of UV signal at 295 nm, at a heating rate of 1 °C/min, in the temperature ranges of 5–100 and 20–100 °C for hTeloC and Tel24, respectively. The melting temperatures (T_m) were mathematically calculated by using the curve fitting function in Origin 7.0 software. T_m values represent the difference between the melting temperature of the DNA with and without ligands.

Nuclear Magnetic Resonance Experiments

A 700 MHz Varian Unity INOVA spectrometer was employed to perform the NMR experiments. One-dimensional proton spectra were recorded at 7 °C using pulsed-field gradient DPGSE for water suppression. All DNA samples were prepared at 0.2 mM strand concentration in 0.22 mL (H₂O/D₂O 9:1) buffer solution. DNA/ligand mixtures were obtained by adding aliquots of a stock solution of the six ligands in DMSO-d₆ directly to the DNA solution inside the NMR tube). NMR data were processed using the iNMRsoftware (www.inmr.net).

FRET studies

A FP-8300 spectrofluorometer (Jasco) equipped with a Peltier temperature controller accessory (Jasco PCT-818) was employed to carry out FRET experiments. The dual-labeled oligonucleotides corresponding to the G4 forming sequence 5'-FAM-d(GGG[TTAGGG]3)-TAMRA-3' (G4-F21T) and the i-motif forming sequence 5'-FAM-d(TAACCC)₄-TAMRA-3' (i-motif-F24T) were used. Such sequences are characterized by the presence of the donor fluorophore FAM (6-carboxyfluorescein) and the acceptor fluorophore TAMRA (6-

carboxytetramethylrhodamine) that are covalently bound at 5'- and 3'-ends, respectively. Labeled oligonucleotides were purchased from Biomers (Germany). G4-F21T and i-motif-F24T were prepared at 1 μ M concentration in 10mM potassium phosphate buffer and 10mM sodium phosphate buffer, respectively. Samples were annealed in a hot water bath at 90 °C for 2 min, and then cooled to room temperature overnight. FRET measurements were performed both in the absence and presence of 5 molar equivalents of compounds. The final concentration of G4-F21T and i-motifF24T was 0.1 μ M. A sealed quartz cuvette with a path length of 1 cm was used. FRET spectra were acquired before (at 5 and 20 °C for i-motif and G4, respectively) and after (at 90 °C) melting assay. The dual-labeled oligonucleotides were excited at 492 nm, and emission spectra were recorded between 500 and 650 nm using 100 nm/s scan speed. Excitation and emission slit widths were both set at 5 nm. FRET-melting experiments were performed by setting the excitation wavelength at 492 nm and the detection wavelength at 522 nm. The emission intensity of FAM was then normalized between 0 and 1. Data analysis was carried out using Origin 7.0 software.

Fluorescent Intercalator Displacement (FID) Assay

For the FID experiments, oligonucleotides d[(TAACCC)₄] (hTeloCFID) and d[(TTTGGG(TTAGGG)₃A)] (Tel24FID) were purchased from Eurogentec and then purified via HPLC. Solid DNA samples were initially dissolved as a 1mM stock solution in MilliQ water. 10mM stock solutions of the candidate ligands were prepared in DMSO. Further dilutions were carried out in buffer: 10 mM NaH₂PO₄ for hTeloCFID and 10 mM KH₂PO₄ for Tel24FID. DNA samples were thermally annealed at 90 μ M in the respective buffers in an Applied Biosystems Veriti 96 well thermal cycler by holding at 95 °C for 5min and cooling at a rate of 1 °C/min to 20 °C. FID experiments were performed on a BMG CLARIO star plate reader using Corning 96-Well Solid Black Flat Bottom plates. A 10 mM stock solution of thiazole orange (TO) was prepared in DMSO and diluted to 2 μ M in the appropriate buffer for either hTeloCFID or Tel24FID. Ninety microliters of the 2 μ M TO solution were added to each well and fluorescence emission at 450 nm measured with excitation at 430 nm; this was normalized to 0% representing background fluorescence. One microliter of DNA was then added, shaken using double orbital shaking at 700 rpm in the plate reader for 15 s, and allowed to equilibrate for 15 min. After equilibration, fluorescence emission was measured as before, and normalized to 100% representing maximal fluorescence enhancement from the TO probe

binding to the DNA secondary structure. 0.9 μ L aliquots of ligand were titrated into each well (in triplicate) and measured as before. Fluorescence measurements after ligand addition were normalized between the 0 and 100% levels determined per the respective well. Percentage TO displacement was calculated as the difference between the normalized 100% fluorescence level and the normalized fluorescence measured after each ligand addition. The concentration for each ligand at which 50% of the TO was displaced (DC_{50}) was calculated by using Origin data analysis software to plot percentage TO displacement against ligand concentration. These data were fitted with dose-response curves and the equations of the curves were solved for $y = 50$ to give the DC_{50} values.

Chapter 6

SYNTHETIZING A NEW THP DERIVATIVE SUITABLE FOR THE CONJUGATION TO NUCLEIC ACID SEQUENCES

In this Chapter, I introduce the research activity that I conducted in the laboratory of Dr. Vincenzo Abbate at King's College London (UK). During this period, I dealt with the synthesis of a Gallium chelator to be used as theranostic.

Positron Emission Tomography (PET) is a method for measuring biochemical and physiological processes *in vivo* in a quantitative way by using radiopharmaceuticals labeled with positron emitting radionuclides such as ^{11}C , ^{13}N , ^{15}O and ^{18}F and by measuring the annihilation radiation using a coincidence technique.

PET employs mainly short-lived positron emitting radiopharmaceuticals. The most widely used radionuclides are: ^{11}C ($t_{1/2} = 20$ min), ^{13}N ($t_{1/2} = 10$ min), ^{15}O ($t_{1/2} = 2$ min) and ^{18}F ($t_{1/2} = 110$ min). Carbon, oxygen, nitrogen and hydrogen are the elements of life and the building blocks of nearly every molecule of biological importance. However, hydrogen has no radioactive isotope decaying with emission of radiation which can be detected outside the human body. For this reason, a fluorine isotope is often used to replace a hydrogen atom in a molecule. Due to these short half-lives, the radionuclides have to be produced in house, preferably with a small, dedicated cyclotron. Since the chemical form of accelerator-produced radionuclides can only be simple, input from organic- and radiochemistry is essential for synthesis of the desired complex molecules. Input from pharmacy is required for the final formulation and pharmacokinetic studies and are also required for (clinical) application. Longer lived positron emitting radionuclides such as ^{52}Fe , ^{55}Co , ^{64}Cu , ^{75}Br , ^{76}Br , ^{124}I are sometimes commercially available or obtainable from research facilities with larger accelerators. Also a few generator systems do exist: the $^{68}\text{Ga}/^{68}\text{Ge}$ and the $^{82}\text{Rb}/^{82}\text{Sr}$ generator.

The $^{68}\text{Ga}/^{68}\text{Ge}$ generator provides a convenient, economic and reliable source of positron-emitting radionuclide, with the advantage of not requiring a local cyclotron. ^{68}Ga has a half-life of 68 min and a high positron yield (90%, 1.9 MeV). To achieve its full potential, incorporation

of this isotope into biomolecules for molecular imaging requires bi-functional chelators able to bind the Ga^{3+} ion rapidly under mild aqueous conditions. The resulting complex has a sufficient kinetic stability *in vivo* to withstand challenge from plasma transferrin, allowing imaging over several hours.

In 2011, in the laboratory of Prof. Robert C. Hider the tripodal tris(hydroxypyridinone) (THP) was synthesized, which is an interesting bifunctional gallium chelator (Figure 6.1). THP incorporation into proteins allowed the easy, fast and high-yield production of ^{68}Ga -labelled derivatives under mild conditions (Berry et al., 2011).

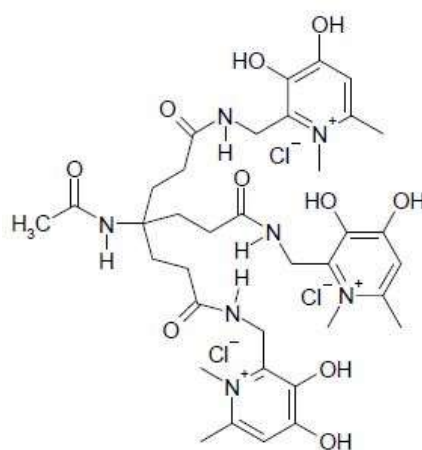


Figure 6.1. Chemical structure of the tripodal tris(hydroxypyridinone).

Later, in the 2017 Young et al. synthesized a conjugate of the THP chelator with the established urea-based prostate-specific membrane antigen (PSMA) inhibitor (Young et al., 2017). Then they radiolabeled this molecule with ^{68}Ga and evaluated its imaging properties. The resulting ^{68}Ga -THP PSMA showed equivalent imaging properties but a greatly simplified radiolabeling procedure compared to other ^{68}Ga -PSMA conjugates, demonstrating that THP offers the prospect of rapid, simple, one-step, room-temperature radiolabeling of ^{68}Ga radiopharmaceuticals (Figure 6.2).

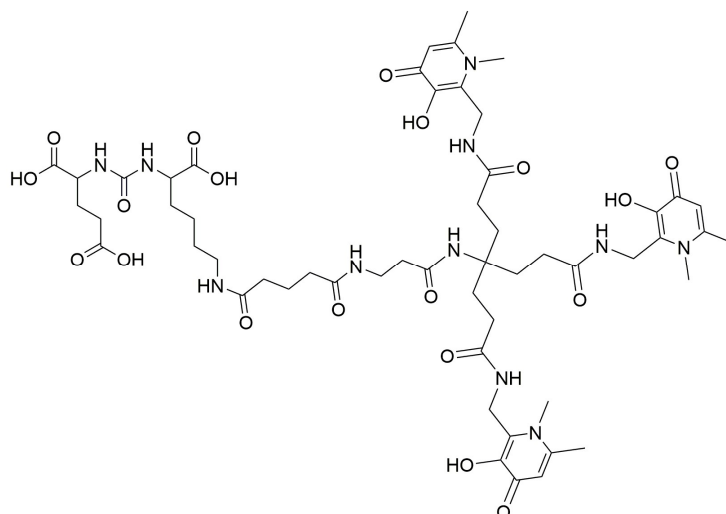


Figure 6.2 Structure of THP-PSMA.

The purpose of my stay at King's College of London was to synthesize a new THP-derivative to be incorporated at the 5'-end of G4-based antiproliferative aptamers.

Actually, THP was designed to be adapted to peptide synthesis. For my purposes, it was necessary to modify some reaction steps in the THP synthetic procedure in order to obtain the suitable THP-derivative for DNA synthesis, reported Figure 6.3.

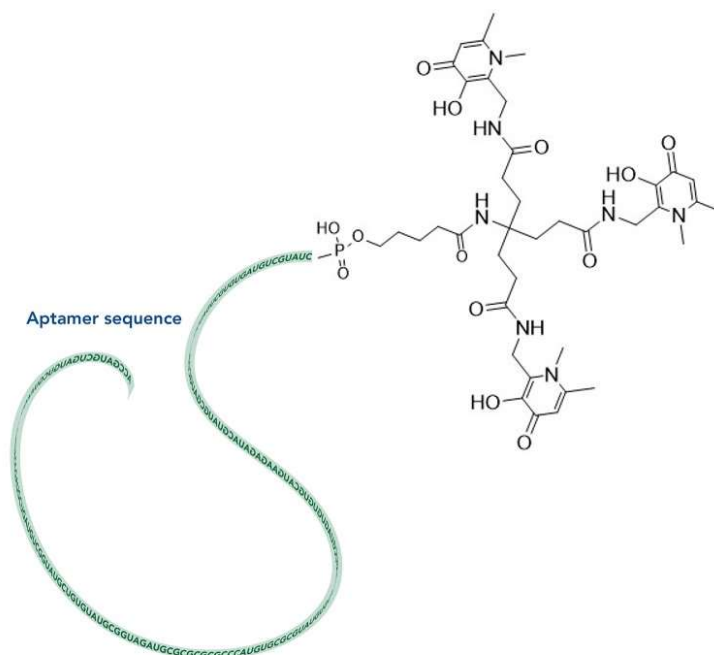


Figure 6.3. Schematic representation of a THP-aptamer.

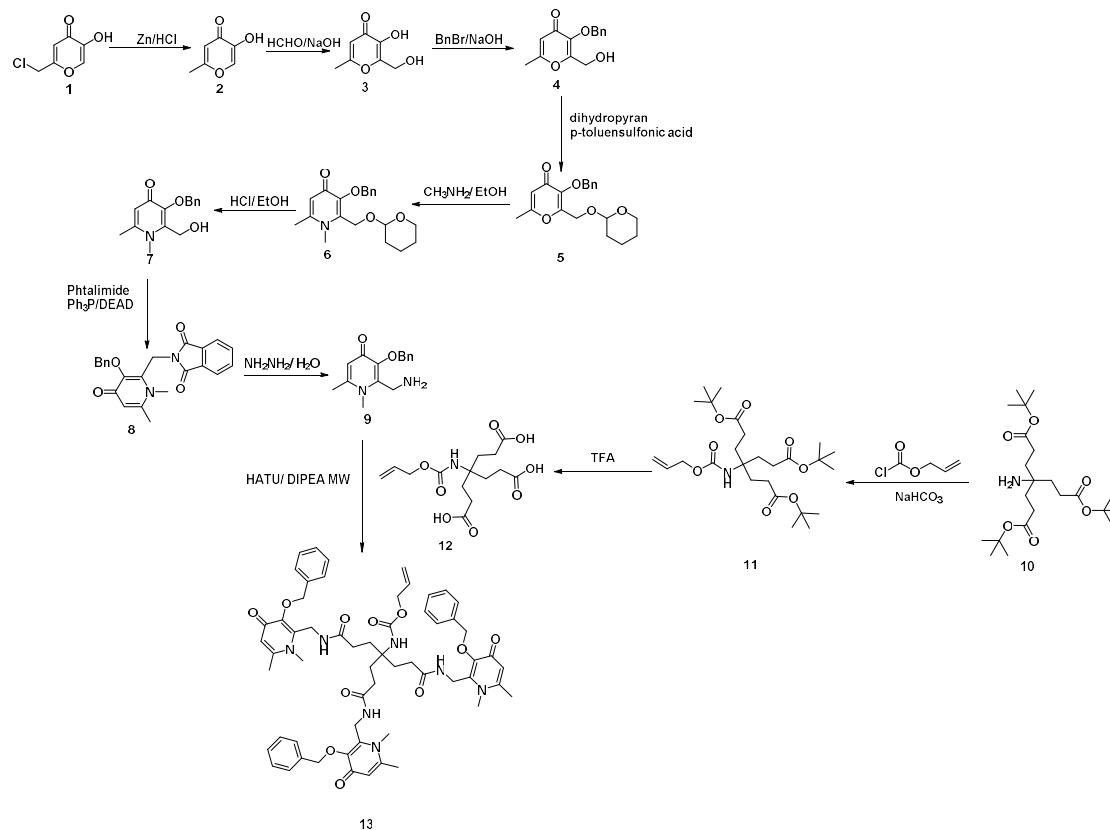
In particular, the incorporation of the THP in the DNA requires the preparation of the corresponding phosphoramidite derivative, and the appropriate protection of the hydroxyl (OH) groups that can interfere during the coupling reaction. As protecting groups of the hydroxyl moiety, we thought to use the acetyl groups due to their stability in the coupling conditions and the ease of deprotection in the same conditions used for the detachment and deprotection of the oligonucleotides from the solid support.

In Scheme 1 is represented the synthetic route for the preparation of THP.

The 3-hydroxy-6-methyl-4*H*-pyran-4-one (**2**) (allomaltol) was synthesized by treating the chlorokojic acid (**1**) with zinc dust and concentrated hydrochloric acid (Ellis et al., 1996). Subsequent condensation of compound **2** with formic aldehyde and sodium hydroxide gave compound **3**. The 3-hydroxyl group was then protected using a benzyl group, thus obtaining compound **4**. The 2-hydroxyl function of (**4**) was protected using 3,4-dihydro-2*H*-pyran before reacting with methylamine. The pyran protecting group was then selectively removed under mild acid conditions without deprotection of the benzyl group, giving the desired protected 2-hydroxymethyl pyridinone (**7**) in a improved yield. The conversion of 1,6-dimethyl-2-hydroxymethyl- 3-benzyloxy pyridin-4(1*H*)-one (**7**) to the corresponding phthalimido intermediate, (**8**), was accomplished using phthalimide, triphenylphosphine, and diethyl azodicarboxylate (DEAD) in good yield. The reaction is believed to proceed through the formation of a quaternary phosphonium salt by addition of triphenylphosphine to DEAD, which reacts in situ with the pyridinone (**7**) to form an alkoxyphosphonium salt (Mitsunobu, 1981). The phosphonium moiety was then displaced by the phthalimido anion to furnish the desired phthalimido derivative (**8**), which was subsequently reacted with hydrazine to afford the amine derivative (**9**).

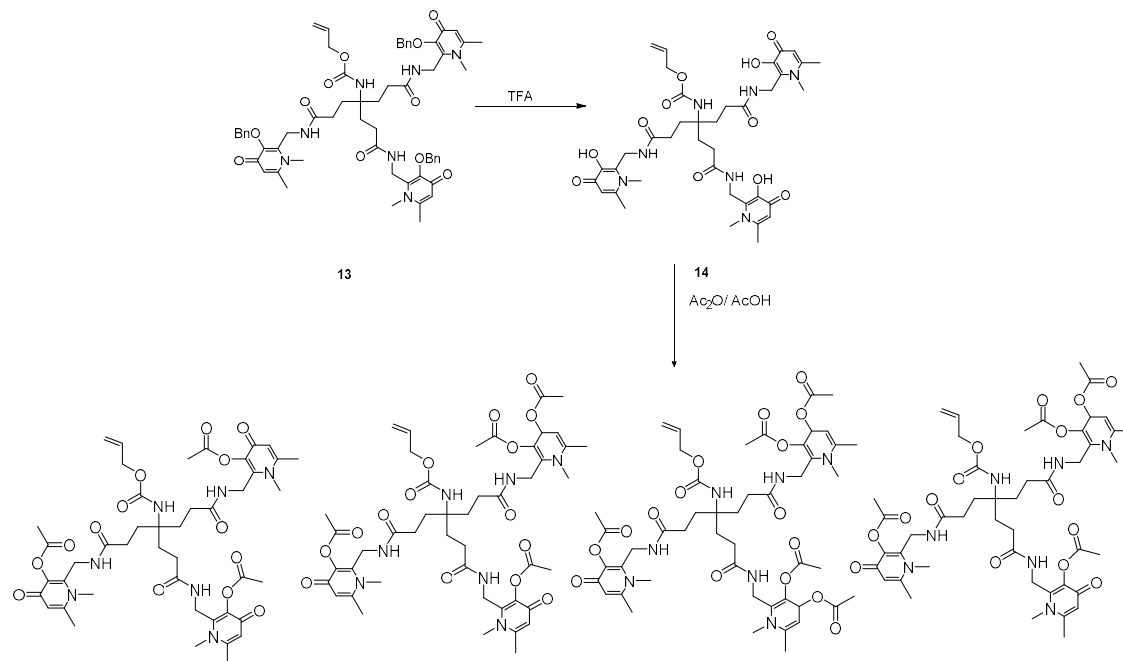
In the meantime, the amino group of di-tert-butyl 4-amino-4-(3-(tert-butoxy)-3-oxopropyl)heptanedioate (**10**) was protected using allyl chloroformate, followed by the hydrolysis of the three tert-butyl ethers in strong acid conditions. The resulting tricarboxylic acid was then coupled to the 1,6-dimethyl-2-aminomethyl-3-benzyloxy pyridin-4(1*H*)-one (**9**) to give the product **13**.

Scheme 1. Synthesis of THP.



Once obtained compound **13**, we proceeded by removing the benzyl ether protective groups (**14**) in strong acidic conditions, and by replacing them with acetate groups (Scheme 2). Unfortunately, the acetylation reaction lead to a mixture of starting material and products with different numbers of acetate groups on them. This mixture was hard to purify, for this reason we decided to change the synthetic strategy.

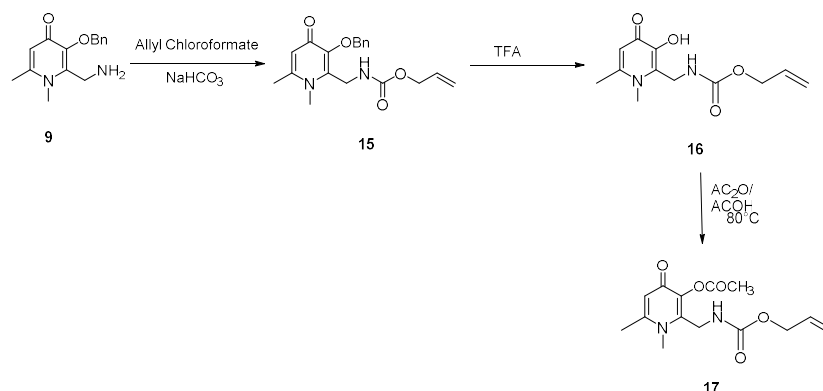
Scheme 2. Synthesis of Ac-THP.



In order to try to overcome this problem we thought to protect the hydroxylic group of the pyridinone ring, used as substrate for preparing the THP, according to the Scheme 3.

The amino group of the 1,6-dimethyl-2-aminomethyl-3-benzoyloxypyridin-4(1H)-one (**9**) was first protected with Alloc-, to avoid its competition in the acetylation reaction. Next, the benzyl-ether protective group was removed in acidic condition, followed by its acetylation in the presence of a mixture of acetic anhydride/acetic acid at 80 °C. The obtained 2-(((allyloxy)carbonyl)amino)methyl)-1,6-dimethyl-4-oxo-1,4-dihydropyridin-3-yl acetate compound (**17**), resulted to be unstable and easily degradable.

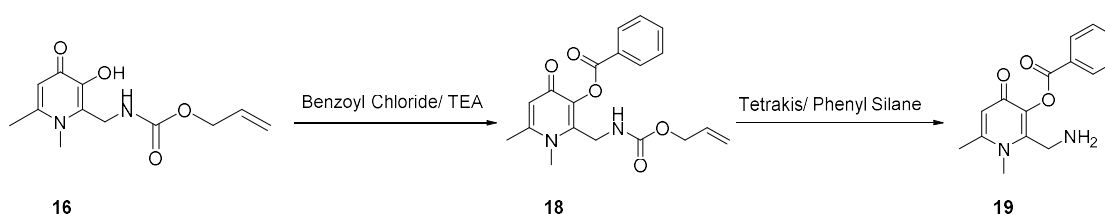
Scheme 3. Synthesis of 2-(((allyloxy)carbonyl)amino)methyl)-1,6-dimethyl-4-oxo-1,4-dihydropyridin-3-yl acetate.



In order to try to obtain a more stable product with an appropriate protective group useful in the oligonucleotide synthesis we decided to mask the hydroxylic function as a benzyl-ester derivative (Scheme 4).

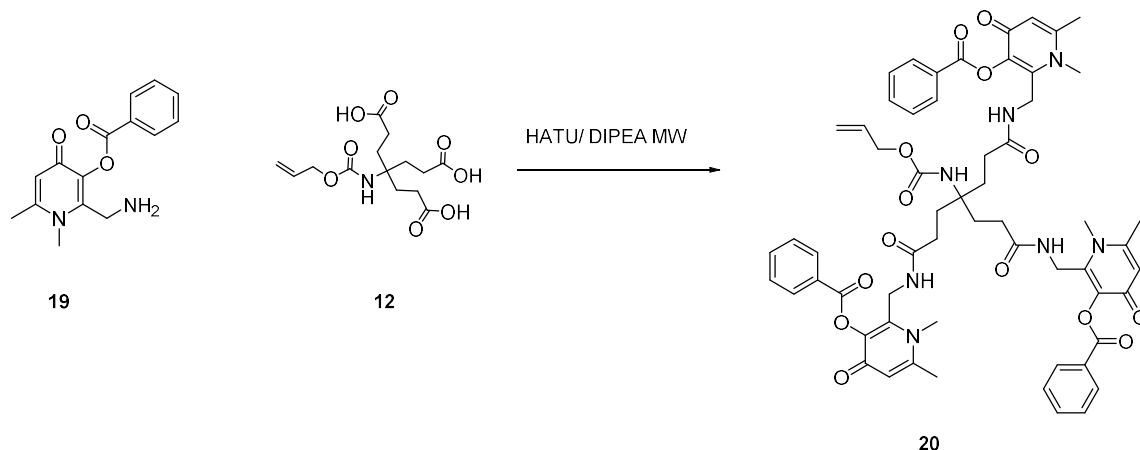
Thus, the allyl ((3-hydroxy-1,6-dimethyl-4-oxo-1,4-dihydropyridin-2-yl)methyl)carbamate (**16**) was treated with benzoyl chloride to give the 2-(((allyloxy)carbonyl)amino)methyl)-1,6-dimethyl-4-oxo-1,4-dihydropyridin-3-yl benzoate (**18**) in a good yield. Next, the Alloc-protective group was removed in presence of Tetrakis and Phenyl Silane to obtain the stable 2-(aminomethyl)-1,6-dimethyl-4-oxo-1,4-dihydropyridin-3-yl benzoate (**19**).

Scheme 4. Synthesis of the 2-(aminomethyl)-1,6-dimethyl-4-oxo-1,4-dihydropyridin-3-yl benzoate.



Compound **19** was used for the conjugation with the 4-(((allyloxy)carbonyl)amino)-4-(2-carboxyethyl)heptanedioic acid (**12**), according to Scheme 5. This reaction was realized in the presence of HATU and DIPEA, under microwave conditions.

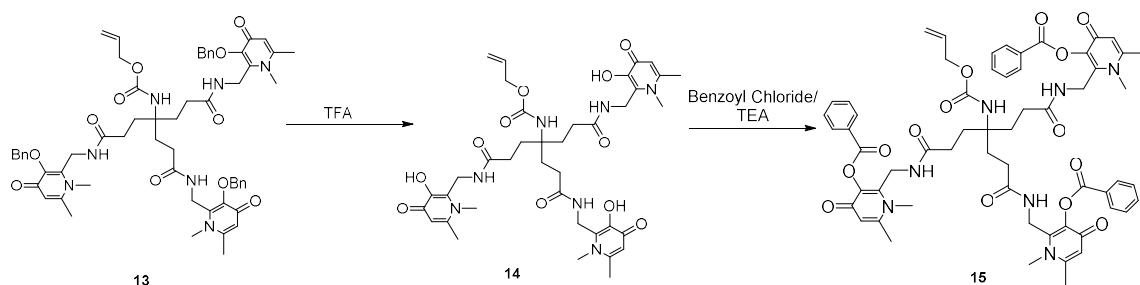
Scheme 5. Synthesis of Benzoyl-THP.



Unfortunately, even if the reaction protocol was the same followed in Scheme 1, we didn't manage to get the desired product **20**, probably due to the interference of the benzyl ester groups on the pyridinone ring.

We are now continuing to collaborate with Dr. Vincenzo Abbate and Dr. Agostino Cilibrizzi to find an alternative synthetic route for obtaining the desired modified-THP to be conjugated to oligonucleotide sequences. In particular, the synthetic route reported in Scheme 6 was designed. Basically, it is based on the removal of benzyl-ether groups from compound **13** in strong acidic conditions, followed by benzoylation of the hydroxyl groups.

Scheme 6. Synthesis of Benzoyl-THP.



Experimental Section

Intermediates **2-9** were synthesized by previously described literature methods (Berry et al., 2014).

2-Methyl-5-hydroxypyran-4(1H)-one (allomaltol) (2). To a three-necked round-bottomed flask (250 mL) equipped with a thermometer, pressure-equalized dropping funnel, condenser, and magnetic stirring bar chlorokojic acid (12.7 g, 79 mmol) and water (100 mL) were added. The mixture was stirred and heated to a temperature of 40 °C in an oil bath. Zinc dust (10.37 g, 160 mmol, 2 mol equiv) was added, and the reaction mixture was stirred vigorously at 60 °C. Concentrated hydrochloric acid (23.5 mL, 3 mol equiv.) was added dropwise via the addition funnel over a period of about 1 h during which the reaction mixture effervesced producing hydrogen. The slurry was left stirring for 2-3 h at 60 °C. The excess zinc was removed from the pale green reaction mixture by hot filtration. The filtrate was adjusted to pH 1 and extracted with dichloromethane (3 x 150 mL). The organic extracts were combined, dried over anhydrous sodium sulfate, filtered, and concentrated. Recrystallization from 2-propanol gave **2** as a white solid (10.4 g).

2-Hydroxymethyl-3-hydroxy-6-methyl-pyran-4(1H)-one (3). Allomaltol **2** (10.4 g, 82.7 mmol) was added to an aqueous solution of sodium hydroxide (3.6 g, 91 mmol) in distilled water (110 mL) and stirred at room temperature for 5 min. Formaldehyde solution (35%, 6.7 mL, 1 mol equivalent) was added dropwise over 10 min, and the solution was stirred for 12 h. Acidification to pH 1 using concentrated hydrochloric acid at 3-5 °C for 12 h gave a crystalline deposit (8.2 g).

2-Hydroxymethyl-3-benzyloxy-6-methyl-pyran-4(1H)-one (4). Sodium hydroxide (2.29 g, 57.4 mmol, 1.1 mol equiv) dissolved in 5 mL of distilled water was added to a solution of **3** (8.2 g, 52.1 mmol, 1 mol equiv) in methanol (50 mL), and the reaction mixture was heated to reflux. Benzyl bromide (6.8 mL, 57.4 mmol, 1.1 mol equivalent) was added dropwise over 30 min and the resulting mixture was refluxed for 12 h. The reaction mixture was concentrated in vacuo, the residue taken up into dichloromethane (150 mL), and the inorganic salts filtered off. The dichloromethane layer was washed with 5% sodium hydroxide solution (2 x 50 mL) and water (50 mL), dried over anhydrous sodium sulfate, filtered, and concentrated in vacuo to yield the crude product as a yellow crystalline solid. Recrystallization from dichloromethane/petroleum ether 40/60 afforded a white crystalline solid (8.3 g).

1,6-Dimethyl-2-hydroxymethyl-3-benzyloxy pyridin-4(1H)-one (7). 3,4-dihydro-2*H*-pyran (3.5 g, 40 mmol, 4 mol equiv) was added to a solution of **9** (2.5 g, 10 mmol, 1 mol equiv) in dichloromethane (30 mL), followed by the addition of *p*-toluenesulfonic acid monohydrate (60 mg, cat.). After being stirred at room temperature for 3 h, the reaction mixture was washed with 5% aqueous sodium carbonate (20 mL) followed by water (2 x 20 mL). The organic fraction was then dried over anhydrous sodium sulfate, filtered, and rotary evaporated to yield a light yellow oil, which was then dissolved in ethanol (10 mL)/40% aqueous methylamine (10 mL). The reaction mixture was sealed in a thick-walled glass tube and stirred at 70 °C for 12 h. After removal of the solvent by rotary evaporation, the residue was redissolved in ethanol (20 mL) and 2 N hydrochloric acid (5 mL) and refluxed for 4 h. The solvent was removed by rotary evaporation prior to addition of water (50 mL) and diethyl ether (2 x 20 mL). Subsequent adjustment of the aqueous fraction to pH 9 with 10 N sodium hydroxide solution was followed by extraction into dichloromethane (4 x 50 mL). The combined organic layers were dried over anhydrous sodium sulfate, filtered, and rotary evaporated to give a light brown solid. Recrystallization from methanol/diethyl ether afforded the pure product (1.8 g) as a white crystalline solid.

1,6-Dimethyl-2-phthalimidomethyl-3-benzyloxy pyridin-4(1H)-one (8). 1.8 g of **7** (6.8 mmol, 1 mol equiv) was added to a solution of triphenyl phosphine (2.1 g, 8.1 mmol, 1.2 mol equiv) and phthalimide (1.2 g, 8.1 mmol, 1.2 mol equiv) in dry THF (20 mL). The mixture was cooled to 0 °C in an ice bath. Diethyl azodicarboxylate (1.3 mL, 8.1 mmol, 1.2 mol equiv) was added dropwise with stirring over 30 min, after which the reaction mixture was allowed to warm slowly to room temperature and then stirred overnight. The resulting precipitate was isolated by filtration, washed with THF (10 mL), and dried under high vacuum to yield a white amorphous powder (2.1 g).

1,6-Dimethyl-2-aminomethyl-3-benzyloxy pyridin-4(1H)-one (9). 6 mL of 5.5% aqueous hydrazine was added to a solution of **8** (3.4 g, 8.7 mmol, 1 mol equiv) in ethanol (28 mL). After being refluxed for 3 h, the reaction mixture was chilled to 0 °C, acidified to pH 1 with concentrated hydrochloric acid, and filtered. The filtrate was concentrated in vacuo, and the residue was dissolved in distilled water (28 mL), adjusted to pH 12 with 10 N sodium hydroxide, and extracted with dichloromethane (3 x 60 mL). The combined organic extracts

were dried over anhydrous sodium sulfate, and the solvent was removed under reduced pressure to yield an orange oil. Purification by column chromatography on silica gel (eluant: methanol:chloroform, 20:80 v/v) furnished a white solid (1.7 g).

di-tert-butyl-4-(((allyloxy)carbonyl)amino)-4-(3-(tert-butoxy)-3-oxopropyl)heptanedioate (11).

Sodium bicarbonate (0.1 g, 2 mol equiv) was added to a solution of **10** (0.25 g, 0.6 mmol, 1 mol equiv) in 50% aqueous dioxane (10 mL) at 0 °C followed by dropwise addition of 0.13 mL of allylchloroformate (2 mol equiv). The reaction mixture was left for 2 h at 0 °C, then at room temperature. After 12 h the reaction mixture was acidified to pH 1 with concentrated hydrochloric acid and filtered. The water solution was extracted with dichloromethane (3 x 50 mL). The combined organic extracts were dried over anhydrous sodium sulfate, and the solvent was removed under reduced pressure to yield a white oil. Purification by column chromatography on silica gel (eluant: methanol:dichloromethane 20:80 v/v) furnished a white solid (0.2 g) ¹H NMR (400 MHz, CDCl₃) δ 5.90 (m, 1H), 5.28 (dd, *J* = 17.2, 1.5 Hz, 1H), 5.20 (dd, *J* = 10.4, 1.3 Hz, 1H), 4.79 (s, 1H), 4.50 (d, *J* = 4.3 Hz, 2H), 2.22 (dd, *J* = 9.1, 6.9 Hz, 6H), 1.91 (m, 6H), 1.44 (s, 27H).

4-(((allyloxy)carbonyl)amino)-4-(2-carboxyethyl)heptanedioic acid (12). 2.0 mL of 100% TFA were added to 0.1 g of **11**. After 24 h the solution was concentrated by nitrogen flush and the residue was crushed with diethyl ether to give a white solid (0.051 g) ¹H NMR (400 MHz, DMSO) δ 6.87 (s, 3H), 5.90 (ddd, *J* = 22.4, 10.5, 5.2 Hz, 1H), 5.28 (m, 1 H), 5.16 (dd, *J* = 10.5, 1.7 Hz, 1H), 4.43 (d, *J* = 5.2 Hz, 2H), 2.12 (m, 6H), 1.80 (dd, *J* = 20.9, 12.4 Hz, 6H).

Allyl-(1,7-bis(((3-(benzyloxy)-1,6-dimethyl-4-oxo-1,4-dihydropyridin-2-yl)methyl)amino)-4-(3-(((3-(benzyloxy)-1,6-dimethyl-4-oxo-1,4-dihydropyridin-2-yl)methyl)amino)-3-oxopropyl)-1,7-dioxoheptan-4-yl)carbamate (13). 0.05 g of **12** (0.15 mmol, 1 mol equiv), 0.228 g of HATU (0.6 mmol, 4 mol equiv), 0.07 g of DIPEA (0.6 mmol, 4 mol equiv), and 0.155 g of **9** (0.6 mmol, 4 mol equiv.) were dissolved in 10 mL anhydrous DMF and reacted in a 100 mL microwave tube. The mixture was stirred for 30 min at 80 °C, 200 W in microwave. DMF was removed in vacuo. The residue was dissolved in methanol and purified by column chromatography on silica gel (eluant: methanol:dichloromethane 20:80

v/v). The fractions were combined, dried (on Na₂SO₄) and the solvent was removed in vacuo to give 0.315 g of a white solid in 50% yield.

Allyl-(1,7-bis(((3-hydroxy-1,6-dimethyl-4-oxo-1,4-dihydropyridin-2-yl)methyl)amino)-4-(3-(((3-hydroxy-1,6-dimethyl-4-oxo-1,4-dihydropyridin-2-yl)methyl)amino)-3-oxopropyl)-1,7-dioxoheptan-4-yl)carbamate (14). 0.1 g of **13** (0.09 mmol, 1 mol equiv) was added to 5 ml of TFA (100%) and stirred for 48 hours. After this time the TFA was concentrated by nitrogen flush and the residue was crushed with diethyl ether to give a white solid (0.07 g).

Allyl-((3-(benzyloxy)-1,6-dimethyl-4-oxo-1,4-dihydropyridin-2-yl)methyl)carbamate (15). Sodium bicarbonate was added to a solution of piridinone **9** (1.72 g, 6.7 mmol, 1 mol equiv) in 50% aqueous dioxane (50 mL) at 0 °C (1.1 g, 13.4 mmol, 2 mol equiv) followed by the dropwise addition of 1.45 mL of allylchloroformate (13.4 mmol, 2 mol equiv). The reaction mixture was left for 2 h at 0 °C then at room temperature. After 12 h the reaction mixture was acidified to pH 1 with concentrated hydrochloric acid and filtered. The water solution was extracted with dichloromethane (3 x 1000 mL). The combined organic extracts were dried over anhydrous sodium sulfate, and the solvent was removed under reduced pressure to yield a white oil. Purification by column chromatography on silica gel (eluant: 2-propanol:dichloromethane 20:80 v/v) furnished a white solid (1.86 g).

Allyl-((3-hydroxy-1,6-dimethyl-4-oxo-1,4-dihydropyridin-2-yl)methyl)carbamate (16). 1.86 g of **15** (5.4 mmol, 1 mol equiv) were added to 20 mL of TFA (100%) and stirred for 48 h. After this time the TFA was concentrated by nitrogen flush and the residue was crushed with diethyl ether to give a white solid (1.3 g). ¹H NMR (400 MHz, DMSO) δ 7.77 (s, 1H), 6.97 (s, 1H), 5.88 (m, 1H), 5.26 (d, *J* = 17.3 Hz, 2H), 5.17 (dd, *J* = 10.5, 1.5 Hz, 15H), 4.53 (s, 2H), 4.48 (d, *J* = 5.3 Hz, 2H), 3.81 (s, 3H) (m/z = 252.96).

2-(((allyloxy)carbonyl)amino)methyl)-1,6-dimethyl-4-oxo-1,4-dihydropyridin-3-yl acetate (17). 0.05 g of **16** (0.2 mmol, 1 mol equiv) were reacted with acetic anhydride (1.8 mL, 19.8 mmol, 99 mol equiv) and glacial acetic acid (0.380 mL, 6.6 mmol, 33 mol equiv). The reaction was held at 80 °C for 18 h under nitrogen. The solution was concentrated and the resulting residue was dissolved in dichloromethane (20 mL) and washed with brine (220 mL).

The organic layer was collected, dried over sodium sulfate, and filtered to obtain (0.04 g) of white oil. ¹H NMR (400 MHz, CDCl₃) δ 6.91 (s, 1H), 5.90 (m, 1H), 5.31 (m, 1H), 5.22 (d, *J* = 10.4 Hz, 1H), 4.58 (d, *J* = 5.6 Hz, 2H), 4.50 (d, *J* = 5.0 Hz, 2H), 3.80 (s, 3H), 2.46 (s, 3H), 2.37 (s, 3H) (*m/z* = 294.91)

2-(((allyloxy)carbonyl)amino)methyl)-1,6-dimethyl-4-oxo-1,4-dihydropyridin-3-yl benzoate (18). To a cooled (0 °C) suspension of allyl-((3-hydroxy-1,6-dimethyl-4-oxo-1,4-dihydropyridin-2-yl)methyl)carbamate (0.4 g, 1.6 mmol, 1 mol equiv) and a catalytic amount of Et₃N (0.011 mL, 0.05 mol equiv) in anhydrous dichloromethane (10 mL), benzoyl chloride (0.232 mL, 2 mmol, 1.25 mol equiv) was added. The mixture was stirred for 1-2 h at 0 °C and then for 1-3 h at room temperature. The precipitate was filtered off, and the solvent was evaporated in vacuo. Cold water (20 mL) was added, the mixture was neutralized with 0.250 N sodium hydroxide, and the precipitate was recovered by vacuum filtration and extracted with dichloromethane (3 x 15 mL). The solvent was dried over sodium sulphate, and then the compound was purified by column chromatography using methanol/dichloromethane 8:2 as eluent to give a white solid (0.05 g). ¹H NMR (400 MHz, CDCl₃) δ 7.50 (dd, *J* = 13.8, 7.3 Hz, H), 7.37 (t, *J* = 7.6 Hz, 2H), 7.03 (s, 1H), 5.43 (m, 1H), 5.25 (s, 2H), 5.04 (d, *J* = 10.4 Hz, 1H), 4.94 (d, *J* = 17.2 Hz, 1H), 4.42 (d, *J* = 5.8 Hz, 2H), 4.06 (s, 3H), 2.55 (s, 3H) (*m/z* = 357.35).

2-(aminomethyl)-1,6-dimethyl-4-oxo-1,4-dihydropyridin-3-yl benzoate (19). To a cooled solution (0 °C) of **18** (0.150 g, 0.42 mmol, 1 mol equiv), 0.004 g of Tetrakis (0.01 mol equiv), and 0.310 mL of phenylsilane (0.252 mmol, 6 mol equiv) were added. The reaction mixture was left on ice bath for 20 min, then it was concentrated in vacuum and purified by column chromatography on silica gel (eluant: methanol:dichloromethane 30:70 v/v) to give a red solid (0.09 g) (*m/z* 273.09).

CONCLUSIONS

Because of their involvement in the regulation of numerous key cellular processes, non-canonical DNA structures are now widely accepted as potential pharmacological targets, and the development of compounds that can selectively interact with these structures is becoming increasingly important. In this context, there are numerous ways that can be explored including *de novo* synthesis of new compounds, or the improvement of existing molecules.

In this thesis, I have dealt with different aspects of this topic. New derivatives were synthesized starting from a previously identified G4 ligand by modifying some chains and substituents, leading to the discovery of a new molecule with high selectivity for the *MYC* oncogene promoter G4.

Furthermore, a one pot reaction was used to synthesize a new class of furopiridazinone compounds. Among these molecules, one has shown to interact with the *BCL2* oncogene promoter G4, inhibiting the overexpression of the anti-apoptotic BCL2 protein.

Moreover, using a known G4 ligand as hit compound, I explored the effects of a single diversity element inserted on a common planar scaffold, which *per se* acts as a suitable structural motif for G4 recognition. Using affinity chromatography-based G4-CPG binding assays coupled with biological and biophysical experiments, I managed to select a promising ligand from a small library of putative G4 binders.

A virtual screening approach combined with experimental investigations led to the discovery of the first G3/G4 interacting compound, thus stimulating further studies aimed at the development of drugs capable of targeting such noncanonical structures.

Biophysical experiments have also highlighted the ability of some well-known G4 binders to interact with i-motif structure, demonstrating the necessity, when studying G4-targeting compounds, of evaluating also their effects on the i-motif counterparts, especially if one is looking for a “specific” drug.

In conclusion, in this thesis, several aspects of G4, G3, and i-motif structures were explored, as well as their molecular interactions with compounds that could represent the starting point for the development of useful drug candidates.

REFERENCES

- Adams, J. M. (1998). The Bcl-2 Protein Family: Arbiters of Cell Survival. *Science (80-.)*. 281, 1322–1326.
- Amato, J., Iaccarino, N., Pagano, B., Morigi, R., Locatelli, A., Leoni, A., et al. (2014a). Bis-indole derivatives with antitumor activity turn out to be specific ligands of human telomeric G-quadruplex. *Front. Chem.* 2, 1–8.
- Amato, J., Iaccarino, N., Randazzo, A., Novellino, E., and Pagano, B. (2014b). Noncanonical DNA secondary structures as drug targets: The prospect of the i-motif. *ChemMedChem* 9, 2026–2030.
- Amato, J., Morigi, R., Pagano, B., Pagano, A., Ohnmacht, S., De Magis, A., et al. (2016). Toward the Development of Specific G-Quadruplex Binders: Synthesis, Biophysical, and Biological Studies of New Hydrazone Derivatives. *J. Med. Chem.* 59, 5706–5720.
- Amato, J., Pagano, A., Capasso, D., Di Gaetano, S., Giustiniano, M., Novellino, E., et al. (2018). Targeting the BCL2 Gene Promoter G-Quadruplex with a New Class of Fuopyridazinone-Based Molecules. *ChemMedChem.* 13, 406–410.
- Amato, J., Pagano, A., Cosconati, S., Amendola, G., Fotticchia, I., Iaccarino, N., et al. (2017). Discovery of the first dual G-triplex/G-quadruplex stabilizing compound: a new opportunity in the targeting of G-rich DNA structures? *Biochim. Biophys. Acta - Gen. Subj.* 1861, 1271–1280.
- Amato, J., Platella, C., Iachettini, S., Zizza, P., Musumeci, D., Cosconati, S., et al. (2019). Tailoring a lead-like compound targeting multiple G-quadruplex structures. *Eur. J. Med. Chem.* 163, 295–306.
- Amrus, A., Chen, D., Dai, J., Jones, R. A., and Yang, D. (2005). Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization. *Biochemistry* 44, 2048–2058.
- Arbel, N., Ben-Hail, D., and Shoshan-Barmatz, V. (2012). Mediation of the antiapoptotic activity of Bcl-xL protein upon interaction with VDAC1 protein. *J. Biol. Chem.* 287, 23152–23161.
- Armond, R. De, Wood, S., Sun, D., Hurley, L. H., and Ebbinghaus, S. W. (2005). Evidence for the Presence of a Guanine Quadruplex Forming Region within a Polypurine Tract of the Hypoxia Inducible Factor 1 R Promoter. *Biochemistry* 44, 16341–16350.
- Asif, M. (2012). Some Recent Approaches of Biologically Active Substituted Pyridazine and Phthalazine Drugs. *Curr. Med. Chem.* 19, 2984–2991.
- Baaske, P., Wienken, C. J., Reineck, P., Duhr, S., and Braun, D. (2010). Optical thermophoresis for quantifying the buffer dependence of aptamer binding. *Angew. Chemie - Int. Ed.* 49, 2238–2241.
- Bacolla, A., and Wells, R. D. (2009). Non-B DNA Conformations as Determinants of Mutagenesis and Human Disease. *Mol. Carcinog.* 285, 273–285.
- Balasubramanian, S., Hurley, L. H., and Neidle, S. (2011). Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discov.* 10, 261–275.
- Bar-Am, O., Weinreb, O., Amit, T., and Youdim, M. B. H. (2005). Regulation of Bcl-2 family proteins, neurotrophic factors, and APP processing in the neurorescue activity of propargylamine. *FASEB J.* 19, 1899–1901.
- Bates, P. J., Kahlon, J. B., Thomas, S. D., Trent, J. O., and Miller, D. M. (1999). Antiproliferative activity of G-rich oligonucleotides correlates with protein binding. *J. Biol. Chem.* 274, 26369–26377.
- Bates, P. J., Laber, D. A., Miller, D. M., Thomas, S. D., and Trent, J. O. (2009). Discovery and development of the G-rich oligonucleotide AS1411 as a novel treatment for cancer. *Exp. Mol. Pathol.* 86, 151–164.
- Baudino, T. (2015). Targeted Cancer Therapy: The Next Generation of Cancer Treatment. *Curr. Drug*

- Discov. Technol.* 12, 3–20.
- Bejugam, M., Sewitz, S., Shirude, P. S., Shahid, R., Balasubramanian, S., Uni, T., et al. (2007). Trisubstituted Isoalloxazines as a New Class of G-Quadruplex Binding Ligands : Small Molecule Regulation of c-kit Oncogene Expression. 12926–12927.
- Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81, 3684–3690.
- Berger, C. M., Gaume, X., and Bouvet, P. (2015). The roles of nucleolin subcellular localization in cancer. *Biochimie* 113, 78–85.
- Berry, D. J., Ma, Y., Ballinger, J. R., Tavaré, R., and Koers, A. (2014). Efficient bifunctional gallium-68 chelators for positron emission tomography : tris (hydroxypyridinone) ligands. *Chem. Commun. (Camb)*. 47, 7068–7070.
- Berry, D. J., Ma, Y., Ballinger, J. R., Tavaré, R., Koers, A., Sunassee, K., et al. (2011). Efficient bifunctional gallium-68 chelators for positron emission tomography: Tris(hydroxypyridinone) ligands. *Chem. Commun.* 47, 7068–7070.
- Biffi, G., Tannahill, D., McCafferty, J., and Balasubramanian, S. (2013). Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* 5, 182–186.
- Bochman, M. L., Paeschke, K., and Zakian, V. A. (2012). REVIEWS DNA secondary structures : stability and function of G - quadruplex structures. *Nat. Rev. Genet.* 13, 770–780.
- Bock, L. C., Griffin, L. C., Latham, J. a, Vermaas, E. H., and Toole, J. J. (1992). Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature* 355, 564–566.
- Bodnar, A. G., Ouellette, M., Frolkis, M., Holt, S. E., Chiu, C., Morin, G. B., et al. (1998). Extension of Life-Span by Introduction of Telomerase into Normal Human Cells. *Science (80-)*. 279, 349–352.
- Boger, D. L., Fink, B. E., and Hedrick, M. P. (2000). Total Synthesis of Distamycin A and 2640 Analogues: A Solution-Phase Combinatorial Approach to the Discovery of New, Bioactive DNA Binding Agents and Development of a Rapid, High-Throughput Screen for Determining Relative DNA Binding Affinity or DNA Bind. *J. Am. Chem. Soc.* 122, 6382–6394.
- Brooks, T. A., Kendrick, S., and Hurley, L. (2010). Making sense of G-quadruplex and i-motif functions in oncogene promoters. *FEBS J.* 277, 3459–3469.
- Brown, R. V., Wang, T., Chappeta, V. R., Wu, G., Onel, B., Chawla, R., et al. (2017). The Consequences of Overlapping G-Quadruplexes and i-Motifs in the Platelet-Derived Growth Factor Receptor β Core Promoter Nuclease Hypersensitive Element Can Explain the Unexpected Effects of Mutations and Provide Opportunities for Selective Targeting. *J. Am. Chem. Soc.* 139, 7456–7475.
- Bugaut, A., and Balasubramanian, S. (2008). A sequence-independent study of the influence of short loop lengths on the stability and topology of intramolecular DNA G-quadruplexes. *Biochemistry* 47, 689–697.
- Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K., and Neidle, S. (2006). Quadruplex DNA: Sequence, topology and structure. *Nucleic Acids Res.* 34, 5402–5415.
- Cantor, C. R., Warshaw, M. M., and Shapiro, H. (1970). Oligonucleotide interactions. III. Circular dichroism studies of the conformation of deoxyoligonucleolides. *Biopolymers* 9, 1059–1077.
- Capasso, D., Gaetano, D., Celentano, V., Diana, D., Festa, L., Stasi, D., et al. (2017). Molecular BioSystems the IQGAP1 protein. *Mol. BioSyst.* 13, 1619–1629.
- Capasso, D., Paola, I. De, Liguoro, A., Gatto, A. Del, Gaetano, S. Di, Guarnieri, D., et al. (2014). RGDechi-hCit : a v b 3 Selective Pro-Apoptotic Peptide as Potential Carrier for Drug Delivery into Melanoma Metastatic Cells. *PLoS One* 9, e106441.
- Cerofolini, L., Amato, J., Giachetti, A., Limongelli, V., Novellino, E., Parrinello, M., et al. (2014). G-triplex structure and formation propensity. *Nucleic Acids Res.* 42, 13393–13404.
- Chaires, J. B., Trent, J. O., Gray, R. D., Dean, W. L., Buscaglia, R., Thomas, S. D., et al. (2014). An improved model for the hTERT promoter quadruplex. *PLoS One* 9, 1–13.
- Chambers, V. S., Marsico, G., Boutell, J. M., Di Antonio, M., Smith, G. P., and Balasubramanian, S. (2015). High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.* 33, 877–881.
- Cheatham, T. E., Cieplak, P., and Kollman, P. A. (1999). A modified version of the cornell et al. force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn.* 16, 845–862.

- Chen, Y., Qu, K., Zhao, C., Wu, L., Ren, J., Wang, J., et al. (2012). Insights into the biomedical effects of carboxylated single-wall carbon nanotubes on telomerase and telomeres. *Nat. Commun.* 3.
- Cheng, M., Modi, C., Cookson, J. C., Hutchinson, I., Heald, R. A., Mccarroll, A. J., et al. (2008). Antitumor Polycyclic Acridines. 20. 1 Search for DNA Quadruplex Binding Selectivity in a Series of 8, 13-Dimethylquino [4, 3, 2-kl] acridinium Salts: Telomere- Targeted Agents. *J. Med. Chem.* 51, 963–975.
- Cogoi, S., Quadrifoglio, F., and Xodo, L. E. (2004). G-rich Oligonucleotide Inhibits the Binding of a Nuclear Protein to the Ki- ras Promoter and Strongly Reduces Cell Growth in Human Carcinoma Pancreatic Cells †. *Biochemistry* 21, 2512–2523.
- Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., et al. (1995). A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* 117, 5179–5197.
- Cosconati, S., Forli, S., Perryman, A. L., Harris, R., Goodsell, D. S., and Olson, A. J. (2010). Virtual screening with AutoDock: theory and practice. *Expert Opin. Drug Discov.* 5, 597–607.
- Cosconati, S., Marinelli, L., Trotta, R., Virno, A., Mayol, L., Novellino, E., et al. (2009). 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.
- Cosconati, S., Rizzo, A., Trotta, R., Pagano, B., Iachettini, S., De Tito, S., et al. (2012). Shooting for selective druglike G-quadruplex binders: Evidence for telomeric DNA damage and tumor cell death. *J. Med. Chem.* 55, 9785–9792.
- d’Adda di Fagagna, F., Reaper, P. M., Clay-Farrace, L., Fiegler, H., Carr, P., von Zglinicki, T., et al. (2003). A DNA damage checkpoint response in telomere-associated senescence. *Nature* 426, 194–198.
- Dai, J., Carver, M., Hurley, L. H., and Yang, D. (2011). 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.
- Dai, J., Carver, M., Punchihewa, C., Jones, R. A., and Yang, D. (2007). 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.
- Dai, J., Carver, M., and Yang, D. (2008). Polymorphism of human telomeric quadruplex structures. *Biochimie* 90, 1172–1183.
- Dai, J., Chen, D., Jones, R. A., Hurley, L. H., and Yang, D. (2006). NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. 34, 0–11.
- Day, H. A., Pavlou, P., and Waller, Z. A. E. (2014). I-Motif DNA: Structure, stability and targeting with ligands. *Bioorganic Med. Chem.* 22, 4407–4418.
- De Cian, A., DeLemos, E., Mergny, J. L., Teulade-Fichou, M. P., and Monchaud, D. (2007a). Highly efficient G-quadruplex recognition by bisquinolinium compounds. *J. Am. Chem. Soc.* 129, 1856–1857.
- De Cian, A., Guittat, L., Kaiser, M., Saccà, B., Amrane, S., Bourdoncle, A., et al. (2007b). Fluorescence-based melting assays for studying quadruplex ligands. *Methods* 42, 183–195.
- Delgado, M. D., Albajar, M., Gomez-Casares, M. T., Batlle, A., and León, J. (2013). MYC oncogene in myeloid neoplasias. *Clin. Transl. Oncol.* 15, 87–94.
- Dexheimer, T. S., Sun, D., and Hurley, L. H. (2006). 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.
- Di Gaetano, S., D’Alessio, G., and Piccoli, R. (2001). Second generation antitumour human RNase: significance of its structural and functional features for the mechanism of antitumour action. *Biochem. J.* 358, 241–247.
- Di Leva, F. S., Zizza, P., Cingolani, C., D’Angelo, C., Pagano, B., Amato, J., et al. (2013). Exploring the chemical space of G-quadruplex binders: Discovery of a novel chemotype targeting the human telomeric sequence. *J. Med. Chem.* 56, 9646–9654.
- Dickerson, R. E., and Drew, H. R. (1981). Structure of a B-DNA dodecamer. II. Influence of base sequence on helix structure. *J. Mol. Biol.* 149, 761–786.

- Drew, H. R., and Dickerson, R. E. (1981). Structure of a B-DNA dodecamer: III. Geometry of hydration. *J. Mol. Biol.* 151, 535–556.
- Drew, H. R., Wing, R. M., Takano, T., Broka, C., Tanaka, S., Itakura, K., et al. (1981). Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci.* 78, 2179–2183.
- Du, Z., Zhao, Y., and Li, N. (2008). Genome-wide analysis reveals regulatory role of G4 DNA in gene transcription. *Genome Res.* 18, 233–241.
- Dzatko, S., Krafcikova, M., Hänsel-Hertsch, R., Fessl, T., Fiala, R., Loja, T., et al. (2018). Evaluation of the Stability of DNA i-Motifs in the Nuclei of Living Mammalian Cells. *Angew. Chemie - Int. Ed.* 57, 2165–2169.
- Ellis, B. L., Duhme, A. K., Hider, R. C., Hossain, M. B., Rizvi, S., and van der Helm, D. (1996). Synthesis, Physicochemical Properties, and Biological Evaluation of Hydroxypyranones and Hydroxypyridinones: Novel Bidentate Ligands for Cell-Labeling. *J. Med. Chem.* 39, 3659–3670.
- Entzian, C., and Schubert, T. (2016). Studying small molecule-aptamer interactions using MicroScale Thermophoresis (MST). *Methods* 97, 27–34.
- Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995). A smooth particle mesh Ewald method. *J. Chem. Phys.* 103, 8577–8593.
- Fedoroff, O. Y., Rangan, A., Chemeris, V. V., and Hurley, L. H. (2000). Cationic porphyrins promote the formation of i-motif DNA and bind peripherally by a nonintercalative mechanism. *Biochemistry* 39, 15083–15090.
- Feng, Y., Yang, D., Chen, H., Cheng, W., Wang, L., Sun, H., et al. (2016). Stabilization of G-quadruplex DNA and inhibition of Bcl-2 expression by a pyridostatin analog. *Bioorganic Med. Chem. Lett.* 26, 1660–1663.
- Fernando, H., Reszka, A. P., Huppert, J., Ladame, S., Rankin, S., Venkitaraman, A. R., et al. (2006). A Conserved Quadruplex Motif Located in a Transcription Activation Site of the Human c-kit Oncogene. *Biochemistry* 45, 7854–7860.
- Finetti, F., Basile, A., Capasso, D., Di, S., Di, R., Pascale, M., et al. (2012). Functional and pharmacological characterization of a VEGF mimetic peptide on reparative angiogenesis. *Biochem. Pharmacol.* 84, 303–311.
- Franceschin, M., Rossetti, L., D'Ambrosio, A., Schirripa, S., Bianco, A., Ortaggi, G., et al. (2006). Natural and synthetic G-quadruplex interactive berberine derivatives. *Bioorganic Med. Chem. Lett.* 16, 1707–1711.
- Gaffarogullari, E. C., Krause, A., Balbo, J., Herten, D. P., and Jäschke, A. (2013). Microscale thermophoresis provides insights into mechanism and thermodynamics of ribozyme catalysis. *RNA Biol.* 10, 1815–1821.
- Gallego, J., Chou, S. H., and Reid, B. R. (1997). Centromeric pyrimidine strands fold into an intercalated motif by forming a double hairpin with a novel T:G:G:T tetrad: Solution structure of the d(TCCCGTTTCCA) dimer. *J. Mol. Biol.* 273, 840–856.
- Gao, Y., Shen, J. K., Milane, L., Hornicek, F. J., Amiji, M. M., and Duan, Z. (2015). Targeted cancer therapy; nanotechnology approaches for overcoming drug resistance. *Curr. Med. Chem.* 22, 1335–1347.
- Gavathiotis, E., Heald, R. A., Stevens, M. F. G., and Searle, M. S. (2001). Recognition and Stabilization of Quadruplex DNA by a Potent New Telomerase Inhibitor: NMR Studies of the 2 : 1 Complex of a Pentacyclic Methylacridinium Cation with. *Angew. Chemie - Int. Ed.* 40, 4749–4751.
- Gehring, K., Leroy, J.-L., and Guéron, M. (1993). A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature* 363, 561–565.
- Giancola, C., and Pagano, B. (2013). Energetics of Ligand Binding to G-Quadruplexes. *Top. Curr. Chem.* 330, 211–242.
- Gill, P., Moghadam, T. T., and Ranjbar, B. (2010). Differential scanning calorimetry techniques: applications in biology and nanoscience. *J. Biomol. Tech.* 21, 167–93.
- Giustiniano, M., Meneghetti, F., Mercalli, V., Varese, M., Giustiniano, F., Novellino, E., et al. (2014). Synthesis of aminocarbonyl N-Acylhydrazones by a three-component reaction of isocyanides, hydrazonoyl chlorides, and carboxylic acids. *Org. Lett.* 16, 5332–5335.
- Giustiniano, M., Mercalli, V., Amato, J., Novellino, E., and Tron, G. C. (2015). Exploiting the

- Electrophilic and Nucleophilic Dual Role of Nitrile Imines: One-Pot, Three-Component Synthesis of Furo[2,3-d]pyridazin-4(5H)-ones. *Org. Lett.* 17, 3964–3967.
- Gomez, D., Aouali, N., Londoño-Vallejo, A., Lacroix, L., Mégnin-Chanet, F., Lemarteleur, T., et al. (2003). Resistance to the Short Term Antiproliferative Activity of the G-quadruplex Ligand 12459 Is Associated with Telomerase Overexpression and Telomere Capping Alteration. *J. Biol. Chem.* 278, 50554–50562.
- Gomez, D., O'Donohue, M. F., Wenner, T., Douarre, C., Macadré, J., Koebel, P., et al. (2006). The G-quadruplex ligand telomestatin inhibits POT1 binding to telomeric sequences in vitro and induces GFP-POT1 dissociation from telomeres in human cells. *Cancer Res.* 66, 6908–6912.
- González, V., and Hurley, L. H. (2010). The c-*MYC* NHE III₁: Function and Regulation. *Annu. Rev. Pharmacol. Toxicol.* 50, 111–129.
- Gowan, S. M., Harrison, J. R., Patterson, L., Valenti, M., Read, M. a, Neidle, S., et al. (2002). A G-quadruplex-interactive potent small-molecule inhibitor of telomerase exhibiting in vitro and in vivo antitumor activity. *Mol. Pharmacol.* 61, 1154–1162.
- Gray, R. D., Trent, J. O., and Chaires, J. B. (2014). Folding and unfolding pathways of the human telomeric G-quadruplex. *J. Mol. Biol.* 426, 1629–1650.
- Greider, C. W., and Blackburn, E. H. (1985). Identification of a specific telomere terminal transferase activity in tetrahymena extracts. *Cell* 43, 405–413.
- Gros, J., Alberti, P., Mergny, J., and Gue, A. (2010). How long is too long? Effects of loop size on G-quadruplex stability. *Nucleic Acids Res.* 38, 7858–7868.
- Guédin, A., Lacroix, L., and Mergny, J.-L. (2010). Thermal melting studies of ligand DNA interactions. *Methods Mol. Biol.* 613, 25–35.
- Guéron, M., and Leroy, J. (2000). The i-motif in nucleic acids. *Curr. Opin. Struct. Biol.* 10, 326–331.
- Gunaratnam, M., Greciano, O., Martins, C., Reszka, A. P., Schultes, C. M., Morjani, H., et al. (2007). Mechanism of acridine-based telomerase inhibition and telomere shortening. *Biochem. Pharmacol.* 74, 679–689.
- Guo, K., Gokhale, V., Hurley, L. H., and Sun, D. (2008). Intramolecularly folded G-quadruplex and i-motif structures in the proximal promoter of the vascular endothelial growth factor gene. *Nucleic Acids Res.* 36, 4598–4608.
- Hackett, J. A., Feldser, D. M., and Greider, C. W. (2001). Telomere Dysfunction Increases Mutation Rate and Genomic Instability. *Cell* 106, 275–286.
- Hammond-kosack, M. C. U., Dobrinski, B., Lurz, R., Docherty, K., and Kilpatrick, M. W. (1992). The human insulin gene linked polymorphic region exhibits an altered DNA structure. *Nucleic Acids Res.* 20, 231–236.
- Hemann, M. T., and Greider, C. W. (1999). G-strand overhangs on telomeres in telomerase-deficient mouse cells. *Nucleic Acids Res.* 27, 3964–3969.
- Hengartner, M. O. (2000). The biochemistry of apoptosis. *Nature* 407, 770–776.
- Hsu, S. D., Varnai, P., Bugaut, A., Reszka, A. P., Neidle, S., and Balasubramanian, S. (2009). A G-Rich Sequence within the c-kit Oncogene Promoter Forms a Parallel G-Quadruplex Having Asymmetric G-Tetrad Dynamics. *J Am Chem Soc.* 13399–13409.
- http://www.glenresearch.com/Technical/TB_OAS.pdf
- Huang, H.-S., Chen, I.-B., Huang, K.-F., Lu, W.-C., Shieh, F.-Y., Huang, Y.-Y., et al. (2007). Synthesis and human telomerase inhibition of a series of regioisomeric disubstituted amidoanthraquinones. *Chem. Pharm. Bull. (Tokyo)*. 55, 284–292.
- Huang, M., Shen, A., Ding, J., and Geng, M. (2014). Molecularly targeted cancer therapy: Some lessons from the past decade. *Trends Pharmacol. Sci.* 35, 41–50.
- HUEY, R., Morris, G. M., Olson, A. J., and Goodsell, D. S. (2007). A Semiempirical Free Energy Force Field with Charge-Based Desolvation. *J. Comput. Chem.* 28, 1545–1152.
- Huppert, J. L. (2010). Structure, location and interactions of G-quadruplexes. *FEBS J.* 277, 3452–3458.
- Huppert, J. L., and Balasubramanian, S. (2007). G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* 35, 406–413.
- Hwang, T. L., and Shaka, A. J. (1995). Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients. *J. Magn. Reson. Ser. A* 112, 275–279.

- Irwin, J. J., and Shoichet, B. K. (2005). ZINC – A Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model* 45, 177–182.
- Izbicka, E., Wheelhouse, R. T., Raymond, E., Davidson, K. K., Lawrence, R. A., Sun, D., et al. (1999). Effects of cationic porphyrins as G-quadruplex interactive agents in human tumor cells. *Cancer Res.* 59, 639–644.
- Jana, J., Mondal, S., Bhattacharjee, P., Sengupta, P., Roychowdhury, T., Saha, P., et al. (2017). Chelerythrine down regulates expression of VEGFA, BCL2 and KRAS by arresting G-Quadruplex structures at their promoter regions. *Sci. Rep.* 7, 1–14.
- Jaroszewski, J. W., Clausen, V., Cohen, J. S., and Dahl, O. (1996). NMR investigations of duplex stability of phosphorothioate and phosphorodithioate DNA analogues modified in both strands. *Nucleic Acids Res.* 24, 829–834.
- Jerabek-Willemsen, M., André, T., Wanner, R., Roth, H. M., Duhr, S., Baaske, P., et al. (2014). MicroScale Thermophoresis: Interaction analysis and beyond. *J. Mol. Struct.* 1077, 101–113.
- Jerabek-Willemsen, M., Wienken, C. J., Braun, D., Baaske, P., and Duhr, S. (2011). Molecular Interaction Studies Using Microscale Thermophoresis. *Assay Drug Dev. Technol.* 9, 342–353.
- Joensuu, H., Pylkkänen, L., and Toikkanen, S. (1994). Bcl-2 protein expression and long-term survival in breast cancer. *Am. J. Pathol.* 145, 1191–1198.
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impley, R. W., and Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935.
- Kanaori, K., Shibayama, N., Gohda, K., Tajima, K., and Makino, K. (2001). Multiple four-stranded conformations of human telomere sequence d (CCCTAA) in solution. *Nucleic Acids Res.* 29, 831–840.
- Kang, H. J., Kendrick, S., Hecht, S. M., and Hurley, L. H. (2014). The transcriptional complex between the BCL2 i-motif and hnRNP LL is a molecular switch for control of gene expression that can be modulated by small molecules. *J. Am. Chem. Soc.* 136, 4172–4185.
- Karsisiotis, A. I., Hessari, N. M. A., Novellino, E., Spada, G. P., Randazzo, A., and Webba Da Silva, M. (2011). Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism. *Angew. Chemie - Int. Ed.* 50, 10645–10648.
- Kendrick, S., Kang, H. J., Alam, M. P., Madathil, M. M., Agrawal, P., Gokhale, V., et al. (2014). The dynamic character of the BCL2 promoter i-motif provides a mechanism for modulation of gene expression by compounds that bind selectively to the alternative DNA hairpin structure. *J. Am. Chem. Soc.* 136, 4161–4171.
- Keren-Kaplan, T., Attali, I., Estrin, M., Kuo, L. S., Farkash, E., Jerabek-Willemsen, M., et al. (2013). Structure-based in silico identification of ubiquitin-binding domains provides insights into the ALIX-V:ubiquitin complex and retrovirus budding. *EMBO J.* 32, 538–551.
- Kim, N. W., Piatyszek, M. A., Prowse, K. R., Harley, C. B., West, M. D., Ho, P. L. C., et al. (2011). Specific Association of Human Telomerase Activity with Immortal Cells and Cancer. 266, 2011–2016.
- Kovanda, A., Zalar, M., Šket, P., Plavec, J., and Rogelj, B. (2015). Anti-sense DNA d(GGCCCC)nexpansions in C9ORF72 form i-motifs and protonated hairpins. *Sci. Rep.* 5, 3–9.
- Krauss, I. R., Merlino, A., Randazzo, A., Novellino, E., Mazzarella, L., Sica, F., et al. (2012). High-resolution structures of two complexes between thrombin and thrombin-binding aptamer shed light on the role of cations in the aptamer inhibitory activity. 40, 8119–8128.
- Kuryavyi, V., Phan, A. T., and Patel, D. J. (2010). Solution structures of all parallel-stranded monomeric and dimeric G-quadruplex scaffolds of the human c-kit2 promoter. *Nucleic Acids Res.* 38, 6757–6773.
- Kypr, J., Kejnovská, I., Renčiuk, D., and Vorlíčková, M. (2009). Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.* 37, 1713–1725.
- Larsen, A., and Weintraub, H. (1982). An altered DNA conformation detected by S1 nuclease occurs at specific regions in active chick globin chromatin. *Cell* 29, 609–622.
- Leonetti, C. (2004). Biological Activity of the G-Quadruplex Ligand RHPS4 (3,11-Difluoro-6,8,13-trimethyl-8H-quinolo[4,3,2-kl]acridinium methosulfate) Is Associated with Telomere Capping

- Alteration. *Mol. Pharmacol.* 66, 1138–1146.
- Li, J. J., Fang, X., and Tan, W. (2002). Molecular aptamer beacons for real-time protein recognition. *Biochem. Biophys. Res. Commun.* 292, 31–40.
- Li, Q., Xiang, J. F., Yang, Q. F., Sun, H. X., Guan, A. J., and Tang, Y. L. (2013). G4LDB: A database for discovering and studying G-quadruplex ligands. *Nucleic Acids Res.* 41, D1115–D1123.
- Li, X., Peng, Y., Ren, J., and Qu, X. (2006). Carboxyl-modified single-walled carbon nanotubes selectively induce human telomeric i-motif formation. *Proc. Natl. Acad. Sci.* 103, 19658–19663.
- Limongelli, V., Tito, S. De, Cerofolini, L., Fragai, M., Pagano, B., Trotta, R., et al. (2013). The G-Triplex DNA. *Angew. Chemie - Int. Ed.* 52, 2269–2273.
- Lin, C. C., Melo, F. A., Ghosh, R., Suen, K. M., Stagg, L. J., Kirkpatrick, J., et al. (2012). Inhibition of basal FGF receptor signaling by dimeric Grb2. *Cell* 149, 1514–1524.
- Lipps, H. J., and Rhodes, D. (2009). G-quadruplex structures : in vivo evidence and function. *Cell.*
- Liu, Z., Delavan, B., Roberts, R., and Tong, W. (2017). Lessons Learned from Two Decades of Anticancer Drugs. *Trends Pharmacol. Sci.* 38, 852–872.
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2- $\Delta\Delta$ CT T Method. *Methods* 25, 402–408.
- Luu, K. N., Phan, A. T., Kuryavyi, V., Lacroix, L., and Patel, D. J. (2006a). Structure of the human telomere in K⁺ solution: An intramolecular (3 + 1) G-quadruplex scaffold. *J. Am. Chem. Soc.* 128, 9963–9970.
- Luu, K. N., Tu, A., Kuryavyi, V., Lacroix, L., and Patel, D. J. (2006b). Structure of the Human Telomere in K⁺ Solution: An Intramolecular (3 + 1) G-Quadruplex Scaffold. *J Am Chem Soc*, 9963–9970.
- Ma, D. L., Ma, V. P. Y., Chan, D. S. H., Leung, K. H., Zhong, H. J., and Leung, C. H. (2012). In silico screening of quadruplex-binding ligands. *Methods* 57, 106–114.
- Maizels, N. (2015). G4-associated human diseases. *EMBO Rep.* 16, 910–922.
- Majima, J., and Tetsuro, C. (2011). Conformational changes of non-B DNA. *Chem. Soc. Rev.* 40, 5893–5909.
- Malliavin, T. E., Snoussi, K., and Leroy, J. L. (2003). The NMR structure of [Xd(C2)]₄ investigated by molecular dynamics simulations. *Magn. Reson. Chem.* 41, 18–25.
- Masiero, S., Trotta, R., Pieraccini, S., De Tito, S., Perone, R., Randazzo, A., et al. (2010a). A non-empirical chromophoric interpretation of CD spectra of DNA G-quadruplex structures. *Org. Biomol. Chem.* 8, 2683.
- Masiero, S., Trotta, R., Pieraccini, S., De Tito, S., Perone, R., Randazzo, A., et al. (2010b). A non-empirical chromophoric interpretation of CD spectra of DNA G-quadruplex structures. *Org. Biomol. Chem.* 8, 2683–2692.
- Mathad, R. I., Hatzakis, E., Dai, J., and Yang, D. (2011). C-MYC promoter G-quadruplex formed at the 5'-end of NHE III element: Insights into biological relevance and parallel-stranded G-quadruplex stability. *Nucleic Acids Res.* 39, 9023–9033.
- Mayer, G. (2009). The chemical biology of aptamers. *Angew. Chemie - Int. Ed.* 48, 2672–2689.
- Mergny, J. L., and Maurizot, J. C. (2001). Fluorescence resonance energy transfer as a probe for G-quartet formation by a telomeric repeat. *ChemBioChem* 2, 124–132.
- Micheli, E., Altieri, A., Cianni, L., Cingolani, C., Iachettini, S., Bianco, A., et al. (2016). Perylene and coronene derivatives binding to G-rich promoter oncogene sequences efficiently reduce their expression in cancer cells. *Biochimie* 125, 223–231.
- Mitsunobu, O. (1981). The use of diethylazodicarboxylate and triphenylphosphine in synthesis and transformation of nature products. *Synthesis (Stuttg).* 1, 1–28.
- Monchaud, D., Allain, C., Bertrand, H., Smargiasso, N., Rosu, F., Gabelica, V., et al. (2008). Ligands playing musical chairs with G-quadruplex DNA: A rapid and simple displacement assay for identifying selective G-quadruplex binders. *Biochimie* 90, 1207–1223.
- Monchaud, D., Allain, C., and Teulade-Fichou, M.-P. (2006). Development of a fluorescent intercalator displacement assay (G4-FID) for establishing quadruplex-DNA affinity and selectivity of putative ligands. *Bioorg. Med. Chem. Lett.* 16, 4842–4845.
- Monchaud, D., and Teulade-Fichou, M. P. (2008). A hitchhiker's guide to G-quadruplex ligands. *Org. Biomol. Chem.* 6, 627–636.

- Moraca, F., Amato, J., Ortuso, F., Artese, A., Pagano, B., and Novellino, E. (2017). Ligand binding to telomeric G-quadruplex DNA investigated by funnel-metadynamics simulations. *Proc. Natl. Acad. Sci.* 114, E2136–E2145.
- Morita, Y., Leslie, M., Kameyama, H., Volk, D. E., and Tanaka, T. (2018). Aptamer therapeutics in cancer: Current and future. *Cancers (Basel)*. 10.
- Morris, G., and Huey, R. (2009). AutoDock4 and autoDockTools4: automated docking with selective receptor flexibility. *J. ...* 30, 2785–2791.
- Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., et al. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* 19, 1639–1662.
- Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., et al. (2009). AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility GARRETT. *J. Comput. Chem.* 30, 2785–2791.
- Moye, A. L., Porter, K. C., Cohen, S. B., Phan, T., Zyner, K. G., Sasaki, N., et al. (2015). Telomeric G-quadruplexes are a substrate and site of localization for human telomerase. *Nat. Commun.* 6, 1–12.
- Müller, S. and, and Rodriguez, R. (2014). G-quadruplex interacting small molecules and drugs : from bench toward bedside. *Expert Rev. Clin. Pharmacol.* 7, 663–679.
- Musumeci, D., Amato, J., Randazzo, A., Novellino, E., Giancola, C., Montesarchio, D., et al. (2014). G-quadruplex on oligo affinity support (G4-OAS): An easy affinity chromatography-based assay for the screening of G-quadruplex ligands. *Anal. Chem.* 86, 4126–4130.
- Musumeci, D., Amato, J., Zizza, P., Platella, C., Cosconati, S., Cingolani, C., et al. (2017a). Tandem application of ligand-based virtual screening and G4-OAS assay to identify novel G-quadruplex-targeting chemotypes. *Biochim. Biophys. Acta - Gen. Subj.* 1861, 1341–1352.
- Musumeci, D., Platella, C., Riccardi, C., Merlino, A., Marzo, T., Massai, L., et al. (2016). A first-in-class and a fished out anticancer platinum compound: cis-[PtCl₂(NH₃)₂] and cis-[PtI₂(NH₃)₂] compared for their reactivity towards DNA model systems. *Dalt. Trans.* 45, 8587–8600.
- Musumeci, D., Platella, C., Riccardi, C., Moccia, F., and Montesarchio, D. (2017b). Fluorescence sensing using DNA Aptamers in cancer research and clinical diagnostics. *Cancers (Basel)*. 9, 1–43.
- Nagatoishi, S., Tanaka, Y., and Tsumoto, K. (2007). Circular dichroism spectra demonstrate formation of the thrombin-binding DNA aptamer G-quadruplex under stabilizing-cation-deficient conditions. *Biochem. Biophys. Res. Commun.* 352, 812–817.
- Neidle, S. (2009). The structures of quadruplex nucleic acids and their drug complexes. *Curr. Opin. Struct. Biol.* 19, 239–250.
- Neidle, S. (2016). Quadruplex Nucleic Acids as Novel Therapeutic Targets. *J. Med. Chem.* 59, 5987–6011.
- Neidle, S., and Parkinson, G. (2002). TELOMERE MAINTENANCE AS A TARGET FOR ANTICANCER DRUG DISCOVERY. *Nat. Rev. Drug Discov.* 1, 383–393.
- Nonin-lecomte, S., and Leroy, J. L. (2001). Structure of a C-rich Strand Fragment of the Human Centromeric Satellite III : A pH-dependent Intercalation Topology. *J. Mol. Biol.* 309, 491–506.
- Nussbaumer, S., Bonabry, P., Veuthey, J. L., and Fleury-Souverain, S. (2011). Analysis of anticancer drugs: A review. *Talanta* 85, 2265–2289.
- Ohnmacht, S. A., and Neidle, S. (2014). Small-molecule quadruplex-targeted drug discovery. *Bioorganic Med. Chem. Lett.* 24, 2602–2612.
- Ou, T. M., Lu, Y. J., Tan, J. H., Huang, Z. S., Wong, K. Y., and Gu, L. Q. (2008). G-quadruplexes: Targets in anticancer drug design. *ChemMedChem* 3, 690–713.
- Paeschke, K., Simonsson, T., Postberg, J., Rhodes, D., and Lipps, H. J. (2005). Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo. *Nat. Struct. Mol. Biol.* 12, 847–854.
- Pagano, A., Iaccarino, N., Abdelhamid, M. A. S., Brancaccio, D., Garzarella, E. U., Di Porzio, A., et al. (2018). Common G-Quadruplex Binding Agents Found to Interact With i-Motif-Forming DNA: Unexpected Multi-Target-Directed Compounds. *Front. Chem.* 6, 1–13.
- Pagano, B., Amato, J., Iaccarino, N., Cingolani, C., Zizza, P., Biroccio, A., et al. (2015). Looking for

- efficient G-quadruplex ligands: Evidence for selective stabilizing properties and telomere damage by drug-like molecules. *ChemMedChem* 10, 640–649.
- Pagano, B., Cosconati, S., Gabelica, V., Petraccone, L., De Tito, S., Marinelli, L., et al. (2012). State-of-the-art methodologies for the discovery and characterization of DNA G-quadruplex binders. *Curr. Pharm. Des.* 18, 1880–1899.
- Pagano, B., Fotticchia, I., De Tito, S., Mattia, C. a, Mayol, L., Novellino, E., et al. (2010). Selective Binding of Distamycin A Derivative to G-Quadruplex Structure [d(TGGGGT)]₄. *J. Nucleic Acids* 2010, 1–7.
- Pagano, B., Mattia, C. A., Virno, A., Randazzo, A., Mayol, L., and Giancola, C. (2007). Thermodynamic analysis of quadruplex DNA-drug interaction. in *Nucleosides, Nucleotides and Nucleic Acids*, 761–765.
- Pagano, B., Randazzo, A., Fotticchia, I., Novellino, E., Petraccone, L., and Giancola, C. (2013). Differential scanning calorimetry to investigate G-quadruplexes structural stability. *Methods* 64, 43–51.
- Pairs, C. Á. C., Catasti, P., Chen, X., Deaven, L. L., Moyzis, R. K., Bradbury, E. M., et al. (1997). Cytosine-rich Strands of the Insulin Minisatellite Adopt Hairpins with Intercalated. *J. Mol. Biol.*
- Palumbo, S. M. L., Ebbinghaus, S. W., and Hurley, L. H. (2009). Formation of a unique end-to-end stacked pair of G-quadruplexes in the hTERT core promoter with implications for inhibition of telomerase by G-quadruplex-interactive ligands. *J. Am. Chem. Soc.* 131, 10878–10891.
- Pane, K., Sgambati, V., Zanfardino, A., Smaldone, G., Cafaro, V., Angrisano, T., et al. (2016). A new cryptic cationic antimicrobial peptide from human apolipoprotein E with antibacterial activity and immunomodulatory effects on human cells. *FEBS J.* 283, 2115–2131.
- Parkinson, G. N., Lee, M. P. H., and Neidle, S. (2002). Crystal structure of parallel quadruplexes from human telomeric DNA. *natu* 417, 376–380.
- Patel, D. J., Phan, A. T., and Kuryavyi, V. (2007). Human telomere , oncogenic promoter and 5' -UTR G-quadruplexes : diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res.* 35, 7429–7455.
- Patnaik, S., Zheng, W., Choi, J. H., Motabar, O., Southall, N., Westbrook, W., et al. (2012). Discovery, structure - Activity relationship, and biological evaluation of noninhibitory small molecule chaperones of glucocerebrosidase. *J. Med. Chem.* 55, 5734–5748.
- Pérez, A., Marchán, I., Svozil, D., Sponer, J., Cheatham, T. E., Laughton, C. A., et al. (2007). Refinement of the AMBER force field for nucleic acids: Improving the description of α/γ conformers. *Biophys. J.* 92, 3817–3829.
- Petraccone, L., Malafrente, A., Amato, J., and Giancola, C. (2012). G-quadruplexes from human telomeric DNA: How many conformations in PEG containing solutions? *J. Phys. Chem. B* 116, 2294–2305.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., et al. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612.
- Phan, A. T. (2010). Human telomeric G-quadruplex : structures of DNA and RNA sequences. *FEBS J.* 277, 1107–1117.
- Phan, A. T., Guéron, M., and Leroy, J. L. (2000). The solution structure and internal motions of a fragment of the cytidine-rich strand of the human telomere. *J. Mol. Biol.* 299, 123–144.
- Phan, A. T., Kuryavyi, V., Burge, S., Neidle, S., and Patel, D. J. (2007a). Structure of an unprecedented G-quadruplex scaffold in the human c-kit promoter. *J. Am. Chem. Soc.* 129, 4386–4392.
- Phan, A. T., Kuryavyi, V., Luu, K. N., and Patel, D. J. (2007b). Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K⁺ solution. *Nucleic Acids Res.* 35, 6517–6525.
- Phatak, P., Cookson, J. C., Dai, F., Smith, V., Gartenhaus, R. B., Stevens, M. F. G., et al. (2007). Telomere uncapping by the G-quadruplex ligand RHPS4 inhibits clonogenic tumour cell growth in vitro and in vivo consistent with a cancer stem cell targeting mechanism. *Br. J. Cancer* 96, 1223–1233.
- Platella, C., and Al., E. The content is object of the Italian patent application n. 102017000030473 filed

- on March 20, 2017.
- Platella, C., Guida, S., Bonmassar, L., Aquino, A., Bonmassar, E., Ravagnan, G., et al. (2017a). Antitumour activity of resveratrol on human melanoma cells: A possible mechanism related to its interaction with malignant cell telomerase. *Biochim. Biophys. Acta - Gen. Subj.* 1861, 2843–2851.
- Platella, C., Musumeci, D., Arciello, A., Doria, F., Freccero, M., Randazzo, A., et al. (2018). Controlled Pore Glass-based oligonucleotide affinity support: towards High Throughput Screening methods for the identification of conformation-selective G-quadruplex ligands. *Anal. Chim. Acta* 1030, 133–141.
- Platella, C., Riccardi, C., Montesarchio, D., Roviello, G. N., and Musumeci, D. (2017b). G-quadruplex-based aptamers against protein targets in therapy and diagnostics. *Biochim. Biophys. Acta - Gen. Subj.* 1861, 1429–1447.
- Prime, M. E., Courtney, S. M., Brookfield, F. A., Marston, R. W., Walker, V., Warne, J., et al. (2011). Phthalazinone pyrazoles as potent, selective, and orally bioavailable inhibitors of Aurora-A kinase. *J. Med. Chem.* 54, 312–319.
- Proctor, W. G., and Yu, F. C. (1950). The dependence of a nuclear magnetic resonance frequency upon chemical compound. *Phys. Rev.* 77, 717.
- Qin, Y., and Hurley, L. H. (2008). Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions. 90.
- Qin, Y., Rezler, E. M., Gokhale, V., Sun, D., and Hurley, L. H. (2007). Characterization of the G-quadruplexes in the duplex nuclease hypersensitive element of the PDGF-A promoter and modulation of PDGF-A promoter activity by TMPyP4. 35, 7698–7713.
- Rajendran, A., Endo, M., Hidaka, K., and Sugiyama, H. (2014). Direct and single-molecule visualization of the solution-state structures of G-hairpin and G-triplex intermediates. *Angew. Chemie - Int. Ed.* 53, 4107–4112.
- Rajendran, A., Nakano, S., and Sugimoto, N. (2010). Molecular crowding of the cosolutes induces an intramolecular i-motif structure of triplet repeat DNA oligomers at neutral pH. *Chem. Commun.* 46, 1299–1301.
- Randazzo, A., Spada, G. P., and Da Silva, M. W. (2013). Circular dichroism of quadruplex structures. *Top. Curr. Chem.* 330, 67–86.
- Rankin, S., Reszka, A. P., Huppert, J., Zloh, M., Parkinson, G. N., Todd, A. K., et al. (2005). Putative DNA quadruplex formation within the human c-kit oncogene. *J. Am. Chem. Soc.* 127, 10584–10589.
- Rawal, P., Bhadra, V., Kummarasetti, R., Ravindran, J., Kumar, N., Halder, K., et al. (2006). Genome-wide prediction of G4 DNA as regulatory motifs: Role in Escherichia coli global regulation. *Genome Res.* 16, 644–655.
- Reed, J. C., Cuddy, M., Haldart, S., Crocet, C., Nowell, P., Makover, D., et al. (1990). BCL2-mediated tumorigenicity of a human T-lymphoid cell line: Synergy with MYC and inhibition by BCL2 antisense. *Proc. Natl. Acad. Sci.* 87, 3660–3664.
- Reyes-Reyes, E. M., Teng, Y., and Bates, P. J. (2010). A new paradigm for aptamer therapeutic AS1411 action: Uptake by macropinocytosis and its stimulation by a nucleolin-dependent mechanism. *Cancer Res.* 70, 8617–8629.
- Rhodes, D., and Lipps, H. J. (2018). G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res.* 43, 8627–8637.
- Riou, J. F., Guittat, L., Mailliet, P., Laoui, A., Renou, E., Petitgenet, O., et al. (2002). Cell senescence and telomere shortening induced by a new series of specific G-quadruplex DNA ligands. *Proc. Natl. Acad. Sci.* 99, 2672–2677.
- Rodriguez, R., Müller, S., Yeoman, J. A., Trentesaux, C., Riou, J. F., and Balasubramanian, S. (2008). A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. *J. Am. Chem. Soc.* 130, 15758–15759.
- Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* 23, 327–341.
- Salvati, E., Leonetti, C., Rizzo, A., Scarsella, M., Mottolese, M., Galati, R., et al. (2007). Telomere

- damage induced by the G-quadruplex ligand RHPS4 has an antitumor effect. *J. Clin. Invest.* 117, 3236–3247.
- Salvati, E., Scarsella, M., Porru, M., Rizzo, A., Iachettini, S., Tentori, L., et al. (2010). PARP1 is activated at telomeres upon G4 stabilization: Possible target for telomere-based therapy. *Oncogene* 29, 6280–6293.
- Salvati, E., Zizza, P., Rizzo, A., Iachettini, S., Cingolani, C., D'angelo, C., et al. (2014). Evidence for G-quadruplex in the promoter of vegfr-2 and its targeting to inhibit tumor angiogenesis. *Nucleic Acids Res.* 42, 2945–2957.
- Sanner, M. F. (1999). Python: A programming language for software integration and development. *J. Mol. Graph. Model.* 17, 57–61.
- Schaffitzel, C., Berger, I., Postberg, J., Hanes, J., Lipps, H. J., and Plu, A. (2001). In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with *Stylynychia lemnae* macronuclei. *Proc. Natl. Acad. Sci.* 98, 8572–8577.
- Schultze, P., Macaya, R. F., and Feigon, J. (1994). Three-dimensional Solution Structure of the Thrombin-binding DNA Aptamer d(GGTTGGTGTGGTTGG). *J. Mol. Biol.* 235, 1532–1547.
- Seidel, S. A. I., Dijkman, P. M., Lea, W. A., van den Bogaart, G., Jerabek-Willemsen, M., Lasic, A., et al. (2013). Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions. *Methods* 59, 301–315.
- Seidel, S. A. I., Wienken, C. J., Geissler, S., Jerabek-Willemsen, M., Duhr, S., Reiter, A., et al. (2012). Label-free microscale thermophoresis discriminates sites and affinity of protein-ligand binding. *Angew. Chemie - Int. Ed.* 51, 10656–10659.
- Seto, M., Jaeger, U., Hockett, R. D., Graninger, W., Bennett, S., Goldman, P., et al. (1988). Alternative promoters and exons, somatic mutation and deregulation of the Bcl-2-Ig fusion gene in lymphoma. *EMBO J.* 7, 123–31.
- Sfeir, A. J., Chai, W., Shay, J. W., and Wright, W. E. (2005). Telomere-End Processing : the Terminal Nucleotides of Human Chromosomes. *Mol. Cell* 18, 131–138.
- Shao, J., Tanner, S. W., Thompson, N., and Cheatham, T. E. (2007). Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms. *J. Chem. Theory Comput.* 3, 2312–2334.
- Sheng, Q., Neaverson, J. C., Mahmoud, T., Stevenson, C. E. M., Matthews, S. E., and Waller, Z. A. E. (2017). Identification of new DNA i-motif binding ligands through a fluorescent intercalator displacement assay. *Org. Biomol. Chem.* 15, 5669–5673.
- Siddiqui-Jain, A., Grand, C. L., Bearss, D. J., and Hurley, L. H. (2002). 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.
- Simonsson, T., Pecinka, P., and Kubista, M. (1998). DNA tetraplex formation in the control region of c-myc. *Nucleic Acids Res.* 26, 1167–1172.
- Simonsson, T., Pribylova, M., and Vorlickova, M. (2000). A nuclease hypersensitive element in the human c-myc promoter adopts several distinct i-tetraplex structures. *Biochem. Biophys. Res. Commun.* 278, 158–166.
- Soundararajan, S., Chen, W., Spicer, E. K., Courtenay-Luck, N., and Fernandes, D. J. (2008). The nucleolin targeting aptamer AS1411 destabilizes Bcl-2 messenger RNA in human breast cancer cells. *Cancer Res.* 68, 2358–2365.
- Soundararajan, S., Wang, L., Sridharan, V., Chen, W., Courtenay-Luck, N., Jones, D., et al. (2009). Plasma Membrane Nucleolin Is a Receptor for the Anticancer Aptamer AS1411 in MV4-11 Leukemia Cells. *Mol. Pharmacol.* 76, 984–991.
- Sparapani, S., Bellini, S., Gunaratnam, M., Haider, S. M., Andreani, A., Rambaldi, M., et al. (2010). Bis-guanylhydrazono diimidazo[1,2-a:1,2-c]pyrimidine as a novel and specific G-quadruplex binding motif. *Chem. Commun.* 46, 5680–5682.
- Strauß, T., Van Poecke, R. M. P., Strauß, A., Römer, P., Minsavage, G. V., Singh, S., et al. (2012). RNA-seq pinpoints a *Xanthomonas* TAL-effector activated resistance gene in a large-crop genome. *Proc. Natl. Acad. Sci. U. S. A.* 109, 19480–19485.
- Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein*

- Expr. Purif.* 41, 207–234.
- Sun, D., Guo, K., Rusche, J. J., and Hurley, L. H. (2005). 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.
- Sun, D., and Hurley, L. H. (2009). The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: Implications for drug targeting and control of gene expression. *J. Med. Chem.* 52, 2863–2874.
- Sun, D., Thompson, B., Cathers, B. E., Salazar, M., Kerwin, S. M., Trent, J. O., et al. (1997). Inhibition of human telomerase by a G-Quadruplex-Interactive compound. *J. Med. Chem.* 40, 2113–2116.
- Takai, H., Smogorzewska, A., and De Lange, T. (2003). DNA damage foci at dysfunctional telomeres. *Curr. Biol.* 13, 1549–1556.
- Temime-Smaali, N., Guittat, L., Sidibe, A., Shin-Ya, K., Trentesaux, C., and Riou, J. F. (2009). The G-quadruplex ligand telomestatin impairs binding of topoisomerase III α to G-quadruplex-forming oligonucleotides and uncaps telomeres in ALT cells. *PLoS One* 4, 1–14.
- Thiriet, C., and Hayes, J. J. (2005). Chromatin in need of a fix: Phosphorylation of H2AX connects chromatin to DNA repair. *Mol. Cell* 18, 617–622.
- Todd, A. K., Johnston, M., and Neidle, S. (2005). Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res.* 33, 2901–2907.
- Van Bogaart, G. Den, Meyenberg, K., Diederichsen, U., and Jahn, R. (2012). Phosphatidylinositol 4,5-bisphosphate increases Ca²⁺ affinity of synaptotagmin-1 by 40-fold. *J. Biol. Chem.* 287, 16447–16453.
- Van Den Bogaart, G., Thutupalli, S., Risselada, J. H., Meyenberg, K., Holt, M., Riedel, D., et al. (2011). Synaptotagmin-1 may be a distance regulator acting upstream of SNARE nucleation. *Nat. Struct. Mol. Biol.* 18, 805–812.
- Viswanadha Vijaya Padma (2015). An overview of targeted cancer therapy. *Biomedicine* 5, 19.
- Voet, D., and Voet, J. (1995). *Biochemistry*. 2nd ed. , ed. John Wiley & Sons.
- Wang, G., and Vasquez, K. M. (2006). Non-B DNA structure-induced genetic instability. *Mutat. Res.* 598, 103–119.
- Wang, J. M., Wolf, R. M., Caldwell, J. W., Kollman, P. a, and Case, D. a (2004). Development and testing of a general amber force field. *J. Comput. Chem.* 25, 1157–1174.
- Wang, X. D., Ou, T. M., Lu, Y. J., Li, Z., Xu, Z., Xi, C., et al. (2010). Turning off transcription of the bcl-2 gene by stabilizing the bcl-2 promoter quadruplex with quindoline derivatives. *J. Med. Chem.* 53, 4390–4398.
- Wang, Y., and Patel, D. J. (1993). Solution Structure of a Parallel-stranded G-Quadruplex DNA. *J. Mol. Biol.* 234, 1171–1183.
- Watson, J. D., and Crick, F. H. D. (1953). A Structure for Deoxyribose Nucleic Acid. *Nature* 171, 737–738.
- Wei, D., Husby, J., and Neidle, S. (2015). Flexibility and structural conservation in a c-KIT G-quadruplex. *Nucleic Acids Res.* 43, 629–644.
- Wei, D., Parkinson, G. N., Reszka, A. P., and Neidle, S. (2012). Crystal structure of a c-kit promoter quadruplex reveals the structural role of metal ions and water molecules in maintaining loop conformation. *Nucleic Acids Res.* 40, 4691–4700.
- Wells, R. D., Dere, R., Hebert, M. L., Napierala, M., and Son, L. S. (2005). Advances in mechanisms of genetic instability related to hereditary neurological diseases. *Nucleic Acids Res.* 33, 3785–3798.
- Wheate, N. J., Walker, S., Craig, G. E., and Oun, R. (2010). The status of platinum anticancer drugs in the clinic and in clinical trials. *Dalt. Trans.* 39, 8113–8127.
- Wienken, C. J., Baaske, P., Rothbauer, U., Braun, D., and Duhr, S. (2010). Protein-binding assays in biological liquids using microscale thermophoresis. *Nat. Commun.* 1, 1–7.
- Williamson, J. R., Raghuraman, M. K., and Cech, T. R. (1989). Monovalent cation-induced structure of telomeric DNA: The G-quartet model. *Cell* 59, 871–880.
- Wilson, S. M., Schmutzler, B. S., Brittain, J. M., Dustrude, E. T., Ripsch, M. S., Pellman, J. J., et al. (2012). Inhibition of transmitter release and attenuation of anti-retroviral- associated and tibial

- nerve injury-related painful peripheral neuropathy by novel synthetic Ca²⁺-channel peptides. *J. Biol. Chem.* 287, 35065–35077.
- Wood, W. I., and Felsenfeld, G. (1982). Chromatin Structure of the Chicken P-Globin Gene Region. 257, 7730–7736.
- Wright, E. P., Day, H. A., Ibrahim, A. M., Kumar, J., Boswell, L. J. E., Huguin, C., et al. Mitoxantrone and analogues bind and stabilize i-motif forming DNA sequences. *Sci. Rep.* 6, 39456.
- Wu, X., Xu, P., Wang, J., Xu, Y., Fu, T., Zhao, M., et al. (2015). Theoretical Studies on the Folding Mechanisms for Different DNA G-quadruplexes. *Adv. Struct. Bioinforma.*, 123–141.
- Xi-Miao, H., Wen-Qiang, W., Xiao-Lei, D., Na-Nv, L., Hai-Hong, L., Jing, F., et al. (2015). Molecular mechanism of G-quadruplex unwinding helicase: sequential and repetitive unfolding of G-quadruplex by Pif1 helicase. *Biochem. J.* 466, 189–199.
- Xiong, X., Coombs, P. J., Martin, S. R., Liu, J., Xiao, H., McCauley, J. W., et al. (2013). Receptor binding by a ferret-transmissible H5 avian influenza virus. *Nature* 497, 392–396.
- Xu, L., Hong, S., Sun, N., Wang, K., Zhou, L., Ji, L., et al. (2016). Berberine as a novel light-up i-motif fluorescence ligand and its application in designing molecular logic systems. *Chem. Commun.* 52, 179–182.
- Xue, Y., Kan, Z. Y., Wang, Q., Yao, Y., Liu, J., Hao, Y. H., et al. (2007). Human telomeric DNA forms parallel-stranded intramolecular G-quadruplex in K⁺-solution under molecular crowding condition. *J. Am. Chem. Soc.* 129, 11185–11191.
- Yang, D., and Okamoto, K. (2010). Structural insights into G-quadruplexes: towards new anticancer drugs. *Futur. Med Chem* 2, 619–646.
- Yip, K. W., and Reed, J. C. (2008). Bcl-2 family proteins and cancer. *Oncogene* 27, 6398–6406.
- Young, J. D., Abbate, V., Imberti, C., Meszaros, L. K., Ma, M. T., Terry, S. Y. A., et al. (2017). ⁶⁸ Ga-THP-PSMA: A PET Imaging Agent for Prostate Cancer Offering Rapid, Room-Temperature, 1-Step Kit-Based Radiolabeling. *J. Nucl. Med.* 58, 1270–1277.
- Zakian, V. A. (1995). Telomeres : Beginning t Understand the End.
- Zeraati, M., Langley, D. B., Schofield, P., Moye, A. L., Rouet, R., Hughes, W. E., et al. (2018). I-motif DNA structures are formed in the nuclei of human cells. *Nat. Chem.* 10, 631–637.
- Zhang, Z., Dai, J., Veliath, E., Jones, R. A., and Yang, D. (2010). Structure of a two-G-tetrad intramolecular G-quadruplex formed by a variant human telomeric sequence in K⁺ solution : insights into the interconversion of human telomeric G-quadruplex structures. *Nucleic Acids Res.* 38, 1009–1021.
- Zhao, J., Bacolla, A., Wang, G., and Vasquez, K. M. (2010). Non-B DNA structure-induced genetic instability and evolution. *Cell. Mol. Life Sci.* 67, 43–62.
- Zhou, J., Wei, C., Jia, G., Wang, X., Feng, Z., and Li, C. (2010). Formation of i-motif structure at neutral and slightly alkaline pH. *Mol. BioSyst.* 6, 580–586.
- Zillner, K., Filarsky, M., Rachow, K., Weinberger, M., Längst, G., and Németh, A. (2013). Large-scale organization of ribosomal DNA chromatin is regulated by Tip5. *Nucleic Acids Res.* 41, 5251–5262.
- Zuffo, M., Doria, F., Botti, S., Bergamaschi, G., and Freccero, M. (2017). G-quadruplex fluorescence sensing by core-extended naphthalene diimides. *Biochim. Biophys. Acta - Gen. Subj.* 1861, 1303–1311.

APPENDIX

Paper I

“Toward the Development of Specific G-Quadruplex Binders: Synthesis, Biophysical, and Biological Studies of New Hydrazone Derivatives”. J. Amato, R. Morigi, B. Pagano, A. Pagano, S. Ohnmacht, A. De Magis, Y. Tiang, G. Capranico, A. Locatelli, A. Graziadio, A. Leoni, M. Rambaldi, E. Novellino, S. Neidle, and A. Randazzo. *J. Med. Chem.*, **2016**, 59, 5706-5720.

Paper II

“Targeting the *BCL2* Gene Promoter G-Quadruplex with a New Class of Furopyridazinone-Based Molecules”. J. Amato, A. Pagano, D. Capasso, S. Di Gaetano, M. Giustiniano, E. Novellino, A. Randazzo, and B. Pagano. *ChemMedChem*, **2018**, 13, 406-410.

Paper III

“Tailoring a lead-like compound targeting multiple G-quadruplex structures”. J. Amato, C. Platella, S. Iachettini, P. Zizza, D. Musumeci, S. Cosconati, A. Pagano, E. Novellino, A. Biroccio, A. Randazzo, B. Pagano, and D. Montesarchio. *Eur. J. Med. Chem.* **2019**, 163, 295-306.

Paper IV

“Discovery of the first dual G-triplex/G-quadruplex stabilizing compound: a new opportunity in the targeting of G-rich DNA structures?”. J. Amato, A. Pagano, S. Cosconati, G. Amendola, I. Fotticchia, N. Iaccarino, J. Marinello, A. De Magis, G. Capranico, E. Novellino, B. Pagano, A. Randazzo. *Biochimica et Biophysica Acta – General Subjects*, **2017**, 1861, 1271-1280.

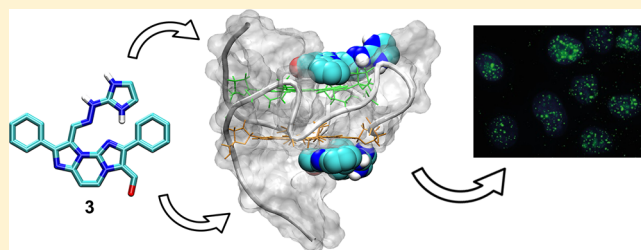
Paper V

“Common G-Quadruplex Binding Agents Found to Interact With i-Motif-Forming DNA: Unexpected Multi-Target-Directed Compounds”. A. Pagano, N. Iaccarino, M. A. S. Abdelhamid, D. Brancaccio, E. U. Garzarella, A. Di Porzio, E. Novellino, Z. A. E. Waller, B., J. Amato, and A. Randazzo. *Frontiers in Chemistry*, **2018**, 6, 281.

Toward the Development of Specific G-Quadruplex Binders: Synthesis, Biophysical, and Biological Studies of New Hydrazone Derivatives

Jussara Amato,^{†,‡} Rita Morigi,^{‡,‡} Bruno Pagano,[†] Alessia Pagano,[†] Stephan Ohnmacht,[§] Alessio De Magis,[‡] Yee-Peng Tiang,[‡] Giovanni Capranico,[‡] Alessandra Locatelli,[‡] Alessandra Graziadio,[‡] Alberto Leoni,[‡] Mirella Rambaldi,[‡] Ettore Novellino,[†] Stephen Neidle,[§] and Antonio Randazzo^{*,†}[†]Department of Pharmacy, University of Naples "Federico II", via D. Montesano 49, 80131 Napoli, Italy[‡]Department of Pharmacy and Biotechnology, University of Bologna, 40126 Bologna, Italy[§]UCL School of Pharmacy, University College London, London WC1N 1AX, U.K.**S** Supporting Information

ABSTRACT: G-Quadruplex-binding compounds are currently perceived as possible anticancer therapeutics. Here, starting from a promising lead, a small series of novel hydrazone-based compounds were synthesized and evaluated as G-quadruplex binders. The *in vitro* G-quadruplex-binding properties of the synthesized compounds were investigated employing both human telomeric and oncogene promoter G-quadruplexes with different folding topologies as targets. The present investigation led to the identification of potent G-quadruplex stabilizers with high selectivity over duplex DNA and preference for one G-quadruplex topology over others. Among them, selected derivatives have been shown to trap G-quadruplex structures in the nucleus of cancer cells. Interestingly, this behavior correlates with efficient cytotoxic activity in human osteosarcoma and colon carcinoma cells.

**■ INTRODUCTION**

The interest in noncanonical DNA secondary structures as targets for novel and selective anticancer drugs is exponentially growing, in particular due to their involvement in a number of critical cancer-related genomic aberrations and pathways.^{1–3} G-Quadruplexes (G4s) are now viewed as promising drug targets, and much effort is currently underway to discover effective and selective ligands for them.^{2,4} G4s are four-stranded structures composed of stacked guanine tetrads (G-tetrads) that can form within G-rich tracts of genomic DNA and depend on the presence of cations, especially K⁺ and Na⁺.⁵ G4-forming sequences are located in different critical positions of the human genome, notably at telomeres and in oncogene promoters.^{2,4}

Telomeric regions at the chromosome ends comprise 2–20 kb of double-stranded TTAGGG repeats together with a 3' single-stranded overhang of 50–500 nucleotides which can readily adopt G4 structures.^{4,5} It is known that telomeres are involved in the regulation of cellular proliferation and become gradually shorter, in parallel with normal cell proliferation, triggering replicative senescence, leading to irreversible cell-cycle arrest.⁶ Almost all cancer cells circumvent this problem by activating telomere-length maintenance mechanisms that facilitate their proliferation and survival. In about 85% of human tumors, the reverse transcriptase enzyme telomerase drives telomere

maintenance by adding copies of the TTAGGG repeat motif to the end of the single-stranded overhang.⁷ In the remaining ca. 15% of cancer cells, the same result is obtained by a different mechanism known as alternative lengthening of telomere (ALT).⁸ The inhibitory effect of G4 formation at telomere ends on the activity of telomerase is well established.⁹ The effect of G4-binding molecules, upon interaction with telomeric G4 DNA, is to induce G4 formation and then inhibit telomerase⁴ and also in a number of instances to induce rapid onset of DNA damage responses, leading to cellular apoptosis.¹⁰

The presence of putative G4s in the promoter regions of a number of oncogenes (e.g., *c-MYC*, *c-KIT*, *VEGF*, *RET*) and their function in transcriptional regulation is a more recent discovery that may have significant therapeutic implications.² Indeed, compounds that can promote formation of G4s in oncogene promoters or modulate their stability could, in turn, regulate expression of the target gene. This has developed into a novel strategy for cancer drug discovery.

Many research groups have subsequently attempted to identify G4-binding compounds and, to date, over 1000 ligands have been reported, representing a wide range of chemotypes.^{11,12} Some of these have been identified in order to target the grooves

Received: January 27, 2016

Published: May 25, 2016

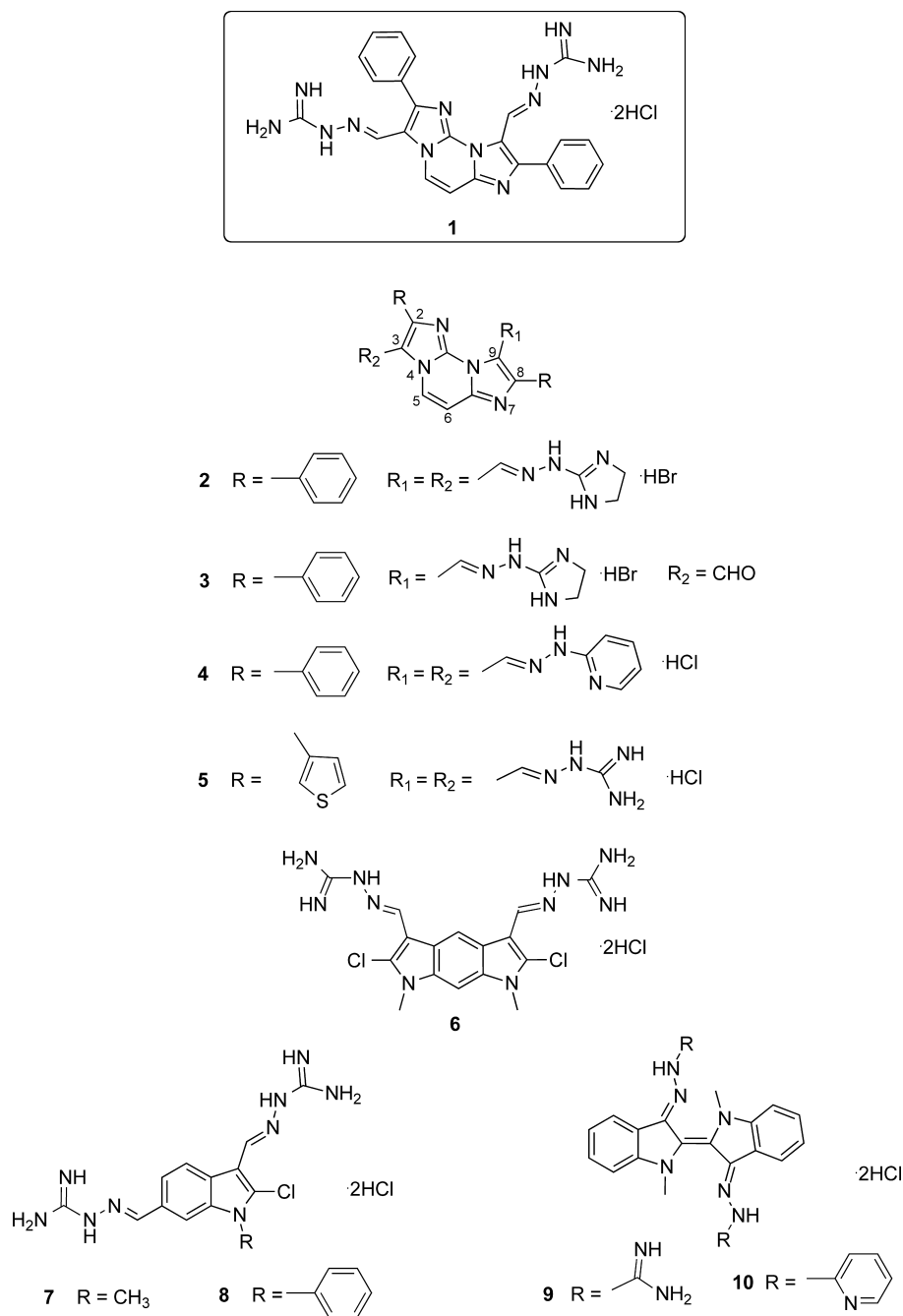


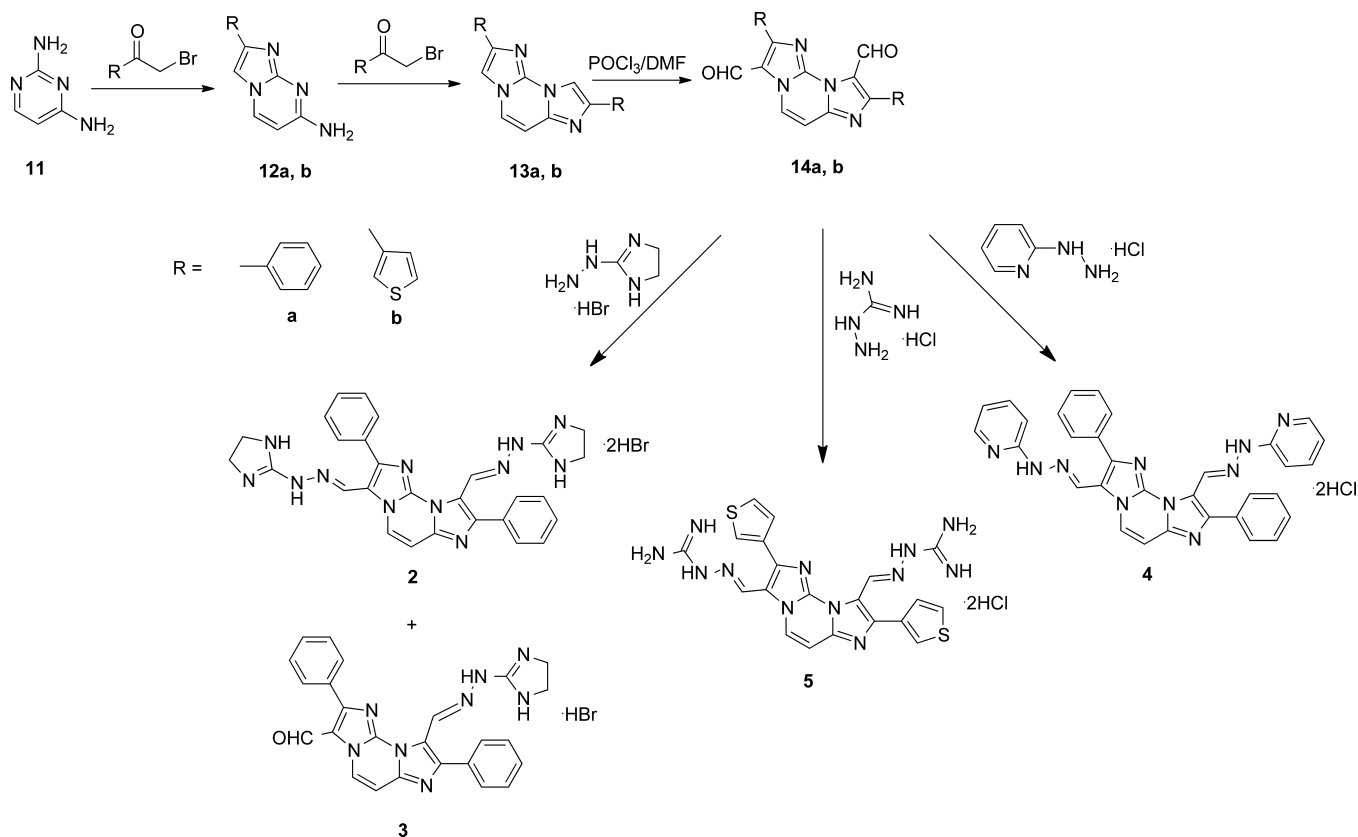
Figure 1. Compound lead **1** and structural modifications.

of G4 structures, although most have been devised to primarily stack on the external G-tetrads of G4s.^{11–15} The structural features that appear to facilitate the latter binding mode are (i) the presence of an extended planar aromatic scaffold that can efficiently stack on terminal G-tetrads through π – π stacking, (ii) an electron-deficient aromatic core that can enhance stacking interactions on G-tetrads via cation– π interaction, and (iii) the presence of cationic side-chains. These positively charged tethered substituents can interact with the grooves and loops of G4 and with the negatively charged phosphate backbone.

In 2010, some of us reported that the bis-guanyldiazene derivative of diimidazo[1,2-*a*:1,2-*c*]pyrimidine **1** (Figure 1) was a potent stabilizer of several G4 DNAs, whereas no significant interaction with duplex DNA was found.¹⁶ Its high affinity and selective G4-stabilizing ability led us to consider **1** as a promising

lead compound and to design a small series of analogues. Structural modifications of **1** involved in particular the guanyl chains, the substituents at positions 2 and 8 and the core chromophore itself (Figure 1). The guanyl groups have been modified by introducing a more rigid frame (**2**) or an aromatic system (**4**). Together with compound **2**, the monohydrazone **3** was isolated from the reaction mixture and has also been investigated. The phenyl rings at positions 2 and 8 have been replaced with the bioisostere thiophene to obtain compound **5**. The tricyclic core has been substituted with different systems: the pyrroloindole nucleus to obtain compound **6** or a chloroindole moiety (**7**, **8**), to achieve core simplification, and finally, core expansion was explored by introducing the planar bisindole system of indigo, which was functionalized with iminoguanyl or hydrazinylpyridyl groups in order to obtain compounds **9** and **10**.

Scheme 1. Synthetic Route for Hydrazones 2–5



The *in vitro* G4-binding properties of the synthesized compounds were investigated by employing both human telomeric and oncogene promoter G4s with different topologies as targets. The compounds that showed the best G4-binding properties in terms of specificity and stabilizing effects were further investigated employing biophysical methodologies. The subsequent analysis of the biological properties of selected compounds demonstrated that the molecular effects on G4 stability in living cells were associated with the biological activity of these hydrazone derivatives.

RESULTS AND DISCUSSION

Chemistry. The hydrazones 2–5 (Scheme 1) and 6–10 (Scheme 2) were prepared by reaction between an aldehyde or a ketone and the appropriate hydrazine: aminoguanidine hydrochloride, 2-hydrazino-2-imidazoline hydrobromide, or 2-hydrazino-2-pyridine hydrochloride. When this reaction was performed with 2-hydrazino-2-imidazoline on aldehyde 14a,¹⁷ in addition to the bis hydrazone 2, the mono hydrazone derivative 3 was isolated from the reaction mixture. Two NOE experiments were used to determine which aldehyde group reacted. The irradiation of the CHO at 10.09 ppm provided a NOE correlation with one of the pyrimidine proton at 9.12 ppm, whereas when the CH=N at 9.47 ppm was irradiated, no NOE effect at the pyrimidine proton was observed. Taken together, these two findings suggest that the unreacted CHO is at position 3.

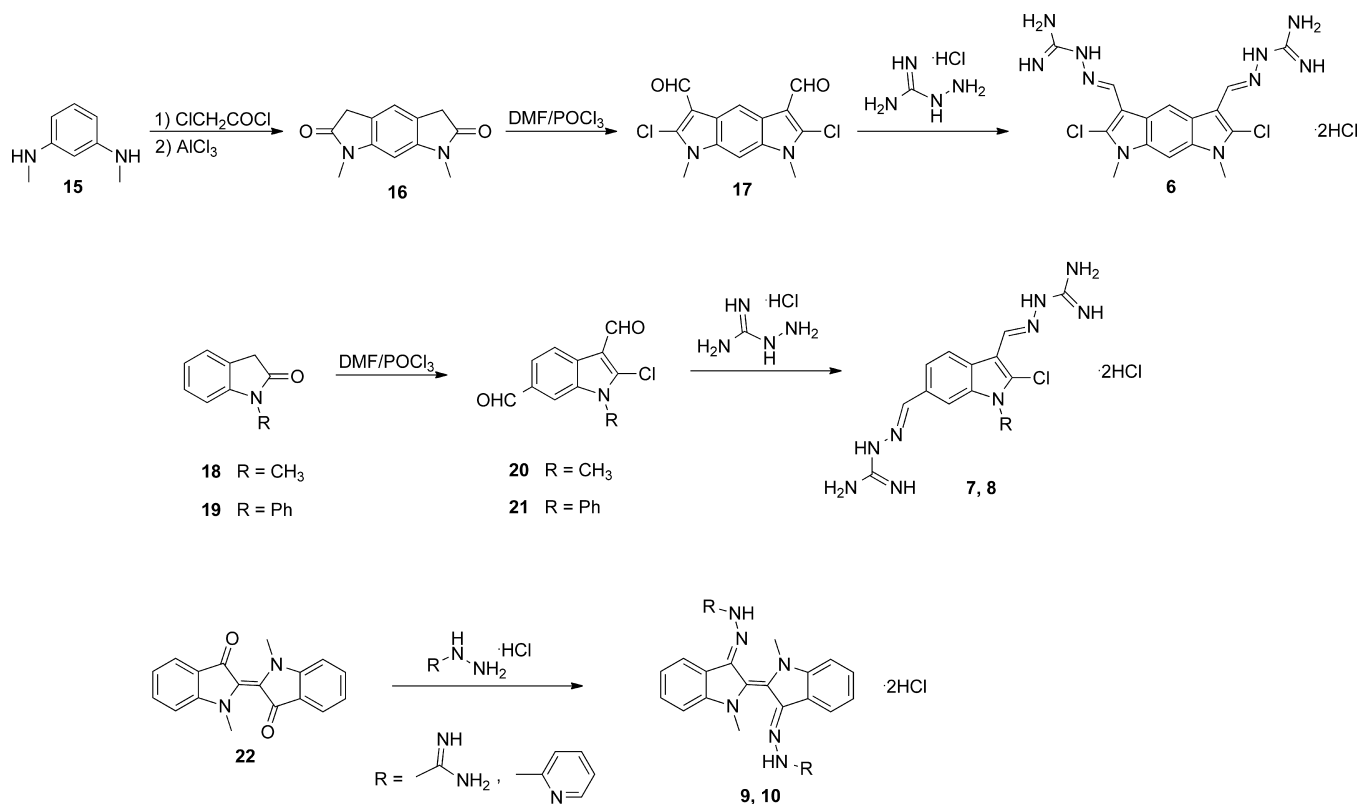
All the hydrazones have been obtained as hydrochlorides or hydrobromides, in which, as previously reported,¹⁸ the side chains are protonated and the positive charge is delocalized on the nitrogens of the guanyl or of the imidazoline groups. Indeed, ¹H NMR spectra of guanylhydrazone derivatives display at about 8 ppm a broad signal integrating to eight protons and exchanging

after D₂O addition, which may be ascribed to the two protonated guanyl groups. Similarly, the ¹H NMR spectra of compounds 2 and 3 show a broad signal at about 8.5 ppm integrating to two protons due to the protonated imidazoline nitrogens. Interestingly, owing to protonation and charge delocalization, the imidazoline ring become symmetric, therefore the two CH₂ give only one ¹H NMR signal: a singlet that integrates to four protons.

The new starting aldehydes 14b, 17, 20, and 21 were obtained by means of the Vilsmeier reaction on compounds 13b, 16, 18, and 19 (Scheme 1, 2) performed under reflux in order to promote bis-formylation. The bis-aldehydes 20 and 21 were isolated by column chromatography and the positions of the formyl groups were determined using NOE experiments. The aromatic signals of compound 20 were assigned by irradiating the methyl group at 3.93 ppm. NOE correlation was observed with the signal at 8.24 ppm, therefore this singlet was assigned to the proton at position 7 (ind-7). Afterward, we proceeded by irradiating the CHO signals. When the CHO proton at 10.03 ppm was irradiated, no NOE effect was observed, thus this signal was assigned to the CHO group at position 3. On the other hand, irradiation at 10.07 ppm provided a NOE with signals at 8.24 ppm (ind-7) and 7.83 ppm (ind-5), thus indicating that this CHO group is at position 6. Similar NOE experiments performed on compound 21 confirmed that, even in this case, the CHO groups are at positions 3 and 6.

Compound 13b (Scheme 1) was obtained in two steps by reacting pyrimidine-2,4-diamine 11 (commercially available) with 2-bromo-1-(thiophen-3-yl)ethan-1-one. As observed for the previously described compound 12a,¹⁹ the first cyclization involved the nitrogen at position 1 and the amino group at position 2 and led to the bicyclic compound 12b, whose structure was determined by two NOE experiments. The first experiment

Scheme 2. Synthetic Route for Hydrazones 6–10



confirmed the spatial closeness of the pyrimidine proton at position 6 and NH_2 , as suggested by NOE correlation between the doublet at 6.37 ppm (H-6 *pym*) and the signals at 8.36 ppm (the other pyrimidine signal) and 7.16 ppm (NH_2). The second NOE experiment, performed by irradiating on the doublet at 8.36 ppm, provided a correlation with the signals at 6.37 ppm (H-6 *pym*) and 7.74 ppm (*im*), thus confirming the structure of **12b**. An attempt to prepare **13b** directly from **11** with excess of 2-bromo-1-(thiophen-3-yl)ethan-1-one was unsuccessful: the intermediate **12b** is formed as the hydrobromide and it was necessary to prepare the corresponding base before treating it again with 2-bromo-1-(thiophen-3-yl)ethan-1-one.

Compound **16** (Scheme 2) was obtained in two steps treating *N,N'*-dimethylbenzene-1,3-diamine (**15**) with chloroacetyl chloride in order to achieve the acylation of the amino groups. Afterward, the intermediate thus obtained was cyclized by Friedel–Crafts alkylation with AlCl_3 . The other starting materials, 2-bromo-1-(thiophen-3-yl)ethan-1-one, **14a**, **15**, **18**, **19**, and **22**, were prepared according to the literature.^{17,20–24}

Circular Dichroism Studies. To evaluate the interaction of compounds **1–10** with G4 structures, several diverse G4-forming sequences able to form parallel, antiparallel, and hybrid G4 structures were selected for this study. It is known that human telomeric DNA G4s can adopt different topologies depending on the selected sequence and experimental conditions.²⁵ In this study, two human telomeric DNA sequences, namely Tel23 and Tel26, were used. They are known to form the so-called hybrid-1 and hybrid-2 folds, respectively. In addition, because several studies suggest the parallel G4 fold²⁶ as the prevalent one in the overcrowded solution conditions present inside a cell,²⁷ we also prepared a Tel23 sample at high DNA concentration conditions in order to promote the parallel G4 conformation (hereafter referred to as Tel23-*p*, see [Experimental Section](#)). Further, to

determine the affinity of the ligands toward other G4 structures, the two G4-forming sequences from the nuclease hypersensitive region of the *c-KIT* promoter (*c-kit1* and *c-kit2*) and one from the *c-MYC* promoter (*c-myc*) were also used.

The structures adopted by each G4 sample were verified by circular dichroism (CD) spectroscopy. Tel23-*p*, *c-kit1*, *c-kit2*, and *c-myc* displayed a positive band at 264 nm and a negative one around 240 nm in the CD spectrum (Supporting Information, [Figure S1](#)). These bands are characteristic of parallel-stranded G4 topologies.²⁸ On the other hand, Tel23 and Tel26 showed very similar CD spectra having a positive band at 289 nm with a shoulder at ca. 268 nm and a weak negative band at around 240 nm (Supporting Information, [Figure S1](#)). These data are consistent with the presence of hybrid structures as major conformations.

CD experiments were also performed to examine the potential of compounds **1–10** to alter the native folding topology of these G4s. Thus, DNA/ligand complexes were obtained by adding ligands (4 mol equiv) to the folded G4 structures. No significant variations of CD signal were observed for any of the structures analyzed here (Supporting Information, [Figure S1](#)), suggesting an overall preservation of their G4 architectures upon the addition of each ligand.

The structure of a duplex DNA (ds12), which is formed by a self-complementary 12-nt sequence, was also verified by CD spectroscopy because it has been used in subsequent experiments to evaluate the selectivity of the ligands for G4 over duplex DNA (see below). In the presence of K^+ , the CD spectrum of ds12 was characterized by a positive band at around 280 nm and a negative one at 250 nm, typical of values observed for duplex DNA (Supporting Information, [Figure S1](#)). These bands were not modified upon ligand addition.

Table 1. Ligand-Induced Thermal Stabilization of G4 and Duplex DNAs Measured by CD Melting Experiments

compd	ΔT_m (°C) ^a						
	c-kit1 ^b	c-kit2 ^c	c-myc ^c	Tel23-p ^b	Tel23 ^b	Tel26 ^b	ds12 ^b
1 ^d	>15.0	>20.0	>20.0	>20.0	-10.0	-4.5	0.1
2	1.6	>20.0	>15.0	10.8	-14.0	-8.9	1.3
3	2.7	9.5	>20.0	9.1	-2.4	-3.0	-0.8
4	0.4	0.0	-0.2	1.3	-1.0	-1.6	0.9
5	>15.0	>20.0	>20.0	>20.0	-4.9	-6.5	3.5
6	>15.0	>15.0	>20.0	>20.0	3.0	1.3	ND ^e
7	>15.0	>20.0	>20.0	>20.0	6.0	5.8	ND ^e
8	>15.0	>15.0	>20.0	>20.0	5.6	5.1	ND ^e
9	0.9	8.2	11.0	7.7	-0.5	0.2	-0.8
10	1.5	2.0	1.6	4.0	-2.5	-2.0	-0.7

^a ΔT_m represents the difference in melting temperature [$\Delta T_m = T_{m(\text{DNA}+\text{4 ligand equiv})} - T_{m(\text{DNA})}$]. The T_m values of DNAs alone are c-kit1 = 69.8 ± 0.5 °C, c-kit2 = 62.5 ± 0.5 °C, c-myc = 72.9 ± 0.5 °C, Tel23-p = 68.5 ± 0.5 °C, Tel23 = 65.2 ± 0.5 °C, Tel26 = 59.5 ± 0.5 °C, ds12 = 71.3 ± 0.5 °C. All experiments were duplicated and the values reported are average of two measurements. ^b100 mM KCl buffer. ^c20 mM KCl buffer. ^dLead compound. ^eThese compounds significantly increase the thermal stability of ds12 (see Supporting Information, Figure S2), however, ΔT_m values cannot be accurately determined.

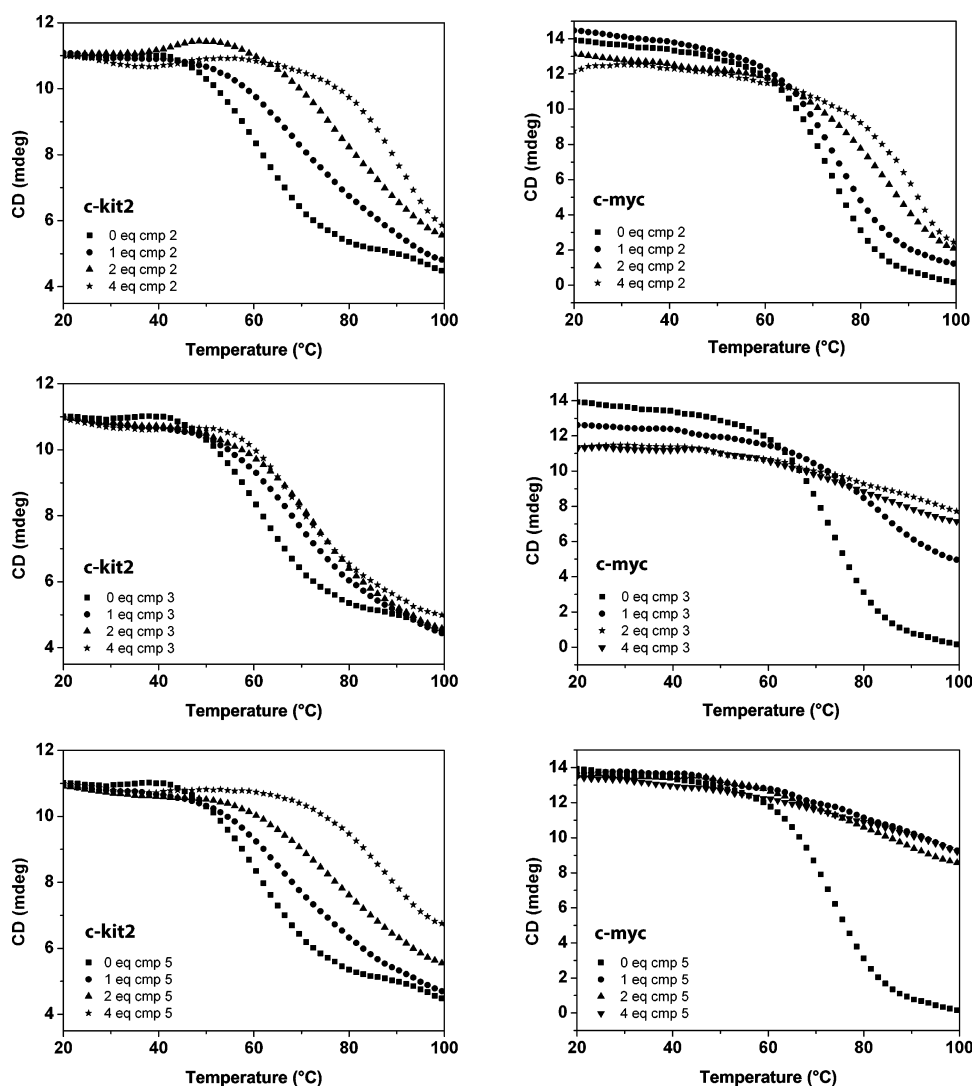


Figure 2. CD melting experiments of c-kit2 (left column) and c-myc (right column) G4s with increasing amounts of compounds 2, 3, and 5.

CD Melting Studies. The stabilizing properties of the compounds were evaluated by CD-melting experiments measuring the ligand-induced change in the melting temperature (ΔT_m) of G4 and duplex structures. CD-melting curves of DNAs

in the absence and presence of each ligand were obtained by following the variations of the intensities of CD signals at the wavelength of 264, 290, and 280 nm for parallel G4s, antiparallel G4s, and duplex, respectively (Supporting Information, Figure

S2). For the first CD-melting experiments, a 4:1 ligand/DNA ratio was used so that all possible G4 binding sites were available. Results of these experiments are reported in Table 1. On the basis of these data, ligands have been ranked in the following categories: (i) compounds that exhibit poor G4-stabilizing properties toward any G4 ($\Delta T_m < 5^\circ\text{C}$) (compounds 4 and 10), (ii) compounds showing medium G4-stabilizing effect ($\Delta T_m = 5\text{--}15^\circ\text{C}$) (compound 9), and (iii) potent G4 stabilizers ($\Delta T_m > 15^\circ\text{C}$) (compounds 2, 3, 5–8). As with the lead compound 1, most of the compounds investigated here showed high selectivity for G4 over duplex DNA. However, compounds 6, 7, and 8, having a different core from compound 1, were the only ligands to significantly increase duplex stability and therefore were not considered to be selective. Interestingly, compounds 2, 3, and 5 showed some preference for parallel G4s over antiparallel ones. In particular, compound 3 stabilized to a large extent only the c-myc G4 ($\Delta T_m > 20^\circ\text{C}$). Because we were mainly interested in finding G4 ligands specific for one topology or, even better, for one structure, further studies were only performed on compounds 2, 3, and 5. Therefore, we carried out CD melting experiments to investigate the effect of increasing concentrations of 2, 3, and 5 on the stability of both c-kit2 and c-myc G4s (Figure 2). Three different [ligand]/[DNA] ratio were examined (1:1, 2:1, and 4:1). In the case of c-kit2, we observed a dose-dependent increase of thermal stability for all three compounds although with some differences between them. In particular, compounds 2 and 5 increased the T_m of the G4 with every addition of ligand, while 3, which had shown a modest increase in c-kit2 stability, did not induce any further increase of T_m beyond the 2:1 [ligand]/[DNA] ratio. In the case of c-myc, increasing thermal stability of the G4 with compound 3 also ceased at a 2:1 ratio but after a considerable increase of T_m . On the other hand, concentration dependency was observed for 2 (although slight compared to c-kit2), while the high G4-stabilizing effect of 5 on c-myc did not enable us to assess any dose-dependent effects.

Overall, these results indicate that six out of nine synthesized ligands are potent G4 stabilizers. The diimidazo[1,2-*a*:1,2-*c*]pyrimidine derivatives having a positive charge on the side chain(s) in the used experimental conditions (i.e., 2, 3, and 5) showed high affinity toward some of the investigated G4 structures. Conversely, compound 4, whose side chain protonation is less favorable at pH 7.0, showed generally poor G4-stabilizing properties. In any case, beyond the overall charge of these ligands, their duplex-binding activity is negligible. Concerning the compounds with a different core, the positively charged ones (6, 7, and 8) showed affinity for some of the G4s investigated as well as for duplex DNA. On the other hand, compounds 9 and 10, whose side chain protonation is less likely to occur at pH 7.0, showed medium to low G4-stabilizing properties and negligible effects on duplex. In summary, these results indicate that the overall charge of compounds is important in DNA binding but, at the same time, suggest that the charge is necessary but not sufficient to ensure potent and selective interaction with a specific DNA structure.

In addition, CD melting data suggest that two molecules of 3 could bind and stabilize the G4 target (the addition of more than two equivalents of ligand does not lead to a further enhancement of G4 stability), possibly interacting in an end-stacking mode at the terminal G-tetrads of the G4 (see also NMR and Molecular Docking section), a common binding mode exhibited by the majority of effective G4-targeting ligands.

It has to be noted that a significant destabilizing effect on Tel23 and Tel26 G4s occurs with some ligands. In particular, this

phenomenon occurs with compounds 1, 2, and 5, namely the bis-hydrazine derivatives of diimidazo[1,2-*a*:1,2-*c*]pyrimidine. The mechanism of destabilization of these G4 structures by such derivatives is unclear and would need further study. However, we hypothesize that it may involve nonspecific interactions of compounds with the DNA molecule that might favor single-stranded DNA, resulting in a shift of folded-unfolded equilibrium toward the unfolded form during the melting experiment, which, in turn, may result in a decrease of T_m .

FRET Melting Studies. The FRET methodology was also used to evaluate the investigated compounds and to estimate their G4 vs duplex selectivity.²⁹ The compounds were first screened with the derivatized human telomeric G4 sequence F21T and a duplex DNA sequence (T-loop) by using various concentrations of ligands. Similarly to the lead 1, compounds 2, 3, and 5–8 showed potent G4 stabilizing capabilities, while compounds 4, 9, and 10 produced no significant change in G4 melting temperature (Supporting Information, Figure S3). Among the effective G4 stabilizers, compounds 6, 7, and 8 showed a significant degree of duplex stabilizing ability, confirming once again their low selectivity for G4 over duplex DNA. On the other hand, FRET results also confirm that compounds 2, 3, and 5 are effective and selective stabilizers of the G4 structures.

However, the results of these FRET melting experiments cannot be directly compared with those obtained from CD melting studies because of differences in DNA sequences and in experimental conditions. Therefore, to more fully compare the results of CD melting experiments obtained for compounds 2, 3, and 5, additional FRET melting experiments were performed on these ligands under the same buffer conditions used for CD, using as targets the labeled G4-forming oligonucleotides from the c-KIT (Fckit2T) and c-MYC (FcmcytT) promoter regions and the telomeric G4-forming sequence annealed at high concentration to promote the parallel conformation (F21T-p) (Supporting Information, Figure S4). The results of FRET melting experiments, shown in Table 2, are in good qualitative

Table 2. Ligand-Induced Thermal Stabilization of Labeled G4-Forming Sequences Evaluated by FRET Melting Analyses

compd	ΔT_m ($^\circ\text{C}$) ^a		
	Fckit2T ^b	FcmcytT ^b	F21T-p ^c
2	>20.0	>15.0	>15.0
3	≈15.0	>20.0	7.0
5	>20.0	>20.0	>20.0

^a ΔT_m represents the difference in melting temperature [$\Delta T_m = T_{m(\text{G4}+\text{4 ligand equiv})} - T_{m(\text{G4})}$]. T_m values of G4s alone: Fckit2T = $66.3 \pm 1.0^\circ\text{C}$, FcmcytT = $68.6 \pm 1.0^\circ\text{C}$, F21T-p = $67.1 \pm 1.0^\circ\text{C}$. All experiments were duplicated and the values reported are average of two measurements. ^b20 mM KCl buffer. ^c100 mM KCl buffer.

agreement with those obtained from CD melting studies. The FRET ranking order reflects that compared to the CD melting data; the fact that the ΔT_m values themselves are somewhat different is likely to be a consequence of different techniques and does not have particular significance.

CD Titration Experiments. CD has been also used to investigate the ability of the best ligands (2, 3, and 5) to induce G4 formation from unfolded single-stranded DNAs by performing CD titration experiments in the absence of metal ions. In particular, the sequences c-kit1, c-kit2, Tel23, and c-myc in 50 mM Tris buffer were investigated. The CD spectra of c-kit1 and

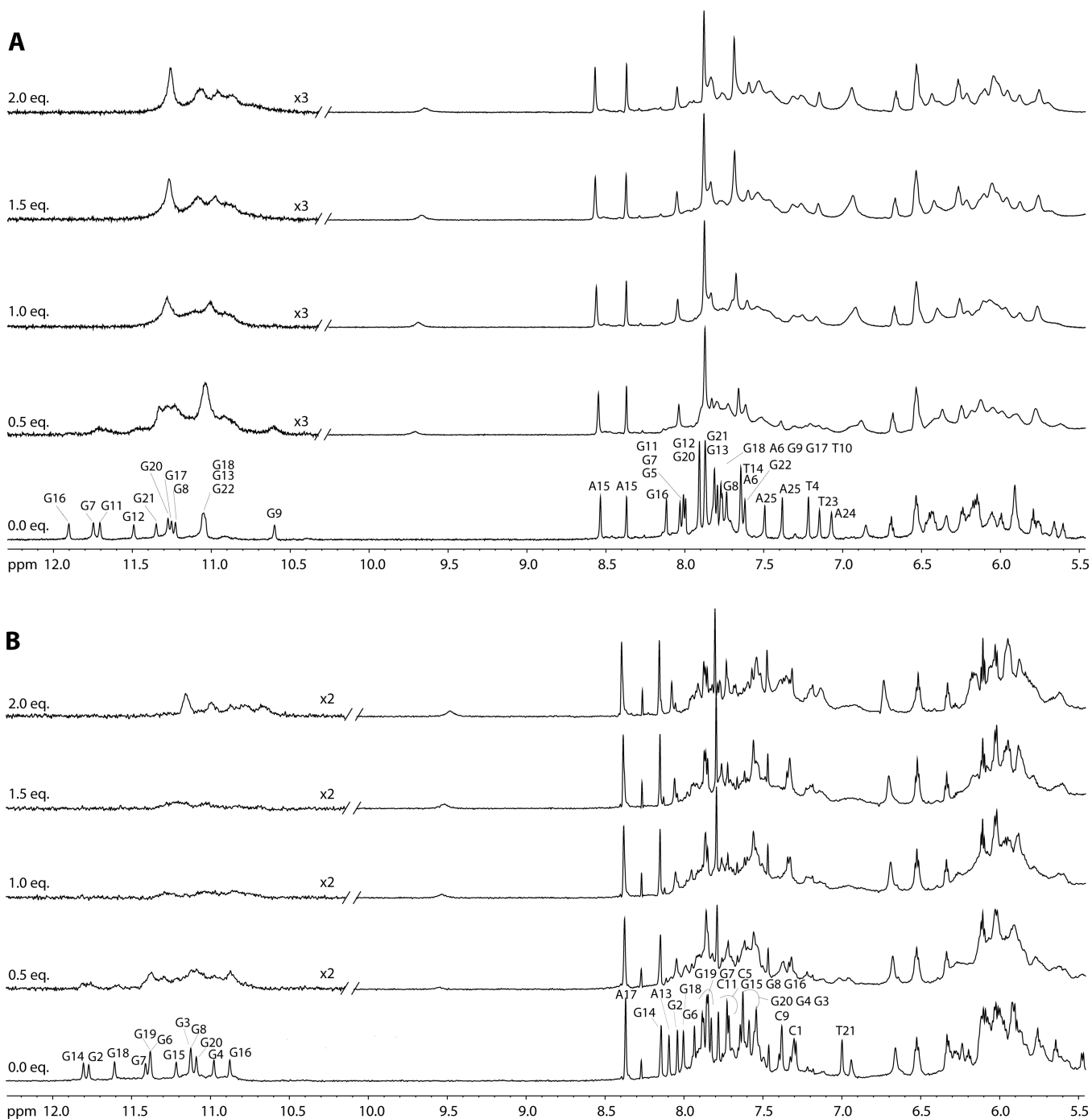


Figure 3. Imino and aromatic proton regions of (A) *c-MYC* promoter G4 (Myc22) in 25 mM phosphate, 70 mM KCl solution (pH 7.0), and (B) *c-KIT* promoter G4 (*c-kit*_{T12/T21}) in 5 mM phosphate, 20 mM KCl solution (pH 6.8) titrated with compound **3**.

c-kit2 sequences in the absence of monovalent metal cations showed a positive band at around 262 nm and a negative one around 240 nm that, even if small, are indicative of the presence of a small percentage of a parallel G4 structure in solution (Supporting Information, Figure S5). Upon addition of ligands to G4s, a dose-dependent increase of the CD bands was observed in some case, suggesting the induction of the parallel G4 structure formation. In particular, as far as *c-kit2* is concerned, compounds **2** and **5** caused a major increase of ellipticity compared to **3**. Similarly, the addition of **2** and **5** to the *c-kit1* sequence also caused a dose-dependent increase in CD signals (more intense for **5**), whereas the addition of **3** did not affect the

CD profile, thus indicating that **3** is unable to induce *c-kit1* G4 formation.

The *c-myc* G4 DNA sequence exists in a parallel topology, in the absence of any added monovalent metal cation, showing a positive band at 263 nm and a negative one around 240 nm. The addition of ligands **2**, **3**, and **5** to *c-myc* G4 did not perturb the parallel G4 structure already formed in Tris buffer. On the other hand, the addition of compounds **2**, **3**, and **5** to the Tel23 sequence did not lead to any increase in CD ellipticity, thus suggesting that none of them is able to induce any telomeric G4 structures.

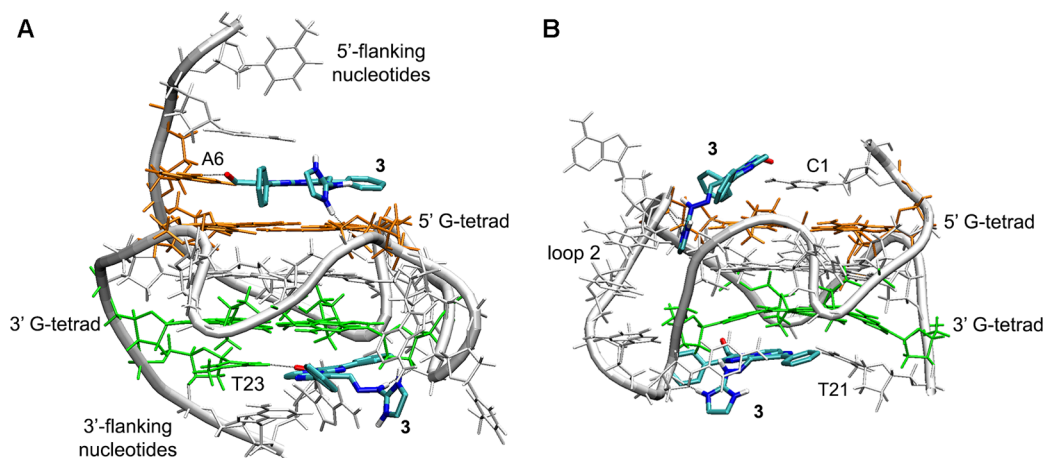


Figure 4. Predicted binding poses upon docking of compound 3 onto the (A) c-myc and (B) c-kit2 G4s. Ligand molecules are represented as sticks. The phosphate backbone and nucleosides of G4s are represented in tube and stick representations, respectively. Black dashed lines indicate hydrogen bond interactions.

In summary, these results suggest that **2** and **5** are able to induce the c-kit1 and c-kit2 sequences to form G4 structures in the absence of metal cations. On the other hand, compound **3**, which is an unexceptional binder of c-kit G4s, does not induce the formation of any G4 structure. Finally, none of the ligands showed an inducing effect on the telomeric G4-forming sequence.

Nuclear Magnetic Resonance Studies. NMR spectroscopy was employed in order to obtain structural information about the binding mode of compound **3**. As stated above, this compound is a potent and specific stabilizer of c-myc G4. Because the wild-type G4 sequence of c-myc is not suitable for study by NMR (its ^1H NMR spectrum is affected by the presence of minor conformations), the well-characterized modified sequence (Myc22) containing two G-to-T mutations at position 14 and 23 was used for the NMR titration experiments. The spectral regions of the imino and aromatic protons in the absence and presence of increasing amounts of compound **3** are shown in Figure 3A. According to the literature, under the experimental conditions used, the Myc22 DNA sequence forms a single G4 conformation characterized by 12 well-resolved imino protons peaks, corresponding to the 12 guanines involved in the three G-tetrad planes.³⁰ Upon addition of compound **3** to the Myc22 solution, both imino and aromatic protons became broad even at higher drug equivalence, indicating a dynamic binding process of **3** to Myc22 with a medium exchange rate on the NMR time scale. The titration was virtually complete at a ligand/DNA ratio of 2:1. Major chemical shift changes were observed for the imino protons belonging to the 5' G-tetrad (G7–G11–G16–G20) as well as for G9 of the 3' G-tetrad (G9–G13–G18–G22), while the remaining imino protons of G13, G18, and G22 at the 3'-end tetrad and those of middle tetrad (G8–G12–G17–G21) were less affected. The origin of this different behavior could be due to inherent structural features associated with the 3'- and 5'-faces as well as the flanking sequences. The 5' side is more accessible for ligand stacking and the only 5'-flanking residue particularly affected by the binding turned out to be the T4, as shown by the corresponding aromatic proton shift. By contrast, the 3'-face is less accessible. A25 from this side folds back to form a base pair with T23 to cover the external G-tetrad, while A24 stays above the T23:A25 base pair. However, these latter flanking residues are not able to completely prevent the binding. Upon addition of **3**, the aromatic protons of T23, A24, and A25 were strongly

perturbed, suggesting that the ligand can be positioned between the four guanines of the 3'-tetrad and the 3'-flanking bases. The stability of the 3'-end bound ligand complex with Myc22 could be related to the possibility of specific H-bond interactions between the imidazole group of **3** and the 3'-flanking bases (as discussed in the docking section).

For comparison, we have also examined the binding behavior of **3** to the c-kit2 G4. Because the c-kit2 G4 wild-type sequence gives poorly resolved imino proton NMR spectra, we used the modified sequence c-kit2_{T12/T21} (in which G12 and G21 are replaced by two T residues), which is known to fold into a monomeric parallel G4 structure with well-resolved imino proton signals.^{31,32} Figure 3B shows the spectral regions of the imino and aromatic protons of c-kit2_{T12/T21} in the absence and presence of increasing amounts of compound **3**. Interestingly, the imino signals belonging to the G-tetrad on the 5'-side (G2, G6, G14, and G18) were equally affected with respect to those belonging to the 3'-terminal G-tetrad, as well as the aromatic protons of C1 and A13 on the 5'-face and those of C9 and T21 on the 3'-face. These results clearly suggest an involvement of both external G-tetrads in binding, despite the potential pairing between residues C1 and A13 on the 5'-side.³¹ Thus, we hypothesize that **3** is able to interact with both the 5' and 3' G-tetrads, possibly by displacing the C1 and A13 base pairing at the former.

Overall, these findings suggest that compound **3** binds to the 5'- and 3'-terminal G-tetrads of both Myc22 and c-kit2_{T12/T21} G4s. However, some differences in the binding to the two G4s were observed that could explain the difference in the thermal stabilization of the two G4s upon binding observed by CD analysis. Particularly, we speculate that this different binding behavior is due to the possibility of **3** forming specific H-bond interactions with Myc22 (see Molecular Docking section), thus ensuring that compound **3** binds more effectively to the c-myc G4 rather than to the c-kit2 G4.

Molecular Docking. Molecular docking calculations were performed to provide a possible molecular basis for the G4 binding of **3**. We docked the ligand to the three-dimensional structures of c-kit2 and c-myc G4s determined by solution state NMR spectroscopy.^{32,33} Taking into account the experimentally suggested binding stoichiometry, 2:1 (ligand/DNA) complex models were built with the G4 targets. Both 5'- and 3'-terminal G-tetrads were considered as potential binding sites for the

Table 3. Cytotoxic Activity of Selected Hydrazone Derivatives in Human U2OS and HCT116 Cells^a

compd	IC ₅₀ (μM) 1 h		IC ₅₀ (μM) 24 h	
	U2OS	HCT116	U2OS	HCT116
1	101.0 ± 1.63	35.07 ± 1.06	15.97 ± 1.41	5.12 ± 1.05
3	5.73 ± 1.48	4.18 ± 1.07	0.845 ± 0.001	0.407 ± 0.106
14a	45.23 ± 2.05	16.79 ± 1.06	9.16 ± 1.44	4.06 ± 1.07

^aThe concentrations killing 50% of cells (IC₅₀) are means ± SE of two independent experiments performed in triplicates.

ligand and were therefore examined in the docking procedures. For each binding site, the most favorable complex was selected from the docked structures on the basis of the calculated binding energies. The low-energy binding poses for the ligand are shown in Figure 4A,B.

The shape of **3** effectively fills much of the available space in the plane of the two G-tetrads in the c-myc G4. The tricyclic diimidazo[1,2-*a*:1,2-*c*]pyrimidine moiety is stacked onto the planes at a 3.4 Å separation at both 5'- and 3'-faces and overlaps part of two guanine bases (G11 and G16 at the 5' G-tetrad; G9 and G13 at the 3' G-tetrad), making extensive π -stacking interactions (Figure 4A). In both cases, one of the two attached phenyl rings overlap part of a guanine of the tetrad, contributing to the stabilization, while the other one is twisted out of the diimidazo[1,2-*a*:1,2-*c*]pyrimidine plane. The 2-hydrazino-2-imidazoline groups of the two ligand molecules make close hydrogen bonds with the oxygen atoms of the backbone phosphate groups of the G4. It is noteworthy that the formyl group of **3** is, in both cases, hydrogen bonded to the first flanking base (A6 and T23 at 5'- and 3'-end, respectively) that stacks over the G-tetrad, thus forming a plane capping the 5' and 3' external G-tetrads of the c-myc G4.

In the case of c-kit2 G4, different binding poses were observed for the ligand on the two external G-tetrad surfaces (Figure 4B). Indeed, the molecular model of **3** bound to the c-kit2 G4 shows that at the 3'-end, the planar aromatic scaffold of the ligand makes π -stacking interactions with the 3'G-tetrad, while the 2-hydrazino-2-imidazoline group is positioned to participate in a hydrogen bond with a phosphate group. However, by contrast with the c-myc G4, compound **3** does not form any additional interaction to further stabilize the complex. On the other hand, the docking calculations at the 5'-face of the c-kit2 G4 structure reveal that to make hydrogen-bonding/electrostatic contacts between the charged 2-hydrazino-2-imidazoline group and the backbone phosphate groups of the G4, the side chain of compound **3** needs to be located in the deep groove formed by the long loop 2 of the G4. As a consequence, the diimidazo[1,2-*a*:1,2-*c*]pyrimidine core of **3** results in not being parallel to the plane of the terminal G-tetrad. This is probably due to the 5'-3' strand polarity generated by the propeller loop topology of this G4, which does not expose the phosphodiester backbone on this side, thus making only the 3'-end more favorable to interact with ligands with positively charged side chains.

Overall, the docking results reveal that **3** has the appropriate shape and electronic structure to form specific interactions. Interestingly, the binding of **3** could involve the first flanking bases to form a plane covering the external G-tetrads and contributing greatly to the G4 stabilization. This could also explain the selectivity of **3** in increasing the thermal stability of the c-myc G4.

Cytotoxic Activity and Stabilization of G4 Structures by Selected Hydrazone Derivatives in Human Cancer Cells. As the studied hydrazone derivatives showed distinct specificity for different G4 structures, we next determined the

effects of selected compounds at cellular levels. We selected compound **3** as it was the most selective agent in vitro, and we compared its effects with those of **1** (the lead) and **14a** (Scheme 1, used as a negative control because it is unable to bind G4s in vitro).¹⁶ First, we determined the cell killing activity of the three selected molecules with the MTT assay in two human cancer cell lines: the osteosarcoma U2OS and colon cancer HCT116 lines. In this assay, after exposure to the compounds, cells were allowed to recover in drug-free medium for 24–48 h before adding MTT to measure surviving cells (see Experimental Section). HCT116 cells were more sensitive than U2OS cells to the tested compounds, however, both cell lines were more sensitive to compound **3** than other agents (Table 3). In particular, compound **3** showed high cell killing activity in U2OS cells after 24 h of treatment (IC₅₀ = 0.845 μM). In addition, **3** was 20- and 12-fold more cytotoxic than compounds **1** (IC₅₀ = 15.97 μM) and **14a** (IC₅₀ = 9.16 μM), respectively. Similarly, **3** was more cytotoxic than compounds **1** and **14a** after 1 h of cell treatment (Table 3). The sensitivity pattern was very similar in HCT116 cells, with compound **3** showing highest cell killing activity as compared with compounds **1** and **14a** both after 24 and 1 h of treatment (Table 3). Therefore, the results established that compound **3** is more cytotoxic than the other tested compounds, at least in the two cell lines used here.

Next, to show whether **3** and **1** stabilized G4s in living cells, we visualized nuclear G4 structures in human U2OS cells by immunofluorescence microscopy using the BG4 antibody, a known specific antibody against G4s.³⁴ U2OS cells were treated for 24 h with compound concentrations very close to the established IC₅₀ values (Table 3). Immediately after treatments with tested compounds, cell death was minimal, as determined by cell FACS analyses and microscopy visualization (data not shown). After incubation with BG4 antibody, the fluorescence signal was generated by a final incubation with a fluorochrome-labeled antibody and measured and analyzed with ImageJ as described in detail in Experimental Section. BG4 showed nuclear staining in control, untreated cells indicating that G4 structures were present at discrete chromatin sites in the nucleus (Figure 5A), in agreement with previous studies.³⁴ Fluorescence spots were of different sizes, and larger spots apparently consisted of a few smaller spots. The fluorescence patterns were similar in cells treated with the three selected hydrazone derivatives. However, the compound effects were very different because compounds **3** and **1**, but not **14a**, markedly increased the number, size, and fluorescence intensity of nuclear spots, thus indicating that they could stabilize G4 structures in the nucleus of living cells. We then measured the total fluorescence signal per cell for each sample, and the results indicated that **3** was the most potent as an increase of G4 signal was detected even at 0.2 μM (a dose 4-fold lower than IC₅₀), and further increased at 1 μM, of **3**. Compound **1** was able to increase the nuclear G4 signal at 10 μM at comparable levels as **3** at 0.2 μM (Figure 5A). In contrast, **14a** did not show any increase in fluorescence intensity at the tested concentrations (Figure 5B). Statistical analyses showed that a

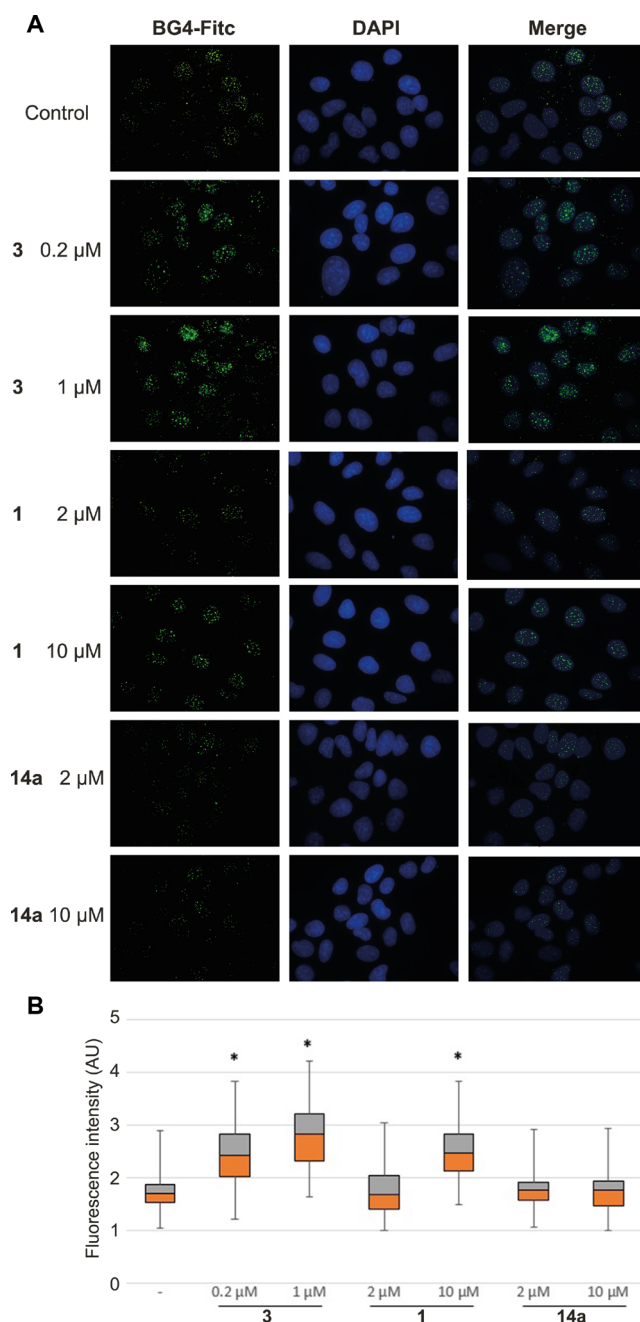


Figure 5. (A) Immunofluorescence showing BG4 foci (green) in human osteosarcoma U2OS cell nuclei (blue). Cells were treated with G4 ligands for 24 h and then fixed. G4s and DNA were stained with BG4 antibody and DAPI, respectively. (B) Box Plot analysis of G4 signal in the nucleus of control and treated cells. At least two experiments were considered, and 150 nuclei were counted per sample. Statistical significance was determined with the *t* test, and * indicates a *p* value less than 10^{-32} .

significant increase of G4 after $0.2 \mu\text{M}$ ($p = 2.05 \times 10^{-32}$) and $1 \mu\text{M}$ ($p = 1.82 \times 10^{-55}$) of compound 3 and $10 \mu\text{M}$ ($p = 2.44 \times 10^{-38}$) of 1. Taken together, these data demonstrate that hydrazone derivatives can stabilize G4s in cultured human cancer cells and that 3 is the most potent agent among the studied derivatives. Although we detected an association of cytotoxic potency with G4 stabilization in living cells among the tested compounds, further investigations at the cellular level are needed

to establish a clear link between G4 stabilization and biological activity of these hydrazone derivatives.

CONCLUSIONS

Starting from the lead compound 1, a series of analogues with structural modifications of the substituents and the core have been designed, synthesized, and evaluated as effective and selective G4 binding ligands. The present study extends our previous work in which a hydrazone derivative was identified as a novel G4-binding motif and provides the basis for the design of other analogues with improved binding properties. The results from CD assays showed that six out of nine synthesized ligands are potent G4 stabilizers that retain the pre-existing architecture of the target G4s upon interaction. However, the compounds having a different core from that of compound 1 (6, 7, and 8) were also shown to increase the stability of duplex DNA and therefore they are not selective for G4s. On the other hand, compounds 2, 3, and 5 displayed a marked preference for binding to parallel G4s over duplex and antiparallel G4s. Interestingly, although not able to induce the formation of any G4 structure, compound 3 is found to be the most selective of the series. Therefore, it was investigated in detail in terms of binding mode. NMR experiments suggested an end-stacking binding mode at the terminal G-tetrads of the G4 with a 2:1 drug/DNA binding stoichiometry. Molecular docking results provided further insight into the recognition between 3 and the G4 DNA, highlighting the possible key structural elements involved in the interaction. Interestingly, compound 3, which is the only monohydrazone of the series, appears to have the appropriate structure to stack on the planar G-tetrad surfaces of c-myc G4 with the tricyclic diimidazo[1,2-*a*:1,2-*c*]pyrimidine core and, at the same time, to make (i) close hydrogen-bonding/electrostatic contact between the charged 2-hydrazino-2-imidazoline group and the oxygen atoms of the backbone phosphate of G4 and (ii) a hydrogen bond interaction between the formyl group at position 3 and the hydrogen-bond donor groups of the first flanking base that stacks on the G-tetrad. Similar intermolecular interactions were also found in the structure of the complex between the quindoline molecule and c-myc G4³³ and could be the molecular determinants for specific G4 recognition. It is noteworthy that decreasing the number of positively charged side chains leads to a significant benefit in terms of selectivity without affecting the G4 stabilizing property of this derivative. In line with the *in vitro* assays, biological experiments demonstrated that hydrazone derivatives can effectively trap G4 structures in the nuclei of cells. Remarkably, derivative 3 is more effective than the lead compound 1 in inhibiting human U2OS and HCT116 cancer cell growth as well as in stabilizing G4s in osteosarcoma cells. Altogether, the reported results indicate that this molecular scaffold could actually form the basis for the development of highly selective drug-like G4 ligands with superior biological activity to accomplish the desired result of achieving new and effective anticancer drug candidates.

EXPERIMENTAL SECTION

Chemistry. All the compounds prepared were at least 98% pure as determined by combustion analysis. Melting points are uncorrected. TLC was performed on Bakerflex plates (Silica gel IB2-F) and column chromatography on Kieselgel 60 (Merck): the eluent was a mixture of petroleum ether/acetone in various proportions. IR spectra were recorded in nujol on a Nicolet Avatar 320 ESP; ν_{max} is expressed in cm^{-1} . ¹H NMR spectra were recorded on a Varian MR 400 MHz (ATB PFG probe); the chemical shift (referenced to solvent signal) is expressed in δ

(ppm) and J in Hz (abbreviations: ar = aromatic, im = imidazole, ph = phenyl, py = pyridine, pym = pyrimidine, th = thiazole, imi = imidazoline). The IR, ^1H NMR, and ^{13}C NMR spectra of the compounds are in agreement with the assigned structures. 2,4-Diaminopyrimidine **11** is commercially available. The following compounds were prepared according to the literature: 2-bromo-1-(thiophen-3-yl)ethan-1-one,²⁰ **14a**,¹⁷ **15**,²¹ **18**,²² **19**,²³ and **22**.²⁴

Synthesis of 2-(Thiophen-3-yl)imidazo[1,2- α]pyrimidin-7-amine (12b) (Scheme 1). 2,4-Diaminopyrimidine **11** (10 mmol) was dissolved in 40 mL of acetone and treated with 2-bromo-1-(thiophen-3-yl)ethan-1-one (15 mmol). The reaction mixture was refluxed for 3 h, and the resulting precipitate (**12b** hydrobromide) was collected by filtration. The free base was obtained by treatment with 15% NH_4OH . The resulting precipitate was collected by filtration and crystallized from ethanol. Yield 46%; mp 285–290 °C. IR (cm^{-1}): 3148, 1667, 1087, 846, 720. ^1H NMR ($\text{DMSO-}d_6$): 6.37 (1H, d, pym, $J = 7.6$), 7.16 (2H, s, NH_2), 7.48 (1H, dd, th, $J = 5.2$, $J = 1.2$), 7.57 (1H, dd, th, $J = 5.2$, $J = 2.6$), 7.74 (1H, s, im), 7.80 (1H, dd, th, $J = 2.6$, $J = 1.2$), 8.36 (1H, d, pym, $J = 7.6$). Anal. Calcd for $\text{C}_{10}\text{H}_8\text{N}_4\text{S}$ (MW 216.26): C, 68.90; H, 5.14; N, 4.46. Found: C, 69.04; H, 5.02; N, 4.78.

Synthesis of 2,8-Di(thiophen-3-yl)diimidazo[1,2- α' ,2'- c]pyrimidine (13b) (Scheme 1). Compound **12b** (5 mmol) was dissolved in ethanol (40 mL), treated with excess of 2-bromo-1-(thiophen-3-yl)ethan-1-one (20 mmol), and kept under reflux for 7 h. The resulting compound **13b** hydrobromide was collected by filtration and treated with 15% NH_4OH in order to prepare the free base which was collected by filtration and crystallized from ethanol. Yield 12%, mp 245–249 °C dec. IR (cm^{-1}): 1635, 1292, 1193, 852, 724. ^1H NMR ($\text{DMSO-}d_6$): 7.16 (1H, d, pym, $J = 7.6$), 7.56 (1H, dd, th, $J = 5.0$, $J = 1.2$), 7.63 (1H, dd, th, $J = 5.0$, $J = 3.0$), 7.65 (1H, dd, th, $J = 5.0$, $J = 3.0$), 7.75 (1H, dd, th, $J = 5.0$, $J = 1.2$), 7.83 (1H, dd, th, $J = 3.0$, $J = 1.2$), 7.99 (1H, dd, th, $J = 3.0$, $J = 1.2$), 8.09 (1H, s, im), 8.28 (1H, d, pym, $J = 7.6$), 8.61 (1H, s, im). Anal. Calcd for $\text{C}_{16}\text{H}_{10}\text{N}_4\text{S}_2$ (MW 322.41): C, 59.61; H, 3.13; N, 17.38. Found: C, 60.03; H, 3.67; N, 17.99.

Synthesis of 1,7-Dimethyl-5,7-dihydropyrrolo[3,2- f]indole-2,6(1H,3H)-dione (16) (Scheme 2). Compound **15** (10 mmol) was dissolved in toluene (20 mL) and treated with a solution of NaOH (20 mmol) in H_2O (4 mL). Chloroacetyl chloride (20 mmol) was added dropwise at 0–15 °C. The reaction mixture was kept for 3 h at room temperature. The aqueous phase was removed, and the organic layer was washed with HCl 2N and with H_2O . Toluene was removed under reduced pressure, and the resulting residue was treated with AlCl_3 (10 g) at 160 °C for 6 h. The mixture was then poured onto ice. The crude compound **16** was collected by filtration and crystallized from ethanol. Yield 45%; mp 275 °C dec. IR (cm^{-1}): 1705, 1625, 1276, 1025, 641. ^1H NMR ($\text{DMSO-}d_6$): 3.14 (6H, s, CH_3), 3.48 (4H, s, CH_2), 6.76 (1H, s, ar), 7.13 (1H, s, ar). Anal. Calcd for $\text{C}_{12}\text{H}_{12}\text{N}_2\text{O}_2$ (MW 216.24): C, 66.65; H, 5.59; N, 12.96. Found: C, 68.03; H, 6.05; N, 13.11.

General Procedure for the Synthesis of the Aldehydes 14b (Scheme 1) and 17, 20, and 21 (Scheme 2). The Vilsmeier reagent was prepared at 0–5 °C by dropping POCl_3 (54 mmol) into a stirred solution of DMF (65 mmol) in CHCl_3 (5 mL). The appropriate starting compound (**13b**, **16**, **18**, and **19**, 5 mmol) was suspended in CHCl_3 (20 mL). The mixture thus obtained was dropped into the Vilsmeier reagent while maintaining stirring and cooling. The reaction mixture was kept for 3 h at room temperature and under reflux for 2–40 h (according to a TLC test). Chloroform was removed under reduced pressure, the resulting oil was poured onto ice, and the precipitate thus obtained was collected by filtration. The crude aldehydes **14b** and **17** were crystallized from ethanol. The aldehydes **20** and **21** were isolated by column chromatography (petroleum ether/acetone, 70/30).

2,8-Di(thiophen-3-yl)diimidazo[1,2- α' ,2'- c]pyrimidine-3,9-dicarbaldehyde (14b). Yield 99%; mp 242–247 °C dec. IR (cm^{-1}): 1662, 1627, 1319, 866, 794. ^1H NMR ($\text{DMSO-}d_6$): 7.57 (1H, d, pym, $J = 7.8$), 7.68 (1H, dd, th, $J = 5.2$, $J = 3.0$), 7.80 (2H, m, th), 7.99 (1H, dd, th, $J = 5.2$, $J = 1.2$), 8.44 (1H, dd, $J = 2.4$, $J = 1.2$), 8.90 (1H, dd, th, $J = 3.0$, $J = 1.2$), 9.21 (1H, d, pym, $J = 7.8$), 10.25 (1H, s, CHO), 11.38 (1H, s, CHO). Anal. Calcd for $\text{C}_{18}\text{H}_{10}\text{N}_4\text{O}_2\text{S}_2$ (MW 378.43): C, 57.13; H, 2.66; N, 14.81. Found: C, 56.98; H, 2.74; N, 15.01.

2,6-Dichloro-1,7-dimethyl-1,7-dihydropyrrolo[3,2- f]indole-3,5-dicarbaldehyde (17). Yield 85%; mp 300 °C dec. IR (cm^{-1}): 1644, 1506, 1096, 1049, 718. ^1H NMR (CF_3COOD): 4.12 (6H, s, CH_3), 7.79 (1H, s, ar), 9.02 (1H, s, ar), 9.82 (2H, s, CHO). Anal. Calcd for $\text{C}_{14}\text{H}_{10}\text{Cl}_2\text{N}_2\text{O}_2$ (MW 309.15): C, 54.39; H, 3.26; N, 9.06. Found: C, 56.22; H, 2.99; N, 8.98.

2-Chloro-1-methyl-1H-indole-3,6-dicarbaldehyde (20). Yield 2%; mp 181–183 °C. IR (cm^{-1}): 1700, 1686, 1646, 1154, 722. ^1H NMR ($\text{DMSO-}d_6$): 3.93 (3H, s, CH_3), 7.83 (1H, dd, ind-5, $J = 8.4$, $J = 1.2$), 8.23 (1H, d, ind-4, $J = 8.4$), 8.24 (1H, d, ind-7, $J = 1.2$), 10.03 (1H, s, CHO), 10.07 (1H, s, CHO). Anal. Calcd for $\text{C}_{11}\text{H}_8\text{ClNO}_2$ (MW 221.64): C, 59.61; H, 3.64; N, 6.32. Found: C, 60.04; H, 3.87; N, 5.98.

2-Chloro-1-phenyl-1H-indole-3,6-dicarbaldehyde (21). Yield 7%; mp 190–195 °C. IR (cm^{-1}): 1654, 1608, 1151, 827, 725. ^1H NMR ($\text{DMSO-}d_6$): 7.66 (1H, d, ind-7, $J = 1.6$), 7.71 (5H, m, ph), 7.89 (1H, dd, ind-5, $J = 8.4$, $J = 1.6$), 8.36 (1H, d, ind-4, $J = 8.4$), 10.00 (1H, s, CHO), 10.16 (1H, s, CHO). Anal. Calcd for $\text{C}_{16}\text{H}_{10}\text{ClNO}_2$ (MW 283.71): C, 67.74; H, 3.55; N, 4.94. Found: C, 68.02; H, 3.86; N, 5.08.

General Procedure for the Synthesis of the Hydrazones 2–5 (Scheme 1) and 6–10 (Scheme 2). The appropriate aldehyde or ketone (5 mmol) was dissolved in ethanol and treated with two equivalents of:

- 2-Hydrazino-2-imidazoline hydrobromide solubilized in ethanol (compounds **2** and **3**);
- 2-Hydrazinopyridine solubilized in ethanol and treated with hydrochloridric acid (compounds **4** and **10**);
- Aminoguanidine hydrogencarbonate suspended in ethanol and treated with hydrochloridric acid in order to achieve a solution (compounds **5–9**).

The reaction mixture was refluxed for 15–30 h according to a TLC test, and the resulting precipitate was collected by filtration and crystallized from ethanol. Compound **3** was obtained as a precipitate from the filtered solution.

3,9-Bis((2-(4,5-dihydro-1H-imidazol-2-yl)hydrazono)methyl)-2,8-diphenyldiimidazo[1,2- α' ,2'- c]pyrimidine Dihydrobromide (2). Yield 22%; mp 300 °C. IR (cm^{-1}): 3360, 1655, 1071, 938, 718. ^1H NMR ($\text{DMSO-}d_6$): 3.72 (4H, s, CH_2), 3.81 (4H, s, CH_2), 7.37 (1H, d, pym, $J = 8.0$), 7.49 (3H, m, ph), 7.58 (3H, m, ph), 7.84 (2H, d, ph, $J = 8.2$), 8.12 (2H, d, ph, $J = 8.2$), 8.29 (2H, broad, $\text{NH}/\text{NH}^+\text{imi}$), 8.64 (1H, s, CH), 8.80 (2H, broad, $\text{NH}/\text{NH}^+\text{imi}$), 9.30 (1H, d, pym, $J = 8.0$), 9.49 (1H, s, CH), 12.50 (2H, broad, NNH). ^{13}C NMR ($\text{DMSO-}d_6$): 42.82, 42.92, 103.69, 115.28, 117.74, 128.00, 128.38, 128.87, 128.97, 129.07, 129.22, 129.51, 131.80, 132.70, 139.73, 139.77, 140.22, 142.16, 145.73, 146.45, 157.07, 157.68. Anal. Calcd for $\text{C}_{28}\text{H}_{26}\text{N}_{12}\cdot 2\text{HBr}$ (MW 692.43): C, 48.57; H, 4.08; N, 24.27. Found: C, 48.93; H, 3.95; N, 24.87.

9-((2-(4,5-Dihydro-1H-imidazol-2-yl)hydrazono)methyl)-2,8-diphenyldiimidazo[1,2- α' ,2'- c]pyrimidine-3-carbaldehyde Hydrobromide (3). Yield 50%; mp 320 °C. IR (cm^{-1}): 3321, 1668, 1625, 774, 699. ^1H NMR ($\text{DMSO-}d_6$): 3.72 (4H, s, CH_2), 7.50 (3H, m, ph), 7.52 (1H, d, pym, $J = 7.4$), 7.61 (3H, m, ph), 8.03 (2H, m, ph), 8.12 (2H, m, ph), 8.29 (2H, broad, $\text{NH}/\text{NH}^+\text{imi}$), 9.12 (1H, d, pym, $J = 7.4$), 9.47 (1H, s, CH), 10.09 (1H, s, CHO), 12.71 (1H, broad, NNH). ^{13}C NMR ($\text{DMSO-}d_6$): 42.93, 105.05, 117.91, 120.24, 126.27, 128.01, 128.95, 129.01, 129.52, 129.63, 130.17, 131.02, 132.51, 139.92, 140.47, 142.39, 145.94, 152.81, 157.71, 180.42. Anal. Calcd for $\text{C}_{25}\text{H}_{20}\text{N}_8\text{O}\cdot \text{HBr}$ (MW 529.39): C, 56.72; H, 4.00; N, 21.17. Found: C, 57.03; H, 3.89; N, 21.84.

2,8-Diphenyl-3,9-bis((2-(pyridin-2-yl)hydrazono)methyl)-diimidazo[1,2- α' ,2'- c]pyrimidine Dihydrochloride (4). Yield 71%; mp 267–270 °C. IR (cm^{-1}): 3380, 1651, 1612, 1276, 707. ^1H NMR ($\text{DMSO-}d_6$): 3.75 (2H, broad, NH^+py), 7.00 (1H, t, py, $J = 6.0$), 7.08 (1H, t, ph, $J = 7.2$), 7.17 (1H, d, pym, $J = 8.0$), 7.28 (1H, d, py, $J = 8.4$), 7.40 (1H, d, py, $J = 8.4$), 7.43–7.53 (4H, m, 3Hph+1Hpy), 7.56 (2H, t, ph, $J = 7.2$), 7.93 (2H, d, ph, $J = 7.2$), 7.95 (1H, t, py, $J = 8.4$), 8.06 (1H, t, py, $J = 8.4$), 8.09 (1H, d, py, $J = 6.0$), 8.14 (2H, d, ph, $J = 7.2$), 8.22 (1H, d, py, $J = 6.0$), 8.83 (1H, s, CH), 9.56 (1H, d, pym, $J = 8.0$), 9.64 (1H, s, CH), 12.82 (1H, broad, NNH), 13.13 (1H, broad, NNH). ^{13}C NMR ($\text{DMSO-}d_6$): 103.77, 111.67, 114.60, 114.96, 115.54, 118.87, 127.92, 128.40, 128.69, 128.88, 129.05, 129.09, 129.32, 131.82, 133.16, 139.67, 141.84, 144.91, 146.08. Anal. Calcd for $\text{C}_{32}\text{H}_{24}\text{N}_{10}\cdot 2\text{HCl}$ (MW 621.52): C, 61.84; H, 4.22; N, 22.54. Found: C, 62.05; H, 4.59; N, 22.87.

2,2'-((2,8-Di(thiophen-3-yl)diimidazo[1,2-a:1',2'-c]pyrimidine-3,9-diyil)bis(methanylylidene))bis(hydrazine-1-carboximidamide) Dihydrochloride (5). Yield 98%; mp 265–269 °C dec. IR (cm⁻¹): 3358, 1675, 1623, 1149, 866. ¹H NMR (DMSO-*d*₆): 7.28 (1H, d, pym, *J* = 8.0), 7.66 (1H, dd, th, *J* = 5.2, *J* = 2.8), 7.69 (8H, broad, NH₂/NH₂⁺), 7.76 (1H, dd, th, *J* = 5.2, *J* = 1.2), 7.78 (1H, dd, th, *J* = 5.2, *J* = 1.2), 7.81 (1H, dd, th, *J* = 5.2, *J* = 2.8), 8.26 (1H, dd, th, *J* = 2.8, *J* = 1.2), 8.47 (1H, dd, th, *J* = 2.8, *J* = 1.2), 8.82 (1H, s, CH), 9.26 (1H, d, pym, *J* = 8.0), 9.69 (1H, s, CH), 12.02 (1H, s, NNH), 12.57 (1H, s, NNH). ¹³C NMR (DMSO-*d*₆): 103.26, 114.63, 117.54, 125.72, 126.32, 126.88, 127.45, 127.77, 128.44, 128.56, 133.95, 134.00, 138.43, 139.19, 139.57, 141.27, 141.77, 142.05, 154.67, 155.20. Anal. Calcd for C₂₀H₁₈N₁₂S₂·2HCl (MW 563.49): C, 42.63; H, 3.58; N, 29.83. Found: C, 42.97; H, 3.02; N, 30.05.

2,2'-((2,6-Dichloro-1,7-dimethyl-1,7-dihydropyrrolo[3,2-*f*]indole-3,5-diyil)bis(methanylylidene))bis(hydrazine-1-carboximidamide) Dihydrochloride (6). Yield 50%; mp 350 °C dec. IR (cm⁻¹): 3329, 1643, 1516, 1148, 1096. ¹H NMR (DMSO-*d*₆): 3.88 (6H, s, CH₃), 7.60 (8H, broad, NH₂/NH₂⁺), 7.88 (1H, s, ar), 8.47 (2H, s, CH), 8.68 (1H, s, ar), 11.70 (2H, broad, NNH). ¹³C NMR (DMSO-*d*₆): 30.65, 91.84, 105.66, 112.71, 119.67, 130.97, 134.61, 143.06, 154.53. Anal. Calcd for C₁₆H₁₈Cl₂N₁₀·2HCl (MW 494.21): C, 38.89; H, 4.08; N, 28.34. Found: C, 39.03; H, 3.87; N, 28.84.

2,2'-((2-Chloro-1-methyl-1-*H*-indole-3,6-diyil)bis(methanylylidene))bis(hydrazine-1-carboximidamide) Dihydrochloride (7). Yield 75%; mp 235–240 °C. IR (cm⁻¹): 3150, 1670, 1605, 1129, 934. ¹H NMR (DMSO-*d*₆): 3.85 (3H, s, CH₃), 7.67 (1H, d, ind, *J* = 7.8), 7.70 (8H, broad, NH₂/NH₂⁺), 8.15 (1H, s, CH), 8.30 (1H, s, CH), 8.35 (2H, d, ind, *J* = 7.8), 12.03 (1H, s, NNH), 12.14 (1H, s, NNH). ¹³C NMR (DMSO-*d*₆): 30.57, 106.59, 109.60, 121.59, 121.95, 124.42, 128.72, 131.93, 136.15, 141.62, 147.32, 154.86, 155.44. Anal. Calcd for C₁₃H₁₆ClN₈·2HCl (MW 406.70): C, 38.39; H, 4.46; N, 31.00. Found: C, 38.78; H, 4.73; N, 29.02.

2,2'-((2-Chloro-1-phenyl-1-*H*-indole-3,6-diyil)bis(methanylylidene))bis(hydrazine-1-carboximidamide) Dihydrochloride (8). Yield 67%; mp 350 °C dec. IR (cm⁻¹): 3163, 1666, 1631, 1158, 934. ¹H NMR (DMSO-*d*₆): 7.57 (5H, m, ph), 7.58 (1H, d, ind-7, *J* = 1.4), 7.76 (8H, broad, NH₂/NH₂⁺), 7.87 (1H, dd, ind-5, *J* = 8.6, *J* = 1.4), 8.25 (1H, s, CH), 8.47 (1H, s, CH), 8.48 (1H, d, ind-4, *J* = 8.6), 12.05 (1H, s, NNH), 12.23 (1H, s, NNH). ¹³C NMR (DMSO-*d*₆): 108.16, 110.57, 121.45, 122.39, 124.54, 128.16, 129.47, 129.55, 129.98, 131.33, 134.31, 136.74, 141.50, 147.71, 154.93, 155.37. Anal. Calcd for C₁₈H₁₈ClN₈·2HCl (MW 468.77): C, 46.12; H, 4.30; N, 26.89. Found: C, 45.97; H, 4.78; N, 27.02.

2,2'-(2-*E*-1,1'-Dimethyl-[2,2'-biindolinylidene]-3,3'-diylidene)bis(hydrazine-1-carboximidamide) Dihydrochloride (9). Yield 32%; mp 315–320 °C. IR (cm⁻¹): 3326, 1667, 1571, 1041, 722. ¹H NMR (DMSO-*d*₆): 3.22 (6H, s, CH₃), 7.16 (2H, d, ind, *J* = 7.6), 7.18 (2H, t, ind, *J* = 7.6), 7.49 (2H, t, ind, *J* = 7.6), 7.75 (2H, d, ind, *J* = 7.6), 8.52 (8H, broad, NH₂/NH₂⁺), 12.48 (2H, s, NNH). ¹³C NMR (DMSO-*d*₆): 25.79, 109.94, 118.88, 121.31, 123.00, 132.08, 136.26, 144.09, 155.88, 160.18. Anal. Calcd for C₂₀H₂₂N₁₀·2HCl (MW 475.38): C, 50.53; H, 5.09; N, 29.46. Found: C, 50.98; H, 4.86; N, 29.78.

2-*E*-1,1'-Dimethyl-3,3'-bis(2-(pyridin-2-yl)hydrazono)-2,2'-biindolinylidene Dihydrochloride (10). Yield 25%; mp 255–260 °C. IR (cm⁻¹): 3397, 1685, 1638, 1117, 1043. ¹H NMR (DMSO-*d*₆): 3.26 (6H, s, CH₃), 3.72 (2H, broad, NH⁺py), 7.16 (6H, m, py), 7.42 (2H, t, ind, *J* = 7.6), 7.67 (4H, m, ind+py), 7.97 (2H, t, ind, *J* = 7.6), 8.32 (2H, m, ind), 12.84 (2H, s, NNH). ¹³C NMR (DMSO-*d*₆): 25.58, 107.53, 109.54, 118.60, 119.04, 119.76, 122.65, 129.31, 129.57, 139.17, 142.02, 147.83, 154.46, 161.25. Anal. Calcd for C₂₈H₂₄N₈·2HCl (MW 545.47): C, 61.65; H, 4.80; N, 20.54. Found: C, 61.20; H, 5.02; N, 20.12.

Oligonucleotide Synthesis and Sample Preparation. The DNA sequences were synthesized using standard β -cyanoethylphosphoramidite solid phase chemistry on an ABI 394 DNA/RNA synthesizer (Applied Biosystem) at the 5 μ mol scale. DNA detachment from support and deprotection were performed by treatment with concentrated ammonia aqueous solution at 55 °C for 12 h. The combined filtrates and washings were concentrated under reduced pressure, dissolved in water, and purified by high-performance liquid chromatography (HPLC) on a Nucleogel SAX column (Macherey-Nagel, 1000–8/46), using buffer A consisting of 20 mM KH₂PO₄/

K₂HPO₄ aqueous solution (pH 7.0), containing 20% (v/v) CH₃CN, buffer B consisting of 1 M KCl, 20 mM KH₂PO₄/K₂HPO₄ aqueous solution (pH 7.0), containing 20% (v/v) CH₃CN, and a linear gradient from 0% to 100% B for 30 min with a flow rate 1 mL/min. The fractions of the oligomers were collected and successively desalted by Sep-pak cartridges (C-18). The isolated oligomer was proved to be >98% pure by NMR. In particular, the following oligonucleotides have been synthesized: d[TAGGG(TTAGGG)₃] (Tel23) and d-[(TTAGGG)₄TT] (Tel26) that are two different truncations of human telomeric sequence, two sequences that occur in the nuclease hypersensitive region of the promoter of *c-KIT* oncogene d(AGGGAG-GGGCTGGGAGGAGGG) (*c-kit*1) and d(CGCGCGGGCGCG-AGGGAGGGG) (*c-kit*2), the modified *c-kit*2 promoter sequence d(CGCGCGGGCGCTAGGGAGGGT) (*c-kit*2_{T12/T21}), the *c-MYC* promoter sequence d(TGAGGGTGGGGAGGGTGGGGGAAGG) (*c-myc*), the modified *c-MYC* promoter sequence d(TGAGGGTGGGTAGGGTGGGTAA) (*Myc*22), and the self-complementary duplex-forming dodecamer d(CGCGAATTCGCG) (*ds*12). The concentration of oligonucleotides was determined by UV adsorption measurements at 90 °C using appropriate molar extinction coefficient values ϵ (λ = 260 nm) calculated by the nearest neighbor model.³⁵ Samples were heated at 90 °C for 5 min and then gradually cooled to room temperature overnight. Parallel arrangement of telomeric sequence (Tel23-p) was prepared as previously described.¹⁵

Circular Dichroism Spectroscopy. Circular dichroism (CD) experiments were recorded on a Jasco J-815 spectropolarimeter equipped with a PTC-423S/15 Peltier temperature controller. All the spectra were recorded at 20 °C in the wavelength range of 230–360 nm and averaged over three scans. The scan rate was set to 100 nm/min, with a 1 s response time and 1 nm bandwidth. Buffer baseline was subtracted from each spectrum. For the CD experiments in the presence of metal cations, 10 μ M G4 DNA and 15 μ M duplex DNA were used. CD spectra of DNA/ligand mixtures were obtained by adding 4 molar equiv of ligands (stock solutions of ligands were 6 mM in DMSO). DNAs were prepared in 10 mM potassium phosphate (for *c-kit*1, Tel23, and Tel26) or 10 mM lithium phosphate (for Tel23-p and *ds*12) buffers (pH 7.0) containing 100 mM KCl. Because the *T_m* value of the *c-myc* G4 alone is above 80 °C in buffer containing 100 mM KCl, it is very difficult to evaluate the stabilizing effects of adding ligands (see Supporting Information, Figure S6), thus a buffer solution containing lower amounts of K⁺ (5 mM potassium phosphate, 20 mM KCl) was used for the experiments on *c-myc* G4. As for *c-kit*2, it was shown that this sequence forms a dimeric G4 in 100 mM K⁺-containing solution,³¹ with a completely different fold compared to the monomeric G4 that is formed in buffer solution containing low amounts of K⁺.^{31,32} Therefore, 5 mM potassium phosphate buffer containing 20 mM KCl was also used in this case.

For the studies in the absence of metal cations, oligonucleotides (12–15 μ M) were prepared in 50 mM Tris buffer and the CD spectra were recorded 10 min after each ligand addition (stepwise additions of 0.5 equiv). All spectra were baseline corrected and analyzed using Origin 7.0 software.

CD melting were carried out in the 20–100 °C temperature range at 1 °C/min heating rate by following changes of CD signal at the wavelengths of the maximum CD intensity. CD melting experiments were recorded in the absence and presence of ligands (4 molar equiv) added to the folded DNA structures. For selected compounds, CD melting experiments were also performed by varying ligand concentration. Three different [ligand]/[DNA] ratio were examined (1:1, 2:1 and 4:1). The melting temperatures (*T_m*) were determined from curve fit using Origin 7.0 software. ΔT_m values were determined as the difference in melting temperature between the G4-DNA with and without ligands.

FRET Melting Studies. The ability of the compounds to stabilize DNA sequences was investigated using a fluorescence resonance energy transfer (FRET) assay modified to be used as a high-throughput screen in a 96-well format. The studied sequences were the G4 forming sequence 5'-FAM-d(GGG[TTAGGG]₃)-TAMRA-3' (F2IT),²⁹ and the duplex-forming sequence 5'-FAM-dTATAGCTATA-HEG-TATAGCTATA-TAMRA-3' (T-loop) (HEG linker: [(-CH₂-CH₂-

O₆)). The labeled oligonucleotides had attached to them the donor fluorophore FAM, 6-carboxyfluorescein, and the acceptor fluorophore TAMRA, 6-carboxytetramethylrhodamine. The FRET probe sequences were diluted from stock to the correct concentration (400 nM) in a 60 mM potassium cacodylate buffer (pH 7.4) and then annealed by heating to 85 °C for 10 min, followed by cooling to room temperature in the heating block. The compounds were stored as a 10 mM stock solution in DMSO; final solutions (at 2× concentration) were prepared using 10 mM HCl in the initial 1:10 dilution, after which 60 mM potassium cacodylate buffer (pH 7.4) was used in all subsequent steps. The maximum HCl concentration in the reaction volume (at a ligand concentration of 20 μM) is thus 200 μM, well within the range of the buffer used. Relevant controls were also performed to check for interference with the assay. 96-Well plates (MJ Research, Waltham, MA) were prepared by aliquoting 50 μL of the annealed DNA into each well, followed by 50 μL of the compound solutions. Measurements were made on a DNA Opticon Engine (MJ Research) with excitation at 450–495 nm and detection at 515–545 nm. Fluorescence readings were taken at intervals of 0.5 °C in the range 30–100 °C, with a constant temperature being maintained for 30 s prior to each reading to ensure a stable value. Final analysis of the data was carried out using a script written in the program Origin 7.0 (OriginLab Corp., Northampton, MA). The advanced curve-fitting function in Origin 7.0 was used for calculation of ΔT_m values. All determinations were performed in triplicate. Esds in ΔT_m are ± 0.2 °C.

For selected compounds, additional FRET melting studies were performed by using the following G4 forming sequences: the 5'-FAM-d(GGGCGGGCGGAGGGAGGGG)-TAMRA-3' (Fckit2T) and 5'-FAM-d(TGAGGGTGGGTAGGGTGGGTAA)-TAMRA-3' (FcmycT) oligonucleotides from the *c-KIT* and *c-MYC* promoter regions, respectively, and the telomeric F21T G4-forming sequence were annealed at high concentrations to promote the parallel conformation (F21T-p). Labeled oligonucleotides were purchased from Biomers (Germany) and purified employing standard HPLC protocols. Such studies were performed under the same buffer and experimental conditions used for the corresponding CD experiments. The FRET melting assay was carried out on a FP-8300 spectrofluorometer (Jasco) equipped with a Peltier temperature controller accessory (Jasco PCT-818) with excitation at 492 nm and detection at 522 nm. Both excitation and emission slit widths were set at 5 nm. A sealed quartz cuvette with a path length of 1 cm was used. The final concentration of the G4s was 0.2 μM. The fluorescence melting of the G4s was monitored at 1 °C/min with and without ligands (at a 4:1 ligand/DNA ratio). Final analysis of the data was carried out using Origin 7.0.

Nuclear Magnetic Resonance Experiments. NMR experiments were performed on a 700 MHz Varian Unity INOVA spectrometer, with data recorded at 25 °C. One-dimensional proton spectra of the sample in H₂O were recorded using pulsed-field gradient DPGSE for H₂O suppression. DNA samples were prepared at 0.2–0.5 mM strand concentration in 0.6 mL of (H₂O/D₂O 9:1) buffer solution. NMR experiments on G4 forming oligonucleotides from *c-MYC* (Myc22) and *c-KIT* (*c-kit*_{T12/T21}) promoters were performed by employing the same buffers and experimental conditions as used for the determination of their 3D structures^{30,31} in order to avoid different G4 folds or conformational heterogeneity in solution, which is detrimental to such experiments. The solution was either 25 mM KH₂PO₄, 70 mM KCl, 0.2 mM EDTA, pH 7.0 (for Myc22), or 5 mM KH₂PO₄, 20 mM KCl, 0.2 mM EDTA, pH 6.8 (for *c-kit*_{T12/T21}). Aliquots of a stock solution of 3 in DMSO-*d*₆ were added directly to the DNA solution inside the NMR tube. The NMR data were processed on an iMAC running iNMR software (www.inmr.net).

Docking Simulations. The solution structures of G4 formed by the human *c-kit* promoter sequence (PDB ID: 2KQH)³² and of *c-myc* G4 bound to quindoline molecule (PDB ID: 2L7V)³³ were used as the targets for docking studies. The ligand found in the *c-myc* complex was removed from the structure to leave empty binding sites. After optimizing the ligand and assigning partial atomic charges, docking calculations were performed with the AutoDock 4.2 program using the Lamarckian genetic algorithm.³⁶ The all-parallel-stranded arrangement

with propeller loops linking adjacent parallel strands on opposite surfaces of the G4s results in accessible outer planar G-tetrad surfaces that, being the potential binding sites for the ligands, were defined as the grid boxes for the docking. The size of the boxes was constrained to 18 Å × 18 Å × 18 Å in the *x*, *y*, and *z* dimensions. Grid maps were generated for each atom type in the ligand using AutoGrid. An active site box was created with a grid spacing of 0.375 Å. The maximum number of energy evaluations was set to 2.5 × 10⁶, the maximum number of genetic algorithm operations was set to 2.7 × 10⁴, the number of individuals in a population was set to be 300, and the rates of mutation and crossover were set to 0.02 and 0.8, respectively. When searching the conformational and orientation spaces of ligand with rotatable bonds having full flexibility, the structure of the G4 was kept rigid. The most favorable binding poses were selected on the basis of both the calculated energies and visual inspection. All the figures were drawn using VMD 1.9 software (www.ks.uiuc.edu/Research/vmd/).

Cell Cultures and Immunofluorescence. Human osteosarcoma U2OS and colon cancer HCT116 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Carlo Erba), supplemented with 1% L-glutamine and 10% FBS (fetal bovine serum). U2OS cells were seeded in 35 mm dishes at a concentration of 100000 cell/mL. Twenty-four hours after seeding, cells were treated with 1 (2 and 10 μM), 3 (0.2 and 1 μM), or 14a (2 and 10 μM). Cells were then fixed in methanol:acetic acid (3:1), permeabilized with 0.1% triton-X100/PBS, and blocking in 2% milk/PBS. Immunofluorescence was performed using standard methods with BG4, anti-FLAG (Cell Signaling Technology), and antirabbit Alexa 488-conjugated (Invitrogen) antibodies. Nuclei were stained with DAPI (Sigma-Aldrich), and coverslips were mounted with Mowiol 4–88 (Sigma-Aldrich). Fluorescence signal was determined using ImageJ software with the following formula: corrected total cell fluorescence (CTCF) = integrated density – (area of selected cell × mean fluorescence of background readings). Significance was determined by parametric tests "Student's *t* test". BG4 antibody was obtained by transfection of BG4 plasmid (kindly obtained by S. Balasubramanian) in BL21 *Escherichia coli* cells. Then BG4 protein expression was induced by the autoinduction method as described by Studier.³⁷ BG4 was purified by using silica-based resin (Protino Ni-IDA) precharged with Ni²⁺ ions, eluted with 250 mM Imidazole/PBS pH 8.0. The eluted antibody was concentrated with Amicon Ultra-15 centrifugal filter units (Millipore), and imidazole was finally removed by buffer exchange with PBS pH 8.0 with Amicon Ultra-15 centrifugal filter units.

MTT Cell Proliferation Assay. U2OS and HCT116 cells were cultured in DMEM as above and were seeded in 24-wells at a concentration of 40000 cell/mL. Twenty-four hours after seeding, cells were treated with the compound at the indicated concentration and time. After 1 or 24 h of treatment, agents were removed and cells were further cultured in complete drug-free medium for 1–2 days. Then, thiazolyl blue tetrazolium bromide (MTT) (Sigma-Aldrich) was added to each well and incubated for 1 h at 37 °C. Then medium was removed and 300 μL of dimethyl sulfoxide (Sigma-Aldrich) were added and incubated for 1 h at room temperature. Then 100 μL of the solution was put in 96-well, and absorbance at 595 nm was measured using a multiplate reader. The linear regression parameters were determined to calculate the IC₅₀ (GraphPad Prism 4.0, Graph Pad Software Inc.).

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jmedchem.6b00129.

CD spectra and CD melting curves of the investigated DNAs in the absence and in the presence of 4 mol equiv of compounds 1–10. FRET-melting experiments at various concentrations of compounds 1–10. CD spectra of labeled oligonucleotides. Additional FRET-melting experiments on compounds 2, 3, and 5. CD titration of the investigated G4s with compounds 2, 3, and 4 (PDF) Molecular formula strings(CSV)

AUTHOR INFORMATION

Corresponding Author

*Phone: (+39) 081 678514. Fax: (+39) 081 678514. E-mail: antonio.randazzo@unina.it.

Author Contributions

#J.A. and R.M. contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Italian Association for Cancer Research (AIRC) (IG-14150 to A.R., IG-16730 to B.P., and IG-15886 to G.C.), by Italian Institute of Technology (IIT), and by "Programma STAR" 2014 of University of Naples "Federico II" to B.P. (14-CSP3-C03-141).

ABBREVIATIONS USED

G4, G-quadruplex; ALT, alternative lengthening of telomere; DMEM, Dulbecco's Modified Eagle Medium; FBS, fetal bovine serum

REFERENCES

- (1) Bacolla, A.; Wells, R. D. Non-B DNA conformations as determinants of mutagenesis and human disease. *Mol. Carcinog.* **2009**, *48*, 273–285.
- (2) Balasubramanian, S.; Hurley, L. H.; Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discovery* **2011**, *10*, 261–275.
- (3) Amato, J.; Iaccarino, N.; Randazzo, A.; Novellino, E.; Pagano, B. Noncanonical DNA secondary structures as drug targets: The prospect of the i-motif. *ChemMedChem* **2014**, *9*, 2026–2030.
- (4) Balasubramanian, S.; Neidle, S. G-quadruplex nucleic acids as therapeutic targets. *Curr. Opin. Chem. Biol.* **2009**, *13*, 345–353.
- (5) Collie, G.; Parkinson, G. The application of DNA and RNA G-quadruplexes to therapeutic medicines. *Chem. Soc. Rev.* **2011**, *40*, 5867–5892.
- (6) Collado, M.; Blasco, M. A.; Serrano, M. Cellular Senescence in cancer and aging. *Cell* **2007**, *130*, 223–233.
- (7) Kim, N. W. Clinical implications of telomerase in cancer. *Eur. J. Cancer* **1997**, *33*, 781–786.
- (8) Shay, J. W.; Reddel, R. R.; Wright, W. E. Cancer and telomeres - an ALternative to telomerase. *Science* **2012**, *336*, 1388–1390.
- (9) Zahler, A. M.; Williamson, J. R.; Cech, T. R.; Prescott, D. M. Inhibition of telomerase by G-quartet DNA structures. *Nature* **1991**, *350*, 718–720.
- (10) Neidle, S. Human telomeric G-quadruplex: The current status of telomeric G-quadruplexes as therapeutic targets in human cancer. *FEBS J.* **2010**, *277*, 1118–1125.
- (11) Monchaud, D.; Teulade-Fichou, M.-P. A hitchhiker's guide to G-quadruplex ligands. *Org. Biomol. Chem.* **2008**, *6*, 627–636.
- (12) Ohnmacht, S. A.; Neidle, S. Small-molecule quadruplex-targeted drug discovery. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 2602–2612.
- (13) Cosconati, S.; Rizzo, A.; Trotta, R.; Pagano, B.; Iachettini, S.; De Tito, S.; Lauri, I.; Fotticchia, I.; Giustiniano, M.; Marinelli, L.; Giancola, C.; Novellino, E.; Biroccio, A.; Randazzo, A. Shooting for selective druglike G-quadruplex binders: Evidence for telomeric DNA damage and tumor cell death. *J. Med. Chem.* **2012**, *55*, 9785–9792.
- (14) Di Leva, F. S.; Zizza, P.; Cingolani, C.; D'Angelo, C.; Pagano, B.; Amato, J.; Salvati, E.; Sissi, C.; Pinato, O.; Marinelli, L.; Cavalli, A.; Cosconati, S.; Novellino, E.; Randazzo, A.; Biroccio, A. Exploring the chemical space of G-quadruplex binders: Discovery of a novel chemotype targeting the human telomeric sequence. *J. Med. Chem.* **2013**, *56*, 9646–9654.
- (15) Pagano, B.; Amato, J.; Iaccarino, N.; Cingolani, C.; Zizza, P.; Biroccio, A.; Novellino, E.; Randazzo, A. Looking for efficient G-quadruplex ligands: Evidence for selective stabilizing properties and

telomere damage by drug-like molecules. *ChemMedChem* **2015**, *10*, 640–649.

- (16) Sparapani, S.; Bellini, S.; Gunaratnam, M.; Haider, S. M.; Andreani, A.; Rambaldi, M.; Locatelli, A.; Morigi, R.; Granaiola, M.; Varoli, L.; Burnelli, S.; Leoni, A.; Neidle, S. Bis-guanyldiazone diimidazo[1,2-a:1,2-c]pyrimidine as a novel and specific G-quadruplex binding motif. *Chem. Commun.* **2010**, *46*, 5680–5682.

- (17) Andreani, A.; Granaiola, M.; Leoni, A.; Locatelli, A.; Morigi, R.; Rambaldi, M.; Giorgi, G.; Garaliene, V. Potential antitumor agents. 34. Synthesis and antitumor activity of guanyldiazones from imidazo[2,1-b]thiazoles and from diimidazo[1,2-a:1,2-c]pyrimidine. *Anticancer Res.* **2004**, *24*, 203–212.

- (18) Andreani, A.; Rambaldi, M.; Locatelli, A.; Bossa, R.; Fraccari, A.; Galatulas, I. Potential antitumor agents. 21. Structure determination and antitumor activity of imidazo[2,1-b]thiazole guanyldiazones. *J. Med. Chem.* **1992**, *35*, 4634–4637.

- (19) Andreani, A.; Granaiola, M.; Leoni, A.; Locatelli, A.; Morigi, R.; Rambaldi, M.; Giorgi, G.; Salvini, L. Cancer fighting cancer: synthesis of the new heterocyclic system diimidazo-[1,2-a:1,2-c]-pyrimidine. *ARKIVOC* **2002**, *2001* (11), 32–38.

- (20) Kipnis, F.; Soloway, H.; Ornfelt, J. 2-Acyloxyacetylthiophenes. *J. Am. Chem. Soc.* **1949**, *71*, 10–11.

- (21) Daszkiewicz, Z.; Domanski, A.; Kyziol, J. B. A simple route to N-methylarylamines. *Chem. Pap.* **1993**, *47*, 109–113.

- (22) Hodges, R.; Shannon, J. S.; Jamieson, W. D.; Taylor, A. Chemical and biological properties of some oxindol-3-ylidene methines. *Can. J. Chem.* **1968**, *46*, 2189–2194.

- (23) Kisteneva, M. S. Azomethine dyes from oxindole derivatives. *Zh. Obshch. Khim.* **1956**, *26*, 1169–1175.

- (24) Russell, G. A.; Kaupp, G. Reactions of resonance stabilized carbanions. XXXI. Oxidation of carbanions. 4. Oxidation of indoxyl to indigo in basic solution. *J. Am. Chem. Soc.* **1969**, *91*, 3851–3859.

- (25) Dai, J.; Carver, M.; Yang, D. Polymorphism of human telomeric quadruplex structures. *Biochimie* **2008**, *90*, 1172–1183.

- (26) Parkinson, G. N.; Lee, M. P.; Neidle, S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **2002**, *417*, 876–880.

- (27) Xue, Y.; Kan, Z. Y.; Wang, Q.; Yao, Y.; Liu, J.; Hao, Y. H.; Tan, Z. Human telomeric DNA forms parallel-stranded intramolecular G-quadruplex in K⁺ solution under molecular crowding condition. *J. Am. Chem. Soc.* **2007**, *129*, 11185–11191.

- (28) (a) Masiero, S.; Trotta, R.; Pieraccini, S.; De Tito, S.; Perone, R.; Randazzo, A.; Spada, G. P. A non-empirical chromophoric interpretation of CD spectra of DNA G-quadruplex structures. *Org. Biomol. Chem.* **2010**, *8*, 2683–2692. (b) Karsisiotis, A. I.; Hessari, N. M.; Novellino, E.; Spada, G. P.; Randazzo, A.; Webba Da Silva, M. Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism. *Angew. Chem., Int. Ed.* **2011**, *50*, 10645–10648. (c) Randazzo, A.; Spada, G. P.; Webba Da Silva, M. Circular dichroism of quadruplex structures. *Top. Curr. Chem.* **2012**, *330*, 67–86.

- (29) Mergny, J. L.; Maurizot, J. C. Fluorescence resonance energy transfer as a probe for G-quartet formation by a telomeric repeat. *ChemBioChem* **2001**, *2*, 124–132.

- (30) Ambrus, A.; Chen, D.; Dai, J.; Jones, R. A.; Yang, D. Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization. *Biochemistry* **2005**, *44*, 2048–2058.

- (31) Kuryavyy, V.; Phan, A. T.; Patel, D. J. Solution structures of all parallel-stranded monomeric and dimeric G-quadruplex scaffolds of the human c-kit2 promoter. *Nucleic Acids Res.* **2010**, *38*, 6757–6773.

- (32) Hsu, S. T. D.; Varnai, P.; Bugaut, A.; Reszka, A. P.; Neidle, S.; Balasubramanian, S. A G-rich sequence within the c-kit oncogene promoter forms a parallel G-quadruplex having asymmetric G-tetrad dynamics. *J. Am. Chem. Soc.* **2009**, *131*, 13399–13409.

- (33) Dai, J.; Carver, M.; Hurley, L. H.; Yang, D. Solution structure of a 2:1 quindoline-c-MYC G-quadruplex: Insights into G-quadruplex-interactive small molecule drug design. *J. Am. Chem. Soc.* **2011**, *133*, 17673–17680.

(34) Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **2013**, *5*, 182–186.

(35) Cantor, C. R.; Warshaw, M. M.; Shapiro, H. Oligonucleotide interactions. III. Circular dichroism studies of the conformation of deoxyoligonucleolides. *Biopolymers* **1970**, *9*, 1059–1077.

(36) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791.

(37) Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expression Purif.* **2005**, *41*, 207–234.

Targeting the *BCL2* Gene Promoter G-Quadruplex with a New Class of Furo[2,3-*d*]pyridazinone-Based Molecules

Jussara Amato,^[a] Alessia Pagano,^[a] Domenica Capasso,^[a] Sonia Di Gaetano,^[b] Mariateresa Giustiniano,^[a] Ettore Novellino,^[a] Antonio Randazzo,^[a] and Bruno Pagano*^[a]

Targeting of G-quadruplex-forming DNA in the *BCL2* gene promoter to inhibit the expression of anti-apoptotic Bcl-2 protein is an attractive approach to cancer treatment. So far, efforts made in the discovery of molecules that are able to target the *BCL2* G-quadruplex have succeeded mainly in the identification of ligands with poor drug-like properties. Here, a small series of furo[2,3-*d*]pyridazin-4(5*H*)-one derivatives were evaluated as a new class of drug-like G-quadruplex-targeting compounds. Biophysical studies showed that two derivatives could effectively bind to *BCL2* G-quadruplex with good selectivity. Moreover, one such ligand was found to appreciably inhibit *BCL2* gene transcription, with a substantial decrease in protein expression levels, and also showed significant cytotoxicity toward the Jurkat human T-lymphoblastoid cell line.

Apoptosis is regulated by many factors, which are involved in the activation and execution of related events.^[1] The Bcl-2 (B-cell lymphoma-2) family of proteins is the main regulator of apoptotic process, acting either to promote or inhibit it.^[2] Indeed, this family includes both pro- and anti-apoptotic proteins that should function in harmony for a controlled regulation of apoptosis pathways, and their relative levels are critical for cells. Overexpression of *BCL2* gene, which encodes the anti-apoptotic Bcl-2 protein, greatly contributes to the resistance of cancer cells to apoptosis, and it has also been reported to play a role in the resistance to conventional cancer treatments.^[3] Targeting the *BCL2* gene to inhibit protein expression may be an effective way to prevent the evasion of apoptosis in cancer cells and to increase chemotherapeutic efficacy. The human *BCL2* gene has two promoters: P1, identified as the principal transcriptional promoter, and P2.^[4] The P1 promoter and the region upstream thereof, contain G-rich elements that have unequivocally been shown to form non-canonical secondary structures called G-quadruplexes.^[5]

G-quadruplexes are four-stranded nucleic acid structures whose key structural features are: 1) a core of two or more stacked G-tetrads (a planar arrangement of four guanine bases held together by Hoogsteen hydrogen bonds) stabilized by π - π stacking interactions and monovalent cations (preferably K^+ and Na^+), which coordinate to the eight O6 atoms of adjacent G-tetrads; and 2) loops formed by sequences of variable length, base composition, and orientation that are arranged on the exterior of the G-tetrad core.^[6] Loops are the major elements that define structural variability in G-quadruplexes; in fact, their differences in sequence and orientation result in highly variable and sometimes flexible cavities that may represent specific binding sites (or part thereof) for small ligands. Interestingly, more than 40% of human gene promoters contain at least one G-quadruplex-forming sequence, and, in several cases, the formation of such structural motifs has been demonstrated to be heavily involved in gene regulation.^[7]

A G-quadruplex-forming sequence that plays a major role in the regulation of *BCL2* transcription is located directly upstream (~30 bases) from the P1 promoter.^[8] This sequence forms a predominant G-quadruplex structure (*BCL2*-G4) with a distinctive hybrid (3 + 1) topology having two lateral loops (of 3 and 7 nucleotides, respectively), one single-nucleotide propeller loop, and four grooves of different widths (Figure 1).^[8]

Several organic molecules have been shown to be able to bind and stabilize *BCL2*-G4 with subsequent down-regulation of Bcl-2 transcription and expression levels,^[9–12] thereby demonstrating the real potential of the G-quadruplex-targeting therapeutic approach, which also has the great advantage of circumventing the problem of acquired resistance, which is the main limitation of current Bcl-2-targeted therapeutic strategies.

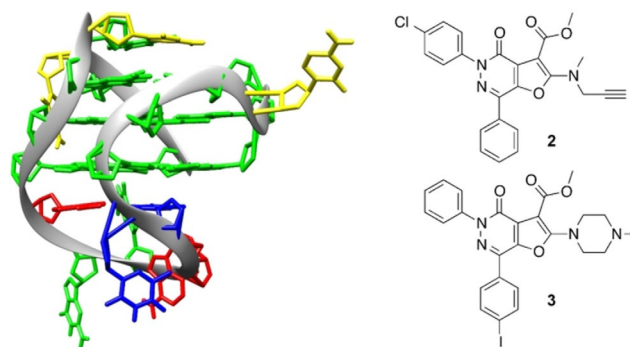


Figure 1. Left: representative model of *BCL2*-G4 NMR structure (guanine, green; adenine, red; thymine, blue; cytosine, yellow). Right: structures of compounds **2** and **3**.

[a] Dr. J. Amato, A. Pagano, Dr. D. Capasso, Dr. M. Giustiniano, Prof. E. Novellino, Prof. A. Randazzo, Dr. B. Pagano
Department of Pharmacy, University of Naples "Federico II", Via D. Montesano 49, 80131, Naples (Italy)
E-mail: bruno.pagano@unina.it

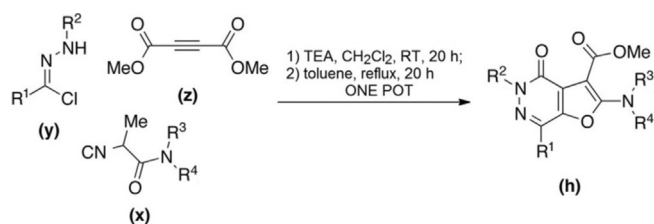
[b] Dr. S. Di Gaetano
Institute of Biostructures and Bioimaging, CNR, Via Mezzocannone 16, 80134, Naples (Italy)

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:
<https://doi.org/10.1002/cmdc.201700749>

Most of the *BCL2*-G4 ligands identified so far are characterized by a large, planar aromatic core with the capability of stacking on the external G-tetrads, a common element in G-quadruplex recognition.^[13–15] Therefore, although these ligands have shown good G-quadruplex over duplex selectivity, they do not exhibit specificity for *BCL2*-G4 versus other G-quadruplex structures. Moreover, most of them share unfavorable chemical properties such as high molecular weight and/or hydrophobic nature, which result in very poor drug-like properties. Therefore, despite being excellent ligands *in vitro*, the unfavorable pharmacokinetics and toxicity due to poor selectivity versus different G-quadruplex structures critically hamper their advancement in chemotherapy.

The next goal in G-quadruplex-based drug design is to enhance the drug-likeness of the ligands and their selectivity for a specific structure, so as to permit pathway-specific targeting. The peculiar structural features of *BCL2*-G4 suggest that specific targeting of such a promoter G-quadruplex may, in principle, be possible. To discover structurally novel compounds that do not show the aforementioned drawbacks, it is necessary to explore a larger chemical space. In this perspective, compounds without a large planar aromatic core may be more promising to achieve the desired drug-likeness and G-quadruplex specificity.

With the aim of identifying compounds with these properties, we decided to exploit the innate ability of multicomponent reactions to generate diversity and to explore a large chemical space in short times, for searching new G-quadruplex-targeting ligands. In particular, we designed and synthesized a small series of furo[2,3-*d*]pyridazinone-based molecules. The pyridazinone nucleus was selected because it is present in a plethora of compounds endowed with various biological activities (antihypertensive, antithrombotic, anti-inflammatory, and anticancer),^[16,17] and it is considered a privileged structure that is currently experiencing a renewed interest in medicinal chemistry. Starting from this nucleus, we recently presented an expeditious multicomponent one-pot reaction methodology for generating a small collection of fully substituted furo[2,3-*d*]pyridazin-4(5*H*)-ones (FPs).^[18,19] The facile and efficient synthetic route for the FP derivatives is shown in Scheme 1. Briefly, the isocyanoacetamide (**x**) and the hydrazonoyl chloride (**y**) were mixed in dichloromethane in the presence of 1 equiv of TEA and stirred overnight at room temperature. After monitoring the formation of a *Z/E* mixture of 1,3-oxazole-2-hydrazone by TLC analysis, the crude mixture was evaporated to dryness, dissolved in toluene and 2 equiv of dimethyl acetylenedicarboxylate (DMAD, **z**) was added. The reaction mixture was fur-



Scheme 1. General synthetic route for the FP derivatives.

ther stirred at reflux overnight, and the final furo[2,3-*d*]pyridazin-4(5*H*)-one (**h**) was purified by chromatography. Notably, this strategy allowed us to readily produce a variety of analogues (1–11, Supporting Information Table S1) with different side arms. It is known that side arms can be important in modulating the affinity of G-quadruplex-targeting ligands, as well as in influencing their binding specificity toward different conformations.^[20] In addition, their structural features should prevent intercalation into the duplex DNA.

To investigate the *in vitro* binding properties of synthesized FP derivatives for *BCL2*-G4 versus other G-quadruplex structures, circular dichroism (CD) spectroscopy and CD-melting experiments were carried out.^[21] In addition to *BCL2*-G4, two G-quadruplex-forming sequences from *c-KIT* (*c-KIT*-G4) and *c-MYC* (*c-MYC*-G4) oncogene promoter regions, and a G-quadruplex from the human telomere (*Tel24*-G4) were used in these experiments. The self-complementary duplex-forming dodecamer d(CGCGAATTCGCG) (ds12) was also used to estimate the G-quadruplex over duplex selectivity of the ligands. The structure adopted by each DNA sample was first verified by CD measurements (Supporting Information Figure S1). In agreement with the presence of parallel G-quadruplex topologies, *c-KIT*-G4 and *c-MYC*-G4 showed a positive band at 264 nm and a negative one at 243 nm in the CD spectra. Otherwise, *BCL2*-G4 and *Tel24*-G4 showed two positive bands at around 265 and 290 nm and a weak negative band at around 240 nm, in agreement with the presence of hybrid structures as major conformation. The structure of ds12 was also verified by CD, showing the typical spectrum of a duplex DNA. CD experiments were then performed to determine if compounds 1–11 alter the native folding topology of the investigated DNA structures. Upon addition of each compound (4 mol equiv) to the DNAs (G-quadruplexes and duplex), no significant alterations of CD spectra were observed, thus suggesting an overall preservation of their architectures (Supporting Information Figure S1).

The DNA stabilizing properties of 1–11 were evaluated by CD-melting experiments measuring the ligand-induced change in the melting temperature (ΔT_m) of both G-quadruplex and duplex structures (Supporting Information Figure S2). The results of these experiments (Table 1) show that none of the derivatives significantly increase the stability of *c-KIT*-G4, *c-MYC*-G4 and *Tel24*-G4 G-quadruplexes, as well as that of ds12 duplex, thus suggesting their poor ability to bind to those DNA structures ($\Delta T_m \leq 2.6^\circ\text{C}$). However, compounds **2** and **3** (Figure 1) showed an appreciable enhancement of the stability of *BCL2*-G4 ($\Delta T_m \geq 3.5^\circ\text{C}$).

Notably, a non-negligible destabilization of the ds12 duplex occurs with some of these compounds. We hypothesize that it may be due to nonspecific interactions of compounds with the DNA molecule that might favor single-stranded DNA, resulting in a shift of the folded–unfolded equilibrium toward the unfolded form during the melting experiment, which, in turn, would result in a slight decrease in T_m .

Despite compounds **2** and **3** not having shown very potent thermal stabilization effects, they exhibited a clear preference for the *BCL2*-G4 over the other G-quadruplex motifs investigated herein. This suggests that the aromatic furo[2,3-*d*]pyridazin-

Table 1. ΔT_m values measured by CD melting experiments for G-quadruplex DNAs (10 μM) and ds12 duplex (15 μM) upon addition of 4 mol equiv of each compound.

Compd	ΔT_m [$^{\circ}\text{C}$] ^[a]				
	<i>BCL2</i> -G4	<i>c-KIT</i> -G4	<i>c-MYC</i> -G4	<i>Tel24</i> -G4	ds12
1	1.9	1.1	0.2	0.0	-0.6
2	4.0	1.3	0.5	1.4	-1.3
3	3.5	0.9	0.0	0.1	-0.4
4	2.9	0.8	0.5	0.0	-2.0
5	2.3	0.4	1.2	0.0	-0.7
6	2.2	0.8	2.0	0.1	-1.2
7	2.8	1.1	0.8	0.2	-0.5
8	2.3	0.7	1.0	0.1	-2.0
9	2.3	1.0	2.6	0.0	-2.0
10	1.6	1.0	1.4	0.0	-1.6
11	2.9	0.4	2.0	0.1	-1.4

[a] $\Delta T_m = T_{m(\text{DNA} + \text{ligand})} - T_{m(\text{DNA})}$; errors were ± 0.4 $^{\circ}\text{C}$.

4(5*H*)-one core can, if appropriately substituted, actually represent a new molecular scaffold for the specific targeting of *BCL2* gene promoter.

To obtain quantitative data regarding the affinity of **2** and **3** for *BCL2*-G4, fluorescence titration experiments were performed.^[22] Fluorescence emission spectra of the ligands in the absence and presence of increasing amounts of G-quadruplex were recorded. The fluorescence intensity of the ligands decreased gradually with addition of DNA until it reached saturation (Figure 2). The binding curves were obtained by plotting the fraction of bound ligands (α), calculated following fluorescence changes at the emission maximum, as a function of G-quadruplex concentration. The curves were fitted using an independent and equivalent binding sites model, by means of nonlinear regression algorithm, giving binding constants (K_b) of $1.0(\pm 0.4) \times 10^6$ and $0.6(\pm 0.2) \times 10^6 \text{ M}^{-1}$ for compounds **2** and **3**, respectively, and a stoichiometry of 1:1 in both cases.

Furthermore, NMR spectroscopy was used to obtain indications about the binding mode of two selected compounds to *BCL2*-G4.^[21] According to the literature, under the experimental

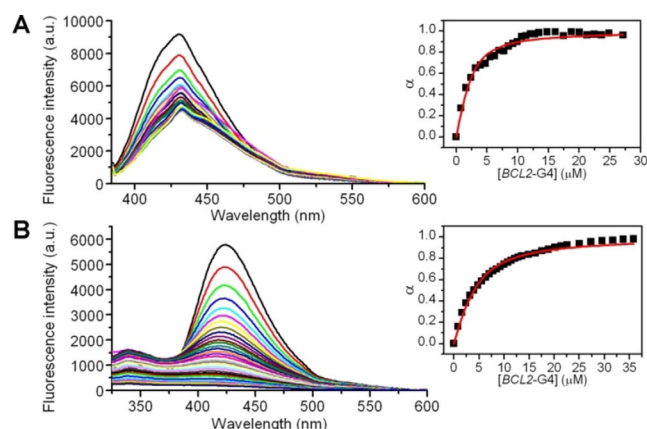


Figure 2. Representative fluorescence emission spectra of A) **2**, and B) **3** (2.5 μM) in the absence and presence of stepwise additions (5 μL) of *BCL2*-G4 (165 μM) at 25 $^{\circ}\text{C}$. Insets show the titration curves obtained by plotting the fraction of bound ligand (α) versus DNA concentration.

conditions used, the investigated sequence forms a single G-quadruplex conformation characterized by 12 well-resolved imino proton peaks, corresponding to the 12 guanines involved in the three G-tetrad planes (Supporting Information Figure S3).^[8] Upon addition of increasing amounts of **2** and **3** to the *BCL2*-G4 solutions, gradual chemical shift changes for some G-quadruplex proton signals were observed. Some signals of the compounds were also detected in the spectra. These peaks only grew in intensity, without showing any significant change in chemical shift values by increasing ligand concentration, clearly suggesting a rapid binding process on the NMR timescale. To evaluate the DNA residues involved in the interaction with **2** and **3**, the chemical shift variations ($\Delta\delta$) of *BCL2*-G4 were calculated (Figure 3 and Supporting Information Table S2). The residues of the G-quadruplex structure are highlighted with different colors to outline different ranges of $\Delta\delta$

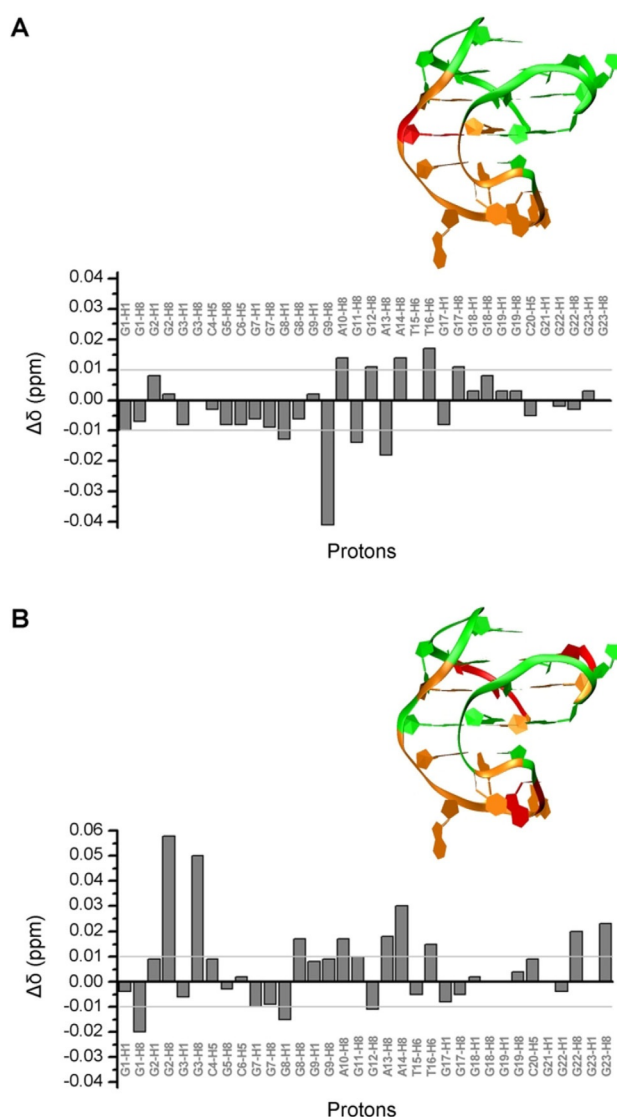


Figure 3. Chemical shift variations ($\Delta\delta$) of selected signals of *BCL2*-G4 DNA (0.2 mM) upon addition (4 mol equiv) of compounds A) **2** and B) **3**, and three-dimensional structure of *BCL2*-G4 colored according to $\Delta\delta$ values: $|\Delta\delta| \leq 0.01$, green; $0.01 < |\Delta\delta| \leq 0.02$, orange; $|\Delta\delta| > 0.02$, red.

values (green, orange, and red for $|\Delta\delta| \leq 0.01$, $0.01 < |\Delta\delta| \leq 0.02$, and $|\Delta\delta| > 0.02$, respectively). Interestingly, both ligands were able to induce significant changes in the chemical shift of the aromatic protons of residues A₁₀–T₁₆ (except for T₁₅), which form the 7-nucleotide lateral loop of the G-quadruplex. Moreover, the binding of **2** also caused the perturbation of the aromatic signals of G9 and G17, that are very close to the loop. On the other hand, **3** also induced variations of the aromatic protons of G1–G3 and G22–G23 residues that face into the same groove of BCL2-G4. All these findings give indications that **2** and **3** bind to the target G-quadruplex mainly through interactions at the long lateral loop and in the groove between the fourth and first parallel G-strands.

Biological assays were also performed to evaluate the cellular effects of compound **3**, and to see whether it could also interact with the G-quadruplex-forming sequence of BCL2 promoter region in cellulo, and thus decrease gene transcription and expression levels. Compound **3** was selected for such further studies because of its higher solubility and BCL2-G4 selectivity over compound **2**. RHPS4, one of the gold-standard ligands for G-quadruplex affinity and G-quadruplex over duplex selectivity,^[23,24] was also tested for comparison. First, the cytotoxic activity of the compound was evaluated in vitro on different human cancer cell lines (HepG2, MCF-7, HeLa, WM266, Jurkat) and normal human dermal fibroblasts (HDF). The experiments, performed at 10 μM ligand concentration for 24 h, showed an interesting cytotoxic activity on all the cancer cell lines examined (Figure 4), with a higher effect (~30% inhibition) observed on Jurkat cells (treated with 10 or 25 μM ligand concentration), where BCL2 is known to play a critical role in regulating the survival and tumorigenicity of cells.^[25] Remarkably, **3** did not exhibit appreciable activity on normal cells. In comparison, RHPS4, used here as a positive control, turned out to be more active on all tested tumor cell lines, but it also exhibited a certain degree of toxicity on normal cells. This was an expected result in view of its high affinity for G-quadruplexes, which, coupled with the lack of selectivity for a specific G-quadruplex structure, leads to strong, but more indiscriminate cytotoxicity.

To determine whether **3** induces apoptosis, Jurkat cells were treated with the compound (and with RHPS4 for comparison) and, after 16 h, apoptosis was evaluated with annexin V-FITC/PI double staining by flow cytometry analysis.^[26] Results of this experiment (Figure 5) showed that the cells treated with **3** exhibited 30% early apoptotic cells with respect to the control; in contrast, for RHPS4, the entire population of cells was found to be in the late apoptotic state. To determine the effect of **3** on BCL2 gene transcription, quantitative analysis of mRNA was carried out by means of real-time PCR. Upon treatment of Jurkat cells with 25 μM of compound **3** for 24 h, a significant decrease (~24%) in BCL2 mRNA was observed (Figure 5). In parallel, to confirm that the observed decrease in mRNA levels leads to a decrease in Bcl-2 protein, western blot analysis was performed by using an anti-Bcl-2 antibody. These experiments showed that treatment with **3** results in a substantial decrease (~40%) in the level of Bcl-2 protein, in agreement with results

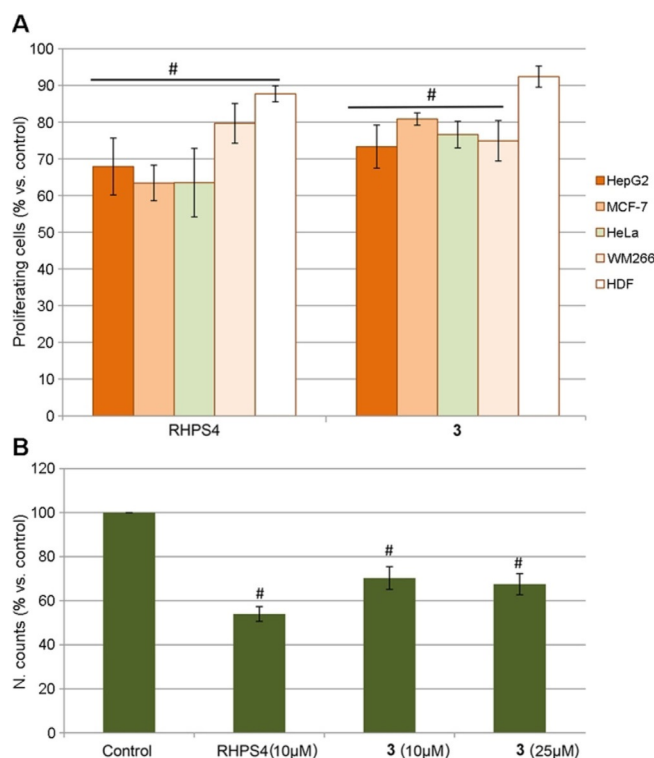


Figure 4. Cytotoxic effect of RHPS4 and **3** on A) HepG2, MCF-7, HeLa, WM266, and HDF cell lines treated with 10 μM ligand concentration for 24 h, and B) Jurkat cells treated with the reported concentrations for 24 h. The results are presented as the percentage of proliferating cells with respect to control (vehicle-treated cells). All data are mean values \pm SEM of at least three independent experiments performed in triplicate ($^{\#}p < 0.05$).

obtained by qPCR. Similar results were obtained in both analyses with RHPS4 (Figure 5).

In summary, a series of furo[2,3-*d*]pyridazin-4(5*H*)-one derivatives were evaluated as G-quadruplex-targeting compounds. Among them, two derivatives were shown to selectively stabilize the G-quadruplex-forming sequence derived from the promoter region of the BCL2 gene. Moreover, one of such ligands effectively down-regulated the transcription and expression of BCL2 in Jurkat cells, inducing a significant increase of the percentage of cells in an early apoptotic state. All these experimental results reinforced the idea of inhibition of Bcl-2 expression through the specific targeting of BCL2 G-quadruplex DNA and provided clear evidence that furo[2,3-*d*]pyridazin-4(5*H*)-one derivatives represent a new class of drug-like compounds useful to this aim, thus stimulating further studies aimed at developing more potent derivatives as effective pathway-specific anticancer agents.

Acknowledgements

The authors thank Dr. Annalisa Tito (Arterra Bioscience, Naples, Italy) for providing human dermal fibroblasts (HDF), and Dr. Gentilcore (IRCCS-Fondazione Pascale, Naples, Italy) for providing the human metastatic melanoma cell line (WM266). This work was

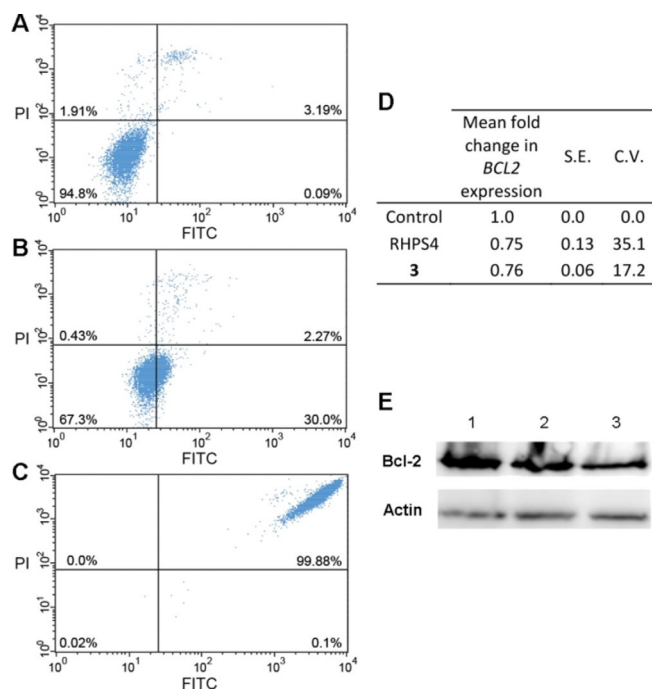


Figure 5. Apoptosis analysis with annexin V-FITC/PI double staining on Jurkat cells treated with A) vehicle, B) **3**, and C) RHPS4 (25 μ M). For all panels: upper left quadrants, necrotic cells; upper right, advanced apoptotic cells; lower left, viable cells; lower right, early apoptotic cells. These results are representative of three independent experiments. D) qPCR analysis of *BCL2* mRNA levels in Jurkat cells. Results are presented as fold change in gene expression relative to the reference gene (*GAPDH*). Data were analyzed using the $2^{-\Delta\Delta Ct}$ method.^[27] S.E. = standard error. C.V. = percent coefficient of variation of $2^{-\Delta\Delta Ct}$. E) Western blot analysis of lysates from Jurkat cells treated with (1) vehicle, (2) RHPS4, and (3) **3** (25 μ M) for 24 h. Specific anti-Bcl-2 antibody was used to detect Bcl-2 protein; actin was used as a control. The blot is representative of three experiments with similar results.

supported by the Italian Association for Cancer Research (AIRC; IG 16730 to B.P., IG 18695 to A.R., and MFAG 18793 to M.G.).

Conflict of interest

The authors declare no conflict of interest.

Keywords: antitumor agents · Bcl-2 transcriptional down-regulation · gene promoters · G-quadruplexes · ligands

- [1] M. O. Hengartner, *Nature* **2000**, *407*, 770–776.
 [2] J. M. Adams, S. Cory, *Science* **1998**, *281*, 1322–1326.
 [3] K. W. Yip, J. C. Reed, *Oncogene* **2008**, *27*, 6398–6406.

- [4] M. Seto, U. Jaeger, R. D. Hockett, W. Graninger, S. Bennett, P. Goldman, S. J. Korsmeyer, *EMBO J.* **1988**, *7*, 123–131.
 [5] T. S. Dexheimer, D. Sun, L. H. Hurley, *J. Am. Chem. Soc.* **2006**, *128*, 5404–5415.
 [6] D. J. Patel, A. T. Phan, V. Kuryavyi, *Nucleic Acids Res.* **2007**, *35*, 7429–7455.
 [7] S. Balasubramanian, L. H. Hurley, S. Neidle, *Nat. Rev. Drug Discovery* **2011**, *10*, 261–275.
 [8] J. Dai, D. Chen, R. A. Jones, L. H. Hurley, D. Yang, *Nucleic Acids Res.* **2006**, *34*, 5133–5144.
 [9] J. Jana, S. Mondal, P. Bhattacharjee, P. Sengupta, T. Roychowdhury, P. Saha, P. Kundu, S. Chatterjee, *Sci. Rep.* **2017**, *7*, 40706.
 [10] E. Micheli, A. Altieri, L. Cianni, C. Cingolani, S. Iachettini, A. Bianco, C. Leonetti, S. Cacchione, A. Biroccio, M. Franceschin, A. Rizzo, *Biochimie* **2016**, *125*, 223–231.
 [11] Y. Feng, D. Yang, H. Chen, W. Cheng, L. Wang, H. Sun, Y. Tang, *Bioorg. Med. Chem. Lett.* **2016**, *26*, 1660–1663.
 [12] X. D. Wang, T. M. Ou, Y. J. Lu, Z. Li, Z. Xu, C. Xi, J. H. Tan, S. L. Huang, L. K. An, D. Li, L.-Q. Gu, Z.-S. Huang, *J. Med. Chem.* **2010**, *53*, 4390–4398.
 [13] S. Neidle, *J. Med. Chem.* **2016**, *59*, 5987–6011.
 [14] B. Pagano, J. Amato, N. Iaccarino, C. Cingolani, P. Zizza, A. Biroccio, E. Novellino, A. Randazzo, *ChemMedChem* **2015**, *10*, 640–649.
 [15] J. Amato, N. Iaccarino, B. Pagano, R. Morigi, A. Locatelli, A. Leoni, M. Rambaldi, P. Zizza, A. Biroccio, E. Novellino, A. Randazzo, *Front. Chem.* **2014**, *2*, 54.
 [16] M. Asif, *Curr. Med. Chem.* **2012**, *19*, 2984–2991.
 [17] M. E. Prime, S. M. Courtney, F. A. Brookfield, R. W. Marston, V. Walker, J. Warne, A. E. Boyd, N. A. Kairies, W. Von Der Saal, A. Limberg, G. Georges, R. A. Engh, B. Goller, P. Rueger, M. Rueth, *J. Med. Chem.* **2011**, *54*, 312–319.
 [18] M. Giustiniano, V. Mercalli, J. Amato, E. Novellino, G. C. Tron, *Org. Lett.* **2015**, *17*, 3964–3967.
 [19] M. Giustiniano, F. Meneghetti, V. Mercalli, M. Varese, F. Giustiniano, E. Novellino, G. C. Tron, *Org. Lett.* **2014**, *16*, 5332–5335.
 [20] J. Amato, R. Morigi, B. Pagano, A. Pagano, S. Ohnmacht, A. De Magis, Y. P. Tiang, G. Capranico, A. Locatelli, A. Graziadio, A. Leoni, M. Rambaldi, E. Novellino, S. Neidle, A. Randazzo, *J. Med. Chem.* **2016**, *59*, 5706–5720.
 [21] B. Pagano, S. Cosconati, V. Gabelica, L. Petraccone, S. De Tito, L. Marinelli, V. La Pietra, F. S. di Leva, I. Lauri, R. Trotta, E. Novellino, C. Giancola, A. Randazzo, *Curr. Pharm. Des.* **2012**, *18*, 1880–1899.
 [22] C. Giancola, B. Pagano, *Top. Curr. Chem.* **2013**, *330*, 211–242.
 [23] M. Cheng, C. Modi, J. C. Cookson, I. Hutchinson, R. A. Heald, A. J. Mccarroll, S. Missailidis, F. Taniou, W. D. Wilson, J.-L. Mergny, C. A. Loughton, M. F. G. Stevens, *J. Med. Chem.* **2008**, *51*, 963–975.
 [24] E. Gavathiotis, R. A. Heald, M. F. G. Stevens, M. S. Searle, *Angew. Chem. Int. Ed.* **2001**, *40*, 4749–4751; *Angew. Chem.* **2001**, *113*, 4885–4887.
 [25] J. C. Reed, M. Cuddy, S. Haldar, C. Croce, P. Nowell, D. Makover, K. Bradley, *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 3660–3664.
 [26] D. Capasso, I. de Paola, A. Liguoro, A. Del Gatto, S. Di Gaetano, D. Guarnieri, M. Saviano, L. Zaccaro, *PLoS One* **2014**, *9*, e106441.
 [27] K. J. Livak, T. D. Schmittgen, *Methods* **2001**, *25*, 402–408.

Manuscript received: December 4, 2017

Revised manuscript received: January 17, 2018

Accepted manuscript online: January 18, 2018

Version of record online: February 9, 2018



Research paper

Tailoring a lead-like compound targeting multiple G-quadruplex structures

Jussara Amato^{a,1}, Chiara Platella^{b,1}, Sara Iachettini^c, Pasquale Zizza^c,
Domenica Musumeci^b, Sandro Cosconati^d, Alessia Pagano^a, Ettore Novellino^a,
Annamaria Biroccio^c, Antonio Randazzo^a, Bruno Pagano^{a,**}, Daniela Montesarchio^{b,*}

^a Department of Pharmacy, University of Naples Federico II, 80131, Naples, Italy

^b Department of Chemical Sciences, University of Naples Federico II, 80126, Naples, Italy

^c Oncogenomic and Epigenetic Unit, IRCCS - Regina Elena National Cancer Institute, 00144 Rome, Italy

^d DiSTABIF, Università della Campania Luigi Vanvitelli, 81100, Caserta, Italy

ARTICLE INFO

Article history:

Received 5 October 2018

Received in revised form

30 October 2018

Accepted 23 November 2018

Available online 27 November 2018

Keywords:

G-quadruplex ligands

Affinity chromatography-based screenings

In vitro biological assays

Biophysical characterization

Molecular docking

ABSTRACT

A focused library of analogs of a lead-like G-quadruplex (G4) targeting compound (**4**), sharing a furanbenzoxazine naphthoquinone core and differing for the pendant groups on the N-atom of the oxazine ring, has been here analyzed with the aim of developing more potent and selective ligands. These molecules have been tested vs. topologically different G4s by the G4-CPG assay, an affinity chromatography-based method for screening putative G4 ligands. The obtained results showed that all these compounds were able to bind several G4 structures, both telomeric and extra-telomeric, thus behaving as multi-target ligands, and two of them fully discriminated G4 vs. duplex DNA. Biological assays proved that almost all the compounds produced effective DNA damage, showing marked anti-proliferative effects on tumor cells in the low μM range. Combined analysis of the G4-CPG binding assays and biological data led us to focus on compound **S4-5**, proved to be less cytotoxic than the parent compound **4** on normal cells. An in-depth biophysical characterization of the binding of **S4-5** to different G4s showed that the here identified ligand has higher affinity for the G4s and higher ability to discriminate G4 vs. duplex DNA than **4**. Molecular docking studies, in agreement with the NMR data, suggest that **S4-5** interacts with the accessible grooves of the target G4 structures, giving clues for its increased G4 vs. duplex selectivity.

© 2018 Elsevier Masson SAS. All rights reserved.

1. Introduction

Improving selectivity and concomitantly reducing toxicity of chemotherapeutic agents is one of the hottest challenges in anti-cancer research [1–4]. Understanding the molecular mechanisms involved in growth and proliferation of cancer has allowed developing therapies that efficiently act on specific cancer cell processes, not damaging normal cells and thus reducing typical side effects of anticancer drugs [5–10]. Among known drug-target candidates, higher and higher attention is currently paid to G-quadruplex (G4)

DNA structures, which are now recognized to play major roles in carcinogenesis. Indeed, putative G4-forming DNA tracts are widely distributed in the human genome, and particularly represented in regulatory genomic regions responsible for uncontrolled cell proliferation and/or oncogenic transformation [11,12]. Interestingly, even if the highest abundance of G-rich sequences is found at telomeres, the majority (~75%) of G4s are present in extra-telomeric regions [13,14], such as in tumor-related gene promoters like *c-myc*, *c-kit*, *KRAS*, *PDGF-A*, *hTERT*, *Rb*, *RET* and *bcl-2* [15]. Thus, compounds able to target G4 structures and discriminate duplex DNA can in principle act as effective candidate antitumor drugs by interfering with cancer onset and progression pathways without impairing healthy cells. In addition, by simultaneously targeting multiple G4 structures located in regulatory regions of different oncogenes or at telomeres, these binders could act on multiple and different pathways via a unique general mechanism of

* Corresponding author.

** Corresponding author.

E-mail addresses: bruno.pagano@unina.it (B. Pagano), daniela.montesarchio@unina.it (D. Montesarchio).

¹ These authors contributed equally to this work.

action, thus resulting into valuable synergistic effects [16].

With a view to expanding the repertoire of available drug-like G4 binders and speeding up the search for true hits, computer-aided virtual screenings can be a valid tool, especially for analyzing large libraries of putative ligands [17]. In this frame, some of us previously performed a virtual screening of a 6000 compounds library – obtained from a commercially available database – against the tetramolecular, parallel G4 formed by the d(TGGGGT) sequence [18]. Among these compounds, six were found to be good G4 binders (Fig. S1), as also confirmed by NMR studies. These ligands were subsequently studied in their interaction with different G4s using a variety of biophysical techniques [19]. Remarkably, two of these compounds (2 and 4, Fig. S1) showed selective G4-stabilizing properties, as well as ability to induce DNA damage response and telomere-dysfunction-induced foci (TIFs) in the μM range. In particular, compound 4 (2-cyclohexyl-2,3-dihydro-1H-naphtho[2',3':2,3]benzofuro[7,6-e][1,3]oxazine-7,12-dione) was shown to effectively stabilize multiple G4s (both telomeric and extra-telomeric), also inducing G4 formation in some of the investigated G-rich oligonucleotides [19].

Encouraged by the promising activity of compound 4, and with the aim of developing more potent and selective ligands efficiently targeting multiple G4 structures, here we identified and analyzed a small focused library of its structural analogs, featured by different pendant groups on the N-atom of the oxazine ring. First, the selected molecules have been tested vs. topologically different G4s by means of the G4-CPG assay, an affinity chromatography-based method we have recently developed for the rapid screening of putative G4 ligands [20]. Then, the G4-binding properties of the most promising analogs of this library – according to the results of the G4-CPG assay – have been investigated in detail employing both biophysical and biological tools.

2. Results and discussion

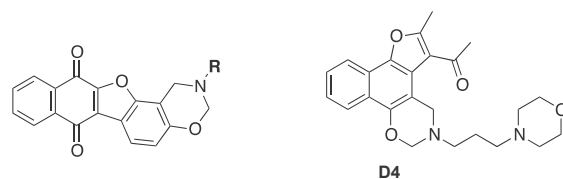
2.1. Selection of compound 4 analogs

A convenient strategy for the discovery of G4-selective ligands with enhanced activity/toxicity ratio can be to start from a known lead compound featured by a suitable core, and then functionalize it with a set of different decorations. In this way, the interactions with the target G4s ensured by the core are preserved, and improved affinity (and/or selectivity) can be achieved by exploiting additional interactions (*i.e.*, electrostatic, stacking, and/or hydrogen bonding interactions) realized by the decorations. Following this strategy, using 4 as the lead compound we here aimed at identifying a set of its analogs featured by the furobenzoxazine naphthoquinone core and bearing different pendant groups on the oxazine N-atom.

Given the relative synthetic accessibility of pentacyclic furobenzoxazine naphthoquinones, we first searched for analogs of 4 available in commercial molecular databases. This approach has the advantage of quickly providing a library of structural analogs of the lead compound and being cost-effective. Therefore, the Dice similarity coefficient was computed between 4 and the compounds present in the ZINC database collection of commercially available compounds (<https://zinc15.docking.org>) setting the similarity threshold to 70%. This search resulted in 40 compounds including, as desired, a number of furobenzoxazine-containing naphthoquinone derivatives, as well as species exhibiting different scaffolds. Ten furobenzoxazine naphthoquinone derivatives obtained from the filtered database were thus selected (S4-1–S4-10, Table 1), along with one furonaphthoxazine analog (D4, Table 1), included in this study as a representative example of a different scaffold.

Table 1

Chemical structures of the here investigated compounds.



Compound	R	Compound	R
4		S4-6	
S4-1		S4-7	
S4-2		S4-8	
S4-3		S4-9	
S4-4		S4-10	
S4-5			

2.2. G4-CPG assay and selection of the model G4-and duplex-forming oligonucleotide sequences

We recently realized a functionalized Controlled Pore Glass (CPG) support, designed for the High Throughput Screening of G4 ligands by affinity chromatography, named G4-CPG [20,21]. This support has been conceived to allow the on-line synthesis of the G4-forming oligonucleotides, as well as the successive binding assays with the potential ligands, providing quick, reliable results. Indeed, in contrast to commercially available supports previously used for affinity chromatography of oligonucleotides [22–24], the new derivatized CPG has low-to-null unspecific interactions with G4 ligands. Thus the binding assays on the G4-CPG provide clean and reproducible data, well reflecting the affinity trend of the tested molecules vs. G4 targets in solution [20,21].

In brief, the G4-CPG assay is based on affinity chromatography screenings, realized on a suitably modified CPG support functionalized with the G-rich oligonucleotide of interest, reproducing the target G4-forming DNA tract. Being bound to the CPG support *via* a long and flexible linker, it can fold in its preferred conformation under various solution conditions and retain high-affinity ligands, thus allowing their easy separation and identification. Indeed, using a fluorescent ligand differently responding to the interaction with various G4 topologies [25], we already proved that the G4-forming oligonucleotides linked to the CPG support preserve their native conformation as in solution [20].

Thus, aiming at expanding the use of the G4-CPG assay to a wide variety of topologically different G4 structures, we here functionalized the CPG support with a set of biologically relevant G4-forming oligonucleotides. In detail, we selected the following DNA sequences: a) the 26-mer $5'(\text{TTAGGG})_4\text{TT}3'$ (tel26) as a well-studied model of the 3'-overhang of the human telomeric DNA [26]; b) the 33-mer $5'\text{TGGGGAGGGTGGGGAGGGTGGGGAA-GGTGGGA}3'$ (cmcy) from the nuclease hypersensitive element in

the human *c-myc* promoter, containing six G-rich tracts of unequal length and characterized by a remarkable polymorphism [27,28], of which typically only shorter and modified variants have been used as models [29]; c) the 22-mer $5'AGGGAGGGCGCTGGGAGGAGGG^{3'}$ (ckit1) - of which both the NMR [30] and crystal structures [31,32] have been solved revealing a unique topology - and the 21-mer $5'CGGGCGGGCGGAGGGAGGG^{3'}$ (ckit2) - which, depending on the K^+ concentration, adopts two distinct parallel-stranded G4 scaffolds in solution [33] - both from the human *c-kit* promoter; d) the 24-mer $5'AAGGGAGGGGCTGGGAGGGCCCGA^{3'}$ (hTERT1) from the G-rich region in the hTERT core promoter, which represents the main regulatory element of the *hTERT* gene [34,35]. Furthermore, in order to probe the ability of the studied compounds to discriminate G4 vs. duplex DNA, the CPG support has been also functionalized with a 27-mer here named ds27 [20,22]. Consisting of two Dickerson 12-mer tracts, $5'CGCGAATTCGCG^{3'}$, i.e. one of the best characterized models of B-DNA duplex [36], connected by a TTT loop, ds27 can in fact fold into a stable hairpin duplex and has been thus selected as a representative model for a duplex DNA tract.

2.3. Experimental Screening by the G4-CPG assay

The binding assays on the functionalized CPG supports have been performed following previously reported procedures [20,22,23], with minimal modifications as described in the Experimental Section. All the oligonucleotide-functionalized supports have been first tested in their ability to bind a set of known G4 ligands with different affinity for G4s. In particular, thiazole orange (TO) and resveratrol have been used as representative molecules, being respectively a strong and a very weak G4 binder [37,38]. The obtained data, showing high affinity (bound ligand >97%) for TO and very low affinity (bound ligand <4%) for resveratrol on all the tested supports, verified the efficacy of the functionalized CPG supports.

A stock solution of each ligand has been prepared by dissolving a weighed amount of the solid compound in pure DMSO; only compound **S4-1** proved to be not completely soluble in DMSO at the concentration required for the binding assays and was therefore discarded. All the other compounds as well as compound **4** proved to be well soluble in the washing/releasing solutions used in the binding assays (see Experimental Section) and at the concentration chosen for the binding experiments. After solubility checks, we verified the absence of unspecific binding on the solid support by incubating the tested compounds with the non-functionalized CPG (here named nude CPG, Fig. S2). Low-to-null unspecific interactions with the nude solid support have been observed (Table 2), thus

further proving that nude CPG is essentially inert in these assays, as demonstrated for known G4 ligands [20], and confirming the general reliability of the method.

Finally, all the new putative ligands have been tested on the G4s- and hairpin duplex-functionalized CPG supports. The results of the binding assays, summarized in Table 2 (see also Fig. S3 in the Supplementary Material), showed that all the analyzed ligands had good affinity for tel26 immobilized on the solid support. However, no marked improvement was found with respect to **4**, still the best binder in the series when tested on this G4-forming sequence, with the only exception of **S4-4**, showing comparable results as **4**. In contrast, when tested on *cmyc*, compounds **S4-2**, **S4-4**, **S4-7**, **S4-8** and **S4-10** displayed higher affinity than **4**. As far as ckit1 is concerned, almost all the tested compounds proved to be stronger binders than compound **4**, with the sole exception of **S4-3**. In the case of ckit2, **S4-2**, **S4-4**, and **S4-10** showed higher binding abilities than **4**. Notably, also in this case compound **S4-3** was the weakest binder of the series. Finally, for hTERT1 all the tested analogs had comparable or higher affinity than **4**.

Noteworthy, all the examined compounds were found to be effective G4 ligands, some of which even better than lead compound **4**. Among all the tested compounds, **D4** was the worst G4 ligand, evidencing that its polycyclic core, different from the other tested molecules, is not suitable for G4 targeting. On the other hand, compound **S4-4** was the strongest G4 binder of the series, able to tightly interact with all the investigated G4s. Overall, although the here screened compounds are able to bind all the investigated G4s, thus behaving as multi-target ligands, the highest percentages of binding have been found for *cmyc*, emerging as the best target of the series.

In order to evaluate the G4s vs. duplex DNA selectivity, all the compounds have been also tested in their interaction with ds27, i.e. the hairpin duplex-forming oligonucleotide. Overall, most of the ligands proved to effectively discriminate G4- vs. duplex-forming oligonucleotides. In particular, **S4-4** and **S4-5** showed a significantly lower affinity for ds27 than **4**. Remarkably, **S4-5** gave null affinity for duplex DNA, thus emerging as the most promising analog in terms of G4 vs. duplex selectivity on the basis of this binding assay.

2.4. Biological activity studies

Following the binding assays, the G4 ligands have been subjected to biological analyses aimed at identifying derivatives with potent and selective DNA damage response (DDR) and antitumoral activity [39,40]. Thus, BJ-EHLT cells - a human foreskin-derived fibroblast cell line, expressing the human telomerase reverse

Table 2
Binding assay results for the investigated compounds on nude and functionalized CPG supports.

Compound	Bound ligand (%) ^a						
	Nude CPG	CPG-tel26	CPG-cmyc	CPG-ckit1	CPG-ckit2	CPG-hTERT1	CPG-ds27
4	0	86	86	70	82	74	18
S4-2	0	80	94	85	91	82	46
S4-3	0	73	84	55	67	73	16
S4-4	0	86	93	90	92	91	2
S4-5	0	80	84	78	80	82	0
S4-6	10	80	87	76	83	78	12
S4-7	11	80	89	77	84	92	33
S4-8	0	78	90	88	83	87	20
S4-9	5	75	73	72	68	76	16
S4-10	0	79	94	90	88	89	28
D4	7	76	74	75	70	78	22

^a Bound ligand is calculated as a difference from the unbound ligand, recovered with 50 mM KCl/10% DMSO/10% ethanol washing solution, and expressed as % of the amount initially loaded on the support. Errors associated with the % are within $\pm 2\%$.

transcriptase (hTERT) and SV40 early region – were treated for 24 h with 0.5 μ M of each compound and the DNA damage was evaluated. To appreciate possible improvements in terms of biological activity of the analogs over lead compound **4**, we used a time and dose that had been previously optimized for **4**, able to induce a slight DNA damage [19]. Immunofluorescence (IF) analysis showed relevant amounts of the phosphorylated form of histone H2AX (γ H2AX), a hallmark of DNA double-strand breaks [41], evidencing that all the tested ligands – with the sole exception of **D4**, **S4-9**, and **S4-10** – induced similar DNA damage extent as the lead compound (Fig. 1A). In parallel, we also evaluated whether the induced DNA damage – or at least part of it – was located at the telomeres. To address this issue, we performed co-staining IF experiments using antibodies against γ H2AX and TRF1 – an effective marker for interphase telomeres – and the co-localization spots, reported as Telomere Induced Foci (TIF) [42], were evaluated (Fig. 1B).

Notably, quantitative analysis confirmed and reinforced the previous observations, evidencing that all the compounds (with the only exception of **D4**, **S4-9**, and **S4-10**) induced a percentage of TIF positive cells (cells with at least 4 co-localization spots) and average number of TIFs per cell similar to lead compound **4** (Fig. 1C). These

data are consistent with the comparable affinities of compound **4** and its analogs towards the telomeric model sequences found using the G4-CPG assay.

Next, to test the antitumor efficacy of compound **4** analogs, the cell colony-forming ability of the human cervical cancer cells, HeLa, untreated or treated with the different compounds, was evaluated. To set up the optimal dose of the drug, we first performed a dose-response experiment by using three different concentrations (0.5, 1 and 2 μ M) of **4**; the 2 μ M concentration proved to be the minimal effective dose and was thus selected for testing all the other derivatives (Fig. 2A). Interestingly, cell survival experiments evidenced that derivatives **S4-4**, **S4-6**, **S4-8**, and **S4-10**, despite their high affinity for G4 structures *in vitro* (Table 2), were less effective than **4** in affecting tumor cell viability (Fig. 2B and C). Compounds **D4** and **S4-3** also showed an antiproliferative activity lower than compound **4** (Fig. 2B and C), which could be related, in this case, to an overall G4-affinity lower than compound **4** (Table 2). Conversely, the potent cytotoxic activity observed for derivatives **S4-2**, **S4-7** and **S4-9** (Fig. 2B and C) might be attributed to G4-independent, off-target effects since these ligands have shown either a good affinity also for duplex DNA (**S4-2** and **S4-7**) or, in the case of **S4-9**, an

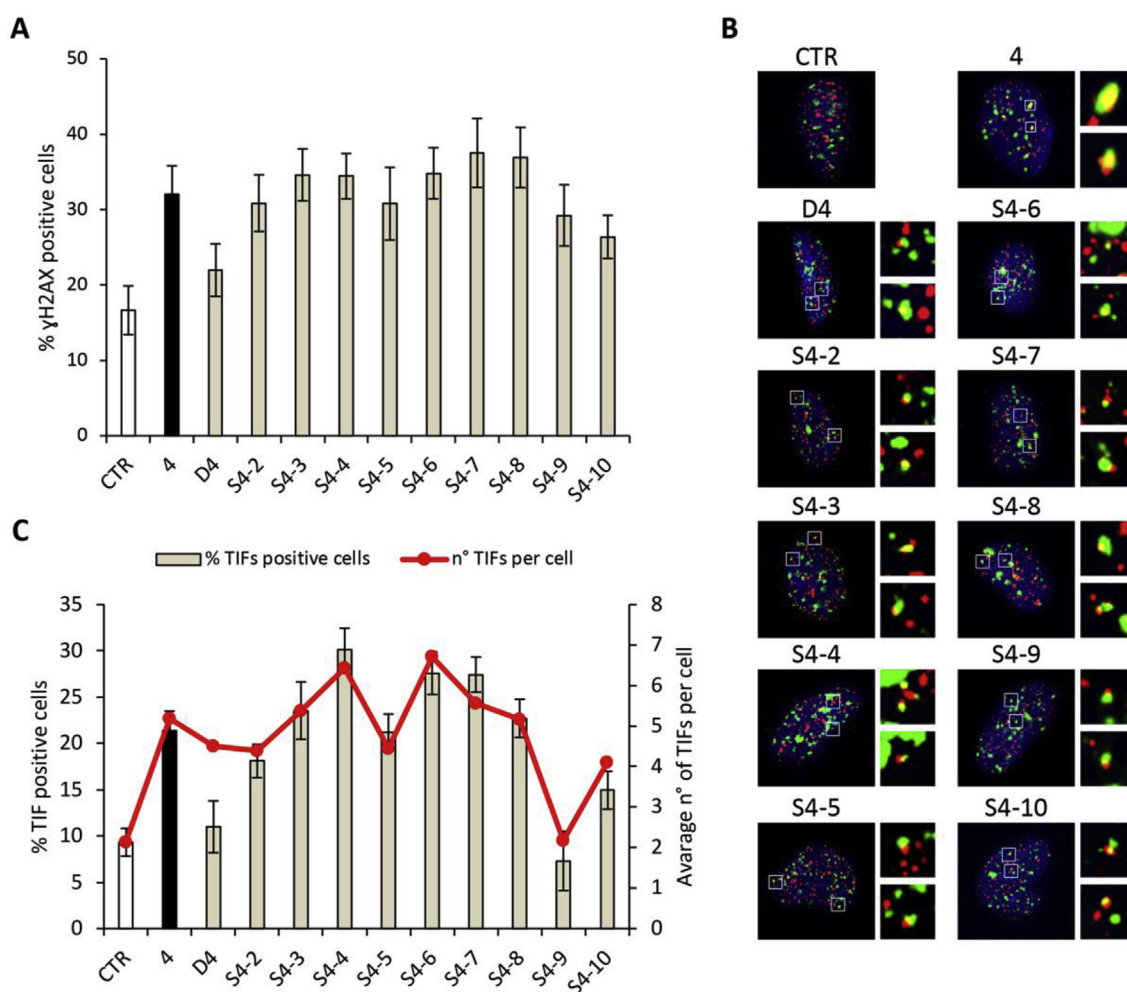


Fig. 1. Compound **4** analogs induce telomeric DNA damage. BJ-EHLT fibroblasts were untreated (CTR, white bar) or treated for 24 h with **4** (black bar) and the indicated analogs (light-grey bars) at 0.5 μ M concentration. Cells were processed for immunofluorescence (IF) using antibodies against γ H2AX and TRF1 to visualize the DNA damage and telomeres, respectively. (A) Percentages of γ H2AX-positive cells. (B) Representative merged images of IF of untreated and treated BJ-EHLT cells; γ H2AX spots in green, TRF1 spots in red and nuclei in blue. Enlarged views of Telomere Induced Foci (TIFs) are reported on the right panels of each picture. The images were acquired with a Leica Deconvolution microscope (magnification 63x). (C) Quantitative analysis of TIFs. The graph represents the percentages of TIF-positive cells (bars) and the mean number of TIFs per cell (red line) in the indicated samples. Cells with at least four γ H2AX/TRF1 foci were scored as TIF positive. Histograms show the mean values \pm S.D. of three independent experiments. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

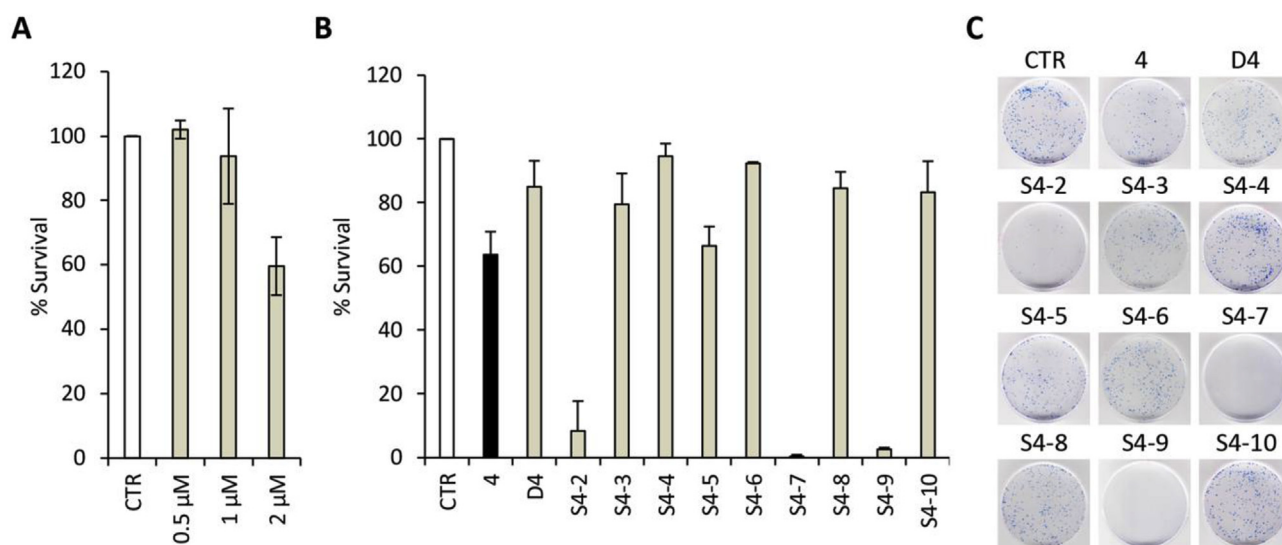


Fig. 2. Anti-tumor efficacy of compound **4** analogs. (A) Clonogenic activity of human cervical cancer cells, HeLa, untreated (CTR, white bar) or treated with **4** at the indicated doses. Surviving fractions were calculated as the ratio of absolute survival of the treated sample/absolute survival of the untreated sample. (B) Clonogenic activity of HeLa cells, untreated (CTR, white bar), treated with **4** (black bar) or the indicated analogs (light-grey bars) at 2 μM dose. Surviving fractions were calculated as reported in A. (C) Representative images of the clonogenic assay described in B. Histograms show the mean values \pm S.D. of three independent experiments.

affinity towards G4 structures generally lower than other analogs (Table 2). In summary, ligand **S4-5** emerged as the most interesting ligand of the series. Indeed, its high affinity for G4 – despite an antitumor activity only slightly superior to lead compound **4** – associated with null affinity for duplex DNA (Table 2) suggests a promisingly improved target specificity compared to the starting compound.

Next, in order to evaluate the selectivity of **S4-5** in cell, the effects of this compound have been tested also on normal cells. Briefly, human immortalized BJ fibroblasts (BJ-hTERT) have been treated with **4** or **S4-5** at 1 and 2 μM concentrations and, after 6 days, the number of viable cells has been evaluated. Notably, these experiments clearly evidenced that the cytotoxic effects of ligand **S4-5** on normal cells are definitely lower than those produced by compound **4** (Fig. 3A). Moreover, IF analysis of γ H2AX evidenced that **S4-5**, in contrast to **4**, was unable to induce an appreciable DNA

damage in normal cells (Fig. 3B and C). Overall, these data unambiguously indicated that **S4-5** has a much higher selectivity in killing cancer cells than compound **4**, making it a promising candidate drug.

2.5. CD and CD-melting experiments

Based on the above discussed results, we investigated by CD experiments in solution the ability of **S4-5** to interact with a human telomeric G4 (tel26), an extra-telomeric G4 (cmyc), and a duplex structure (ds27), in comparison with **4**. In particular, among the extra-telomeric G4s here studied, we chose cmyc since **S4-5**, as well as the other here tested ligands, has shown a slight preference for this G4 structure compared to the others.

CD-monitored titrations have been performed by adding increasing amounts of the tested ligands to samples of these

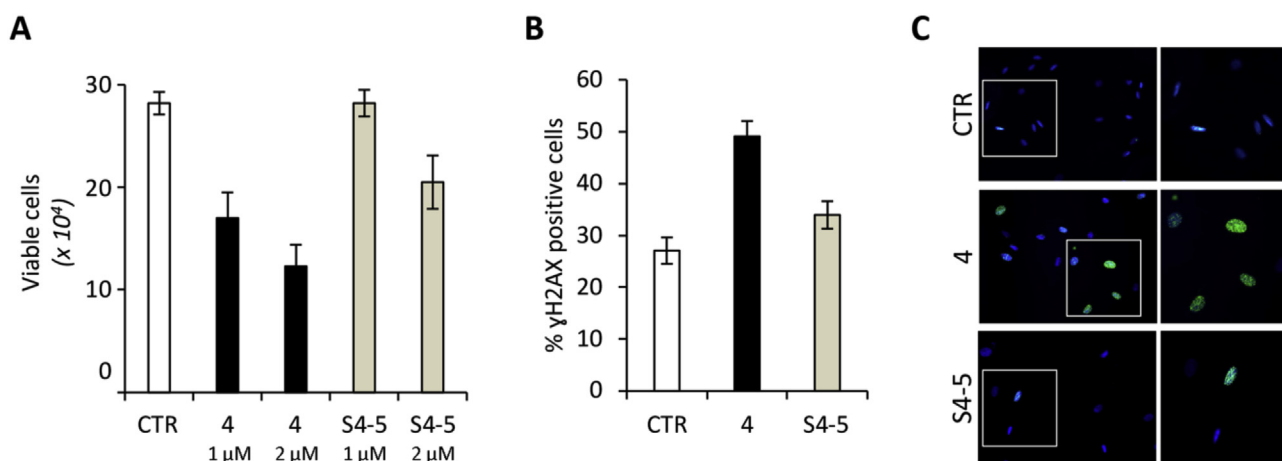


Fig. 3. **S4-5** analog does not affect normal cells. (A) Human immortalized fibroblasts (BJ-hTERT) were treated with compound **4** or derivative **S4-5** at the doses of 1 and 2 μM for 6 days. Viable cell number was determined by the Trypan Blue exclusion test. (B) BJ-hTERT were treated with **4** or **S4-5** at the 1 μM dose for 24 h and processed for IF using antibodies against γ H2AX. The histogram represents the percentages of γ H2AX-positive cells. (C) Representative merged images of IF: γ H2AX spots in green and nuclei in blue. Histograms show the mean values \pm S.D. of three independent experiments. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

oligonucleotides (2 μ M) in 20 mM KCl/5 mM KH₂PO₄/10% DMSO buffer (pH 7). As expected in these buffer conditions, tel26 folded into a hybrid II-type G4 featured by a maximum at 290 nm and a shoulder at 270 nm [43], cmc formed a parallel G4 with a maximum at around 263 nm [28,29], while ds27 showed the characteristic positive band at 280 nm accompanied by a minimum at 251 nm, typical of a B-DNA duplex structure (Fig. 4) [44]. In the case of tel26, an increase of the intensity of the 290 nm band along with the reduction of the shoulder at 270 nm have been observed upon adding increasing amounts of **S4-5** in solution (Fig. 4A), thus suggesting that interactions with the ligand took place causing a conformational rearrangement of the G4. In the titration experiments with cmc, the addition of **S4-5** produced only weak changes, with a slight reduction of the 263 nm band and an increase of the 288 nm signal (Fig. 4B). In contrast, in the titration of ds27 no detectable change in the CD profile of the duplex has been observed, suggesting that this structure was essentially unaffected by the addition of even a large excess of ligand (Fig. 4C). A similar behavior has been found in the titrations of both the G4s and the duplex with **4** (Fig. S4). In all cases, up to 5 equivalents of ligand have been added to the oligonucleotide solutions, as the best compromise between solubility of the analyzed ligands and saturation of the oligonucleotide CD signals. No induced CD signal has been detected for all the investigated systems in the range 320–800 nm (Fig. S5).

CD-melting experiments have been performed for all the oligonucleotide/ligand mixtures to evaluate if stabilizing or destabilizing effects on the G4 and hairpin duplex structures were obtained upon incubation with **4** and **S4-5** (Figs. S4 and S6, respectively). CD melting curves of tel26, cmc and ds27 in the absence or presence of each ligand (at 5:1 ligand/DNA ratio) have been recorded by following the CD signal changes at the wavelength of their intensity maximum (290, 263, and 251 nm for tel26, cmc and ds27, respectively). The results of the CD melting experiments showed that compounds **S4-5** and **4** did not significantly affect the stability of the tel26 and ds27 structures (Table 3), even if for tel26 the drop of CD signal between 20 and 90 °C was much higher in the presence of the ligands than in their absence (Figs. S4 and S6). Conversely, both ligands appreciably stabilized the cmc G4 structure ($\Delta T_m > 8$ °C).

Overall, these results confirmed the ability of the here investigated pentacyclic scaffold to significantly stabilize parallel G4s, not affecting the thermal stability of other G4 conformations and duplex DNA [19].

2.6. NMR experiments

NMR titration experiments were performed to get information

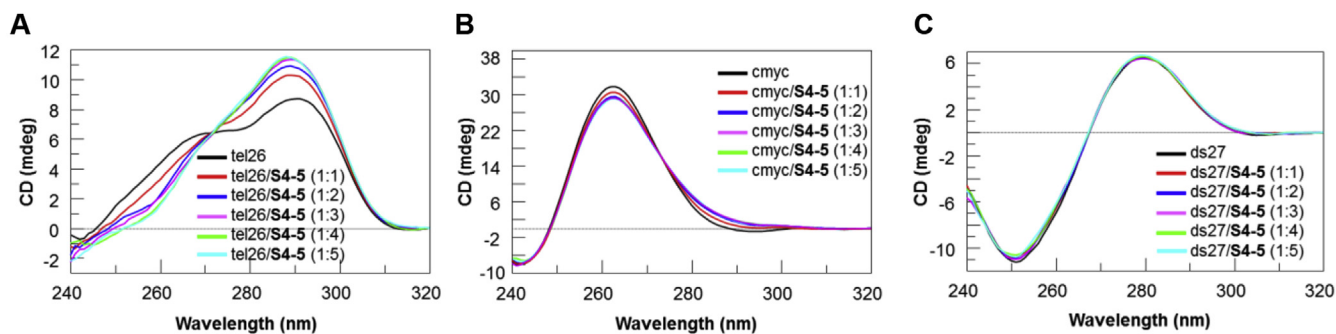


Fig. 4. Effect of **S4-5** on the overall conformation of the DNA structures. CD spectra of 2 μ M solutions of tel26 (A), cmc (B) and ds27 (C) in 20 mM KCl/5 mM KH₂PO₄/10% DMSO buffer (pH 7) in the absence and presence of increasing amount of **S4-5** (up to 5 equivalents).

Table 3

Melting temperatures (T_m) of tel26, cmc and ds27 in the absence or presence of ligands (5 equiv) measured by CD melting experiments.

Ligand	T_m (°C) (± 1)		
	tel26	cmc	ds27
No ligand	49	82	75
4	49	>90	76
S4-5	50	>90	76

on the binding mode of compounds **4** and **S4-5** to the G4 structures formed by the human telomeric and *c-myc* promoter DNA sequences [45]. The modified telomeric sequence 5' TTGGG(TTAGGG)₃A^{3'} (m-tel24) and the modified *c-myc* sequence 5' TGAGGGTGGGTAGGGTGGGTA^{3'} (myc22) were used for this study, since they give higher quality NMR spectra compared to the wild-type sequences, being characterized by a single conformation in solution [46,47]. The ¹H NMR spectra of m-tel24 and myc22 well matched the ones reported in the literature, characterized by 12 well-resolved imino protons peaks, corresponding to the 12 guanines involved in the three G-tetrad planes (Figs. S7–S10) [46,47]. The investigated oligonucleotides were then titrated with **4** and **S4-5**, in parallel experiments, up to a 4:1 ligand/G4 ratio. Upon addition of increasing amounts of ligands to the G4 solutions, considerable proton resonance changes were observed in both imino and aromatic proton regions of the spectra (Figs. S7–S10), thus suggesting that both **4** and **S4-5** may not have a unique binding site. The relative line broadening observed for all the imino resonances and most of the aromatic ones also suggests that the ligands could explore different poses in the same G4 binding pocket, in fast exchange with each other. On the other hand, some distinct aromatic protons of both m-tel24 and myc22 G4s were only weakly affected by the ligand. As far as **4** is concerned, for m-tel24 (Fig. S7) this is the case of residues T6-T7-A8, forming the double-chain reversal loop of the G4; while for myc22 (Fig. S8) this is the case of A6, stacking on top of the 5' G-tetrad, and of residues T10, T14-A15, and T19 (numbering according to Ambrus, A. et al.) [47], which form the three double-chain reversal loops of the G4 structure. Regarding **S4-5**, the NMR titrations showed that in the interaction with m-tel24 G4 (Fig. S9) the ligand affected to a lesser extent the aromatic protons of A8, which is in the double-chain reversal loop, and of residues T18-T19-A20, which form a lateral loop close to the external G-tetrad; while in the interaction with myc22 (Fig. S10), the least affected ones were the aromatic protons of A6, which stacks on top of the 5' G-tetrad, and of residues T14-A15 and T19, which form the second and third double-chain reversal loop of G4, respectively.

Overall, these findings suggest that both **4** and **S4-5** may bind

these G4s mainly through interaction with those grooves that are not made inaccessible by the presence of loops, generally affecting also guanine residues involved in the three G-tetrad layers. However, our data do not allow completely excluding ligand stacking on external G-tetrads, while they clearly indicate that loops are generally very little or not affected at all by ligand binding.

NMR experiments helped us to further investigate the selectivity of **4** and **S4-5** for G4 over duplex DNA. The two ligands were tested on a model duplex DNA formed by the so-called Dickerson sequence $5'CGCGAATTTCGCG3'$ (ds12). The ^1H NMR spectrum of ds12 matched the one reported in the literature (Figs. S11 and S12) [48]. Results of ds12 titrations with **4** and **S4-5** clearly showed that no significant variation of the chemical shift values of ds12 was observed upon addition of **S4-5** (Fig. S11), up to a 4:1 ligand/DNA ratio, thus confirming the absence of binding with the investigated duplex DNA. On the other hand, **4** caused a small perturbation of the ds12 spectrum, in terms of both intensity decrease and line broadening of all the ^1H NMR signals (Fig. S12), thus suggesting the possibility that **4** could somehow interact in a weak and unspecific way with duplex DNA. These results are in full agreement with those obtained from the G4-CPG binding assays.

2.7. Microscale thermophoresis (MST) experiments

To obtain quantitative data about the affinity of **S4-5** for the human telomeric and *c-myc* promoter G4s, microscale thermophoresis (MST) measurements were performed. MST is a fast and easy tool to characterize small molecule–nucleic acid interactions in solution [49]. Basically, it records the thermophoretic movement of a fluorescently-labeled target molecule under microscopic temperature gradients. This molecular motion strongly depends on changes in size, charge, and hydration shell. Since the binding of a ligand to the investigated molecule changes at least one of these parameters, it also alters the thermophoretic behavior of the target. This effect can be used to evaluate equilibrium constants, such as the dissociation constant K_d . To this purpose, serial dilutions of **S4-5** were prepared, mixed with a constant concentration of Cy5-labeled oligonucleotides (tel26 or myc22), loaded into capillaries and analyzed by MST. Experiments were run by using 8% final DMSO, in order to avoid aggregation effects observed in the absence of DMSO. The results of these experiments (Fig. 5) showed that compound **S4-5** was able to bind both tel26 and myc22 G4s. The calculated equilibrium dissociation constants indicated for this ligand ca. twice higher affinity to the parallel *c-myc* promoter G4 [$K_d = 13 (\pm 2) \mu\text{M}$] than to the hybrid human telomeric G4 [$K_d = 26 (\pm 4) \mu\text{M}$].

2.8. Molecular docking

To gain major insights into the binding interactions established by **S4-5** with the target G4 DNA structures, we decided to provide a theoretical model for its interaction with human telomeric and *c-myc* promoter G4-forming sequences. In particular, the ligand was docked into the solution structures of *c-myc* G4 (myc22, PDB ID: 1XAV) [47] and into the hybrid-type G4 structure formed in K^+ solution by the 26-mer oligonucleotide from the human telomere (tel26, PDB ID: 2JPZ) [26]. In this study, the software Autodock 4.2 (AD4.2) [50,51] was employed, which had already been used for our previous virtual screening campaign to identify **4** as a G4 ligand [18].

Analysis of the results obtained when **S4-5** was docked into the myc22 structure revealed that the ligand adopts a well-clustered binding in which the compound is inserted into the G4 groove, unoccupied by the loop, where the ligand finds favorable van der Waals interactions with G21, G22, and T23 residues (Fig. 6). This

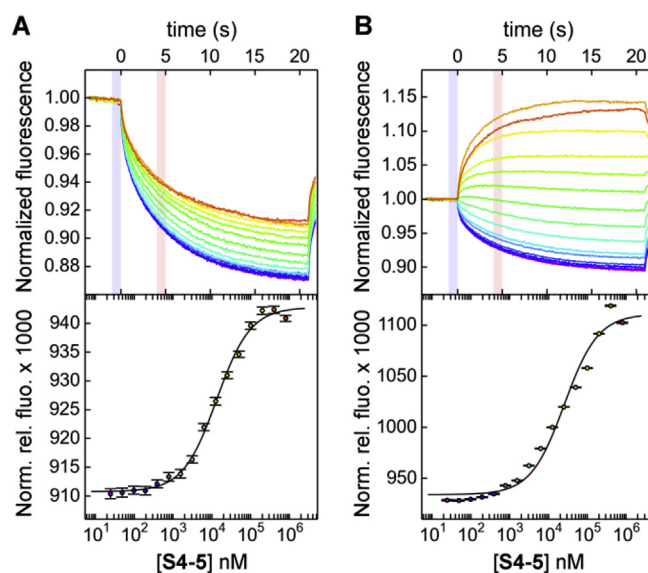


Fig. 5. Interaction of **S4-5** with myc22 (panel A) and tel26 (panel B) G4s studied using microscale thermophoresis experiments. (Top) Time traces recorded by incubating increasing concentrations of **S4-5** with the labeled G4s and (bottom) the corresponding binding curves.

interaction is further reinforced by ionic interactions between the protonated nitrogen of the fused oxazine ring and the backbone phosphate oxygens of A24 residue. Moreover, the pendant benzyl substituent is well positioned to form a parallel displaced π - π interaction with G9 guanine ring. Interestingly, both the (*R*)- and (*S*)-isomer of **S4-5** were docked but no enantiodiscriminating binding was detected. Rather, the methyl substituent on the benzylic carbon seems to optimally orient the phenyl ring to form the aforementioned stacking interactions. Indeed, when comparing the obtained theoretical model with the experimental results of the NMR titrations, a good consistency is recorded, which gives further confidence in the viability of the obtained docking results.

If a well-defined binding pose was detected when docking **S4-5** into the myc22 solution structure, some difficulties were

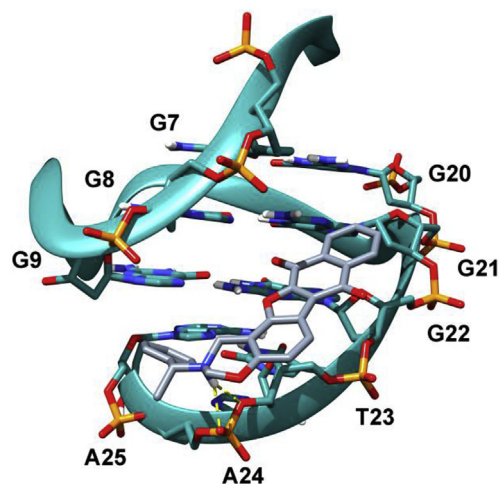


Fig. 6. Binding mode of **S4-5** when docked into the *c-myc* G4 DNA solution structure. The ligand is represented as grey sticks while the DNA as dark cyan ribbons and sticks. H-bonds are represented as dashed yellow lines. DNA residues were numbered according to Ambrus, A. et al. [47]. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

encountered when docking the ligand into the solution structure of the hybrid form of the human telomere DNA G4 structure. More in detail, three possible conformations were suggested by the docking software in which the ligand always occupies the available groove between the first and fourth strand of the G4 structure (Fig. 7). In these three conformations, the ligand adopts the same orientation with respect to the G4 but at different levels of the groove, one towards the 3' end (herein referred to as pose A), one in the middle (herein referred to as pose B), and one towards the 5' end (herein referred to as pose C). In pose A, the ligand establishes a H-bond, through its furan oxygen, with the NH₂ group of G24, and an ionic interaction, through its protonated oxazine nitrogen, with the T25 phosphate group. In pose B, a clear H-bond is formed between the ligand furan oxygen and the G23 NH₂ group. In pose C, one of the two ligand quinone oxygens forms a H-bond with the NH₂ group of the G4 and a charged reinforced H-bond through its protonated oxazine nitrogen and the sugar O4' atom of G22. Moreover, the phenyl ring of the ligand benzyl moiety forms parallel displaced π - π interactions with A3 and G22 aromatic rings. Regardless of their positions in all the three conformations (A, B, and C) the ligand forms van der Waals contacts with the groove atoms of the G4 structure. Moreover, as observed for the myc22 structure, also for tel26 no enantiodiscriminating binding was detected. From the above-described docking results, we cannot give a preference for one of the obtained solutions; rather, it is possible to speculate that the ligand is able to slide into the G4 groove adopting almost iso-energetic binding conformations. This might corroborate the results of the NMR titration experiments, which were not helpful in indicating a specific binding region.

3. Conclusions

In the last decade, increasing efforts in the G4 research field have been devoted to the search of specific G4-binders, able to recognize G4s with high affinity, fully discriminating G4 vs. duplex DNA. Considering the wide distribution and variety of G4 structures within the regulatory regions and telomeres of human genome and their recognized roles in several cancer-related processes, these G4 binders could result into effective multi-target ligands, and expectedly into useful and selective anticancer drugs. Indeed, by simultaneously stabilizing multiple G4 structures at the level of both oncogenes and telomeres, interference with multiple cancer pathways could be achieved, thus resulting into an advantageous approach in innovative anticancer strategies. To reach this ambitious goal, an integrated approach is required, involving the

massive production of large libraries of new putative G4 ligands, coupled with High Throughput Screening methodologies for the fast and reliable analysis of the selected compounds. In this perspective – which is basically inspired to combinatorial approaches for drug discovery – detailed biophysical characterization studies are carried out only after the biological assays, thus devoted only to the best performing G4-ligands, which, following iterative optimization steps, may finally result into promising candidate drugs.

In this work, we have analyzed a focused library of commercially available analogs of a well-characterized G4-binder, here used as a hit compound, differing for the nature of the pendant groups attached on the N-atom of the oxazine ring. The main idea was to explore the effects of a single diversity element inserted on a common planar scaffold, which *per se* acts as a suitable structural motif for G4 recognition. From a methodological point of view, we also aimed at providing a useful working model, here developed as a proof-of-concept, in which our affinity chromatography-based G4-CPG binding assays, coupled with biological screenings, allowed selecting the most promising component of a focused library of putative G4 binders.

The structural analogs of **4** have been screened by our G4-CPG binding assay to identify the most interesting ligands in terms of affinity towards multiple G4 structures, as well as ability to discriminate G4 vs. duplex DNA. Biological screenings on human cancer cells provided a quite varied picture for these compounds, notwithstanding their high structural similarity, with evidence of marked DNA damage and antiproliferative activity for a restricted set of them. Upon analysis of all the available data, one compound, *i.e.* **S4-5**, has been advanced to successive studies for its highly increased selectivity and reduced toxicity compared with the parent compound **4**. The combined use of several biophysical techniques – CD, NMR and MST – allowed evaluating qualitative and quantitative aspects of the binding of **S4-5** towards the target G4 structures, indicating a preference for parallel G4s. Notably, NMR data integrated with molecular modeling studies gave a molecular insight into the interaction of the new identified ligand vs. G4s of different conformations, essentially suggesting for **S4-5** a G4 groove binder ability. Taken together, our results highlighted also the major role played by the pendant groups of selected G4 ligands in fine-tuning their binding preferences and modes to peculiar G4 structures.

Considering that targeting the grooves of G4s could be a successful, even though still poorly explored, approach for the recognition of these structures, and that very few G4-groove binders

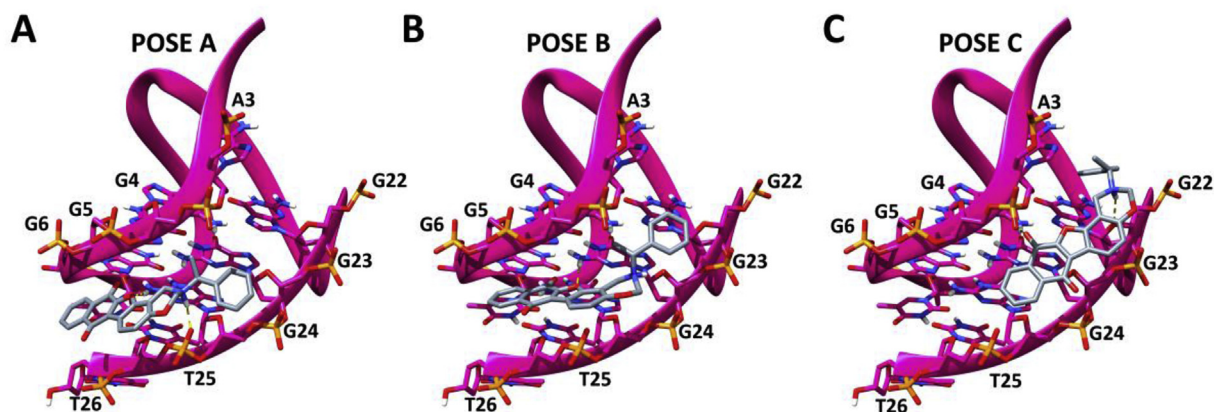


Fig. 7. Alternative binding poses predicted for **S4-5** when docked into the tel26 solution structure. The ligand is represented as grey sticks, the DNA as magenta ribbons and sticks. H-bonds are represented by dashed yellow lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

have been thus far investigated, these results are of great relevance to develop novel effective candidate anticancer drugs. Studies to extend this integrated approach to other classes of putative G4-binders are currently in progress in our laboratories.

4. Experimental section

4.1. Chemistry

All commercial reagents and solvents were purchased from Sigma-Aldrich or Link Technologies, unless otherwise stated. The selected compounds were purchased from Life Chemicals Europe (Germany) and used without further purification. IUPAC name, CAS number, and supplier code of each compound: 2-cyclohexyl-2,3-dihydro-1H-naphtho[2',3':2,3]benzofuro[7,6-e][1,3]oxazine-7,12-dione, 438487-15-1, F1094-0196 (**4**); 2-(*tert*-butyl)-2,3-dihydro-1H-naphtho[2',3':2,3]benzofuro[7,6-e][1,3]oxazine-7,12-dione, 526188-60-3, F1217-0039 (**S4-1**); 2-(3-(1H-imidazol-1-yl)propyl)-2,3-dihydro-1H-naphtho[2',3':2,3]benzofuro[7,6-e][1,3]oxazine-7,12-dione, 438487-37-7, F1094-0208 (**S4-2**); 2-(2-(cyclohex-1-en-1-yl)ethyl)-2,3-dihydro-1H-naphtho[2',3':2,3]benzofuro[7,6-e][1,3]oxazine-7,12-dione, 438487-31-1, F1094-0205 (**S4-3**); 2-(furan-2-ylmethyl)-2,3-dihydro-1H-naphtho[2',3':2,3]benzofuro[7,6-e][1,3]oxazine-7,12-dione, 438487-28-6, F1094-0204 (**S4-4**); 2-(1-phenylethyl)-2,3-dihydro-1H-naphtho[2',3':2,3]benzofuro[7,6-e][1,3]oxazine-7,12-dione, 438487-23-1, F1094-0201 (**S4-5**); 2-benzyl-2,3-dihydro-1H-naphtho[2',3':2,3]benzofuro[7,6-e][1,3]oxazine-7,12-dione, 438487-20-8, F1094-0200 (**S4-6**); 2-(3-morpholinopropyl)-2,3-dihydro-1H-naphtho[2',3':2,3]benzofuro[7,6-e][1,3]oxazine-7,12-dione, 380458-24-2, F1094-0209 (**S4-7**); 2-((tetrahydrofuran-2-yl)methyl)-2,3-dihydro-1H-naphtho[2',3':2,3]benzofuro[7,6-e][1,3]oxazine-7,12-dione, 446269-64-3, F1061-0006 (**S4-8**); 2-(1,1-dioxidotetrahydrothiophen-3-yl)-2,3-dihydro-1H-naphtho[2',3':2,3]benzofuro[7,6-e][1,3]oxazine-7,12-dione, 380458-20-8, F1094-0199 (**S4-9**); 2-(pyridin-3-ylmethyl)-2,3-dihydro-1H-naphtho[2',3':2,3]benzofuro[7,6-e][1,3]oxazine-7,12-dione, 380458-19-5, F1094-0203 (**S4-10**); 1-(6-methyl-3-(3-morpholinopropyl)-3,4-dihydro-2H-furo[3',2':3,4]naphtho[2,1-e][1,3]oxazin-5-yl)ethan-1-one, 438486-98-7, F1094-0190 (**D4**). All the compounds were analyzed by ¹H NMR for purity control.

4.2. G4-CPG assay

Long Chain AlkylAmine-CPG (LCAA-CPG) 1000 Å (purchased from Link Technologies, Bellshill, UK) was functionalized with 3'-O-acetyl-5'-O-(4,4'-dimethoxytrityl)thymidine through a hexaethylene glycol spacer as previously described (Fig. S2) [20,21]. 4.4 mg of this nucleoside-functionalized support (100 nmol) were treated with 3% trichloroacetic acid (TCA) in CH₂Cl₂ (5 min, r.t.) to remove the 4,4'-dimethoxytrityl (DMT) protecting group. After exhaustive washings with CH₂Cl₂, CH₃OH, and again CH₂Cl₂, the support was taken to dryness and then used as a control (nude CPG, Fig. S2). Standard phosphoramidite chemistry on an automated Applied Biosystem 394 DNA/RNA synthesizer was then used exploiting 5'-O-DMT, 2-cyanoethylphosphoramidite base-protected nucleoside monomers to obtain the oligonucleotide-functionalized CPG supports. In particular, using a 1 μmol-scale, "DMT-ON" protocol, the following oligonucleotides were assembled on the CPG supports: 5'(TTAGGG)₄TT^{3'} (tel26), 5'TGGGGAGGGTGGGGAGGGTGGGGAAGGTGGGGA^{3'} (cmyc), 5'AGGGAGGGCGCTGGGAGGAGGG^{3'} (ckit1), 5'CGGGCGGGCGCGAGGGAGGGG^{3'} (ckit2), 5'AAGGGGAGGGGCTGGGAGGGCGGGG^{3'} (hTERT1), and 5'CGCGAATTCGCGTTTCGCGAATTCGCG^{3'} (ds27). The coupling efficiency of each monomer was monitored by spectrophotometric measurements of the DMT cation, released from the support by acidic treatment with

3% TCA in CH₂Cl₂ before the subsequent coupling step. Considering the number of couplings and the average yield per cycle of 99.8%, 99.7%, 99.9%, 99.7%, 99.9% and 99.2%, respectively for tel26, cmyc, ckit1, ckit2, hTERT1 and ds27, the overall yield was determined to be 95%, 91%, 98%, 94%, 98%, and 80%. At the end of each synthesis, the oligonucleotide-functionalized solid supports were treated with 30% ammonium hydroxide/40% methylamine 1:1, v/v (AMA) at r.t. for 2 h to achieve nucleobases and phosphates deprotection.

Stock solutions of each tested compound were prepared by dissolving a weighed amount of the solid compound in pure DMSO. A known volume of the stock solution was then withdrawn and diluted with the proper buffer, so to obtain a 60 μM ligand solution. All the binding assays were carried out in triplicate, always using freshly prepared 60 μM compound solutions. The binding assay protocol consists in leaving a weighed amount of the functionalized CPG supports (corresponding to 100 nmol of oligonucleotide) in contact with 300 μL of the 60 μM ligand solution in a polypropylene column equipped with a polytetrafluoroethylene frit, a stopcock and a cap [20,22]. After incubation on a vibrating shaker for 4 min, defined volumes of the washing solution (50 mM KCl/10% DMSO/10% CH₃CH₂OH in H₂O) or the releasing solution (2.5 M CaCl₂/15% DMSO in H₂O or pure DMSO) were flown through the functionalized CPG supports and all the eluted fractions were separately analyzed by UV measurements [20]. The amount of bound ligand was calculated: i) by subtracting the amount of ligand eluted upon treatment with the washing solution, derived by direct UV measurements, from the ligand amount initially loaded on the supports, or ii) by direct UV measurements of the bound ligand released after treating the supports with the releasing solution (data not shown). The results obtained by the two methods were always in good agreement. After the treatment with the releasing solution inducing G4s and hairpin duplex denaturation, the supports were resuspended in the washing solution and then subjected again to the annealing procedure by taking them at 75 °C for 5 min and then slowly cooling at room temperature. The UV measurements were performed on a JASCO V-550 UV-vis spectrophotometer equipped with a Peltier Thermostat JASCO ETC-505 T. The UV quantification of the ligands was determined by measuring the absorbance relative to the λ_{max} characteristic of each ligand and referring it to the corresponding calibration curves. A quartz cuvette with a path length of 1 cm was used.

4.3. Biological experiments

4.3.1. Cells and culture condition

Human fibroblasts (BJ) and human cervical cancer cells (HeLa) were obtained as previously reported [52]. BJ-hTERT cells were obtained infecting primary BJ cells with a retrovirus carrying hTERT (Addgene plasmid #1773); BJ-EHLT derived from the transformation of BJ fibroblasts with hTERT and SV40 early region (BJ-EHLT). BJ-hTERT, BJ-EHLT and HeLa were grown in Dulbecco Modified Eagle Medium (D-MEM, Invitrogen Carlsbad, CA, USA) supplemented with 10% Fetal Bovine Serum (FBS), 2 mM L-glutamine and antibiotics at 37 °C in a 5% CO₂-95% air atmosphere.

4.3.2. Immunofluorescence

Cells were fixed in 2% formaldehyde and permeabilized in 0.25% Triton X-100 in phosphate buffered saline (PBS) for 5 min at r.t. For immune-labeling, cells were incubated with primary antibody for 2 h at r.t., washed twice in PBS and finally incubated with the secondary antibodies for 1 h. The following primary antibodies were used: Mouse mAb anti-γH2AX (Millipore, Billerica, MA, USA) and Rabbit pAb anti-TRF1 N19 (Santa Cruz Biotechnologies, Santa Cruz, CA, USA). The following secondary antibodies were used: Anti-Mouse IgG (H + L), F(ab')₂ Fragment (Alexa Fluor 488

Conjugate) (Cell Signaling) and Anti-rabbit IgG (H + L), F(ab')₂ Fragment (Alexa Fluor 555 Conjugate) (Cell Signaling). Nuclei were stained with 4',6-diamidino-2-phenylindole (DAPI, Sigma). Fluorescence signals were recorded by using a Leica DMIRE2 microscope equipped with a Leica DFC 350FX camera and elaborated by Leica FW4000 deconvolution software (Leica, Solms, Germany). For quantitative analysis of γ H2AX positivity, 300 cells on triplicate slices were scored. For TIF analysis, a single plane was analyzed, and 30 γ H2AX-positive cells were scored. Cells with at least four colocalizations (γ H2AX/TRF1) were considered as TIF-positive.

4.3.3. Clonogenic assay

Human cervical cancer cells, HeLa, were seeded in 60 mm-Petri dishes at the clonogenic density of 500 cells/plate in DMEM medium with 10% FBS. After 24 h, cells were treated with compound **4** or its analogs at 2 μ M concentration. After 10 days, the cells were stained with 2% methylene blue in 50% ethanol and the number of colonies was counted. Surviving fractions were calculated as the ratio of absolute survival of the treated sample/absolute survival of the untreated sample.

4.3.4. Cell viability

BJ-hTERT fibroblasts were seeded in 60-mm Petri dishes at a density of 5×10^4 cells/plate. After 24 h from plating, cells were treated with compound **4** or **S4-5** at the doses of 1 and 2 μ M. Cell viability (Trypan Blue dye exclusion) was determined after 6 days from treatment.

4.4. CD experiments

CD spectra were recorded in a quartz cuvette with a path length of 1 cm on a Jasco J-715 spectropolarimeter equipped with a Peltier-type temperature control system (model PTC-348WI). The spectra were registered at 20 °C in the range 240–800 nm with 2 s response, 200 nm/min scanning speed, 2.0 nm bandwidth, and corrected by subtraction of the background scan with buffer. All the spectra were averaged over 3 scans. The oligonucleotides 5'(TTAGGG)₄TT^{3'} (tel26), 5'TGGGGAGGGTGGGGAGGGTGGGGAA-GGTGGGGA^{3'} (cmyc) and 5'CGCGAATTCGCGTTTCGCGAATTCGCG^{3'} (ds27) were synthesized by standard automated solid phase oligonucleotide synthesis on an Applied Biosystem 394 DNA/RNA synthesizer. After ammonia treatment (55 °C, 12 h) allowing both deprotection and detachment from the solid support, the crude oligonucleotides were purified by HPLC on a SAX analytical column and then dialyzed against water using a Float-A-Lyzer G2 dialysis device (MWCO 0.5–1.0 kDa, three H₂O changes over 24 h). After lyophilization, the oligonucleotides were dissolved in a 20 mM KCl - 5 mM KH₂PO₄ - 10% DMSO buffer (pH 7) to obtain 2 μ M solutions, then annealed by heating to 95 °C for 5 min, followed by slow cooling to room temperature. The ligand stock solutions were 4 mM in DMSO. CD titrations were obtained by adding increasing amounts of the ligands (up to 5 molar equivalents, corresponding to a 10 μ M solution in ligand) to tel26, cmyc and ds27. After each ligand addition, the system was allowed equilibrating before registering the spectra. For the CD-melting experiments, the ellipticity was recorded at 290, 263 and 251 nm for tel26, cmyc and ds27, respectively, with a temperature scan rate of 0.5 °C/min in the range 20–90 °C.

4.5. NMR experiments

NMR experiments were performed on a 700 MHz Varian Unity INOVA spectrometer. One dimensional ¹H NMR spectra of the samples in H₂O were recorded at 25 °C using pulsed-field gradient DPFGE for H₂O suppression [53]. Data were processed on iMAC

running iNMR software (www.inmr.net). DNA samples were prepared at 0.2–0.4 mM strand concentration in 0.6 mL of H₂O/D₂O (9:1) buffer solution. The following oligonucleotides were used for the NMR experiments: the truncation of human telomeric sequence 5'-TTGGG(TTAGGG)₃A^{3'} (m-tel24) [46], the 5'TGAGGGTGG-GTAGGGTGGGTA^{3'} sequence from the NHE III element of the c-myc oncogene (myc22), containing two G to T substitutions (G14/T and G23/T, numbering according to Ambrus, A. et al.) [47], and the self-complementary duplex-forming dodecamer 5'CGCGAA-TTCGCG^{3'} (ds12) [48]. In order to avoid different G4 folds or conformational heterogeneity in solution, DNA samples were prepared using the appropriate experimental conditions as used for the determination of their 3D structures. Thus, the following buffers containing 10% D₂O were used: 25 mM KH₂PO₄, 70 mM KCl, 0.2 mM EDTA (pH 7.0) for myc22 and m-tel24; 20 mM NaH₂PO₄, 200 mM NaCl (pH 7.0) for ds12. The samples were heated at 90 °C for 5 min and then slowly cooled to room temperature overnight to achieve the correct folding of the oligonucleotides. Aliquots of **4** and **S4-5** stock solutions in DMSO-*d*₆ were directly added to the DNA solution inside the NMR tube; the final DMSO concentration was 14%. A control titration was also performed by adding DMSO-*d*₆ alone to the DNA solution and subtracted from each titration experiment to remove effects due to DMSO on the chemical shift changes.

4.6. Microscale thermophoresis (MST) experiments

MST measurements were performed using the Monolith NT.115 (Nanotemper Technologies, Munich, Germany). The Cy5-fluorescently labeled tel26 and myc22 oligonucleotides (Biomers) were prepared in 5 mM potassium phosphate buffer (pH 7.0) containing 20 mM KCl supplemented with 0.1% Tween. The concentration of the labeled oligonucleotide was kept constant at 80 nM, while a serial dilution of the ligand (1:2 from 4 mM ligand stock solution in 100% DMSO) in the same buffer used for DNAs was prepared and mixed with the oligonucleotide solution with a volume ratio of 1:1. All the samples, containing 8% DMSO as the final concentration, were loaded into standard capillaries (NanoTemper Technologies). Measurements were performed at 20 °C, using auto-tune LED power and medium MST power. MST data analysis was performed by employing the MO.Affinity Analysis software (v2.3) provided with the instrument. Plots were rendered with GUSSI version 1.2.1 software (<http://biophysics.swmed.edu/MBR/software.html>).

4.7. Docking

The latest version of the docking software AD4 (version 4.2) [54], along with its graphical user interface AutoDockTools (ADT), was employed. The NMR structures used for the docking studies had the following PDB codes: 2JPZ, and 1XAV [26,47]. The DNA G4s were prepared for the docking using the Maestro suite (Schrödinger Release 2017–2: Maestro, Schrödinger, LLC, New York, NY, 2017), which assigns bond orders, add hydrogen atoms, deletes water molecules and generates the appropriate protonation states. The 2D Sketcher tool of Maestro was used to build **S4-5**. Of this ligand, the protonation, tautomeric, and isomeric states were calculated through LigPrep, part of the same suite. The ligand and the G4 DNA structures were converted to the AD4 specific file format (PDBQT) using the python scripts prepare_ligand4.py and prepare_receptor4.py, part of ADT, applying the standard settings. The docking area was centered on the DNA center of mass. For each G4 structure, a set of grids of 60 Å × 60 Å × 60 Å with 0.375 Å spacing was calculated around the docking area for the ligand atom types using AutoGrid4. For each G4, 100 separate docking

calculations were performed. Each docking calculation consisted of 10 million energy evaluations using the Lamarckian genetic algorithm local search (GALS) method. The GALS method evaluates a population of possible docking solutions and propagates the most successful individuals from each generation into the subsequent generation of possible solutions. A low-frequency local search according to the method of Solis and Wets is applied to docking trials to ensure that the final solution represents a local minimum. All dockings were performed with a population size of 250, and 300 rounds of Solis and Wets local search were applied with a probability of 0.06. A mutation rate of 0.02 and a crossover rate of 0.8 were used to generate new docking trials for subsequent generations, and the best individual from each generation was propagated over the next generation. The docking results from each of the 100 calculations were clustered on the basis of root-mean square deviation (rmsd) (solutions differing by less than 2.0 Å) between the Cartesian coordinates of the atoms and were ranked on the basis of the free energy of binding (ΔG_{AD4}). Molecular modeling figures were rendered using the UCSF Chimera software [55].

Author contributions

Da.Mo., B.P., J.A., C.P., S.C. and A.B. conceived and designed the experiments. J.A., C.P., P.Z., S.I., Do.Mu., S.C. and A.P., performed the experiments. Da.Mo., B.P., J.A., C.P., Do.Mu., S.C., E.N., A.B. and A.R. analyzed the data. Da.Mo., B.P., J.A., C.P., S.C. and A.B. wrote the paper: All authors have given approval to the final version of the manuscript.

Conflicts of interest

The authors declare no competing financial and non-financial interest.

Abbreviations

AD4.2, AutoDock 4.2; ADT, AutoDockTools; CPG, Controlled Pore Glass; DAPI, 4',6'-diamidino-2-phenylindole; DDR, DNA Damage Response; DMEM, Dulbecco's Modified Eagle's Medium; DMT, 4,4'-dimethoxytriphenylmethyl; DPGFSE, Double Pulsed Field Gradient Spin Echo; FBS, Fetal Bovine Serum; G4, G-quadruplex; GALS, Genetic Algorithm Local Search; γ H2AX, phosphorylated form of histone H2AX; hTERT, human Telomerase Reverse Transcriptase; IF, ImmunoFluorescence; LCAA-CPG, Long Chain AlkylAmine-CPG; MST, MicroScale Thermophoresis; MWCO, Molecular Weight Cut-Off; NHE, Nuclease Hypersensitivity Element; PDGF-A, Platelet-Derived Growth Factor-A; Rb, Retinoblastoma; RET, REarranged during Transfection; SAX, Strong Anion Exchange; TIF, Telomere Induced Foci; TO, Thiazole Orange; TRF1, Telomere Repeat Factor 1.

Acknowledgements

This work was supported by the Italian Association for Cancer Research (IG2015 n. 16730 to B.P.; IG2015 n. 17037 to D.M.; IG2016 n. 18695 to A.R.; IG2016 n. 16910 to A.B.).

Appendix A. Supplementary data

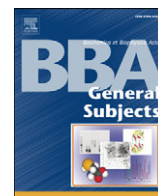
Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ejmech.2018.11.058>.

References

[1] Z. Liu, B. Delavan, R. Roberts, W. Tong, Lessons learned from two decades of anticancer drugs, *Trends Pharmacol. Sci.* 38 (2017) 852–872, <https://doi.org/>

- 10.1016/j.tips.2017.06.005.
- [2] D. Musumeci, C. Platella, C. Riccardi, A. Merlino, T. Marzo, L. Massai, L. Messori, D. Montesarchio, A first-in-class and a fished out anticancer platinum compound: cis-[PtCl₂(NH₃)₂] and cis-[PtI₂(NH₃)₂] compared for their reactivity towards DNA model systems, *Dalton Trans.* 45 (2016) 8587–8600, <https://doi.org/10.1039/C6DT00294C>.
- [3] N.J. Wheate, S. Walker, G.E. Craig, R. Oun, The status of platinum anticancer drugs in the clinic and in clinical trials, *Dalton Trans.* 39 (2010) 8113–8127, <https://doi.org/10.1039/c0dt00292e>.
- [4] S. Nussbaumer, P. Bonnabry, J.L. Veuthey, S. Fleury-Souverain, Analysis of anticancer drugs: a review, *Talanta* 85 (2011) 2265–2289, <https://doi.org/10.1016/j.talanta.2011.08.034>.
- [5] Viswanadha Vijaya Padma, An overview of targeted cancer therapy, *Biomedicine* 5 (2015) 19, <https://doi.org/10.7603/s40>.
- [6] T.A. Baudino, Targeted cancer therapy: the next generation of cancer treatment, *Curr. Drug Discov. Technol.* 12 (2015) 3–20, <https://doi.org/10.2174/1570163812666150602144310>.
- [7] M. Huang, A. Shen, J. Ding, M. Geng, Molecularly targeted cancer therapy: some lessons from the past decade, *Trends Pharmacol. Sci.* 35 (2014) 41–50, <https://doi.org/10.1016/j.tips.2013.11.004>.
- [8] Y. Gao, J.K. Shen, L. Milane, F.J. Hornicek, M.M. Amiji, Z. Duan, Targeted cancer therapy; nanotechnology approaches for overcoming drug resistance, *Curr. Med. Chem.* 22 (2015) 1335–1347, <https://doi.org/10.2174/0929867322666150209151851>.
- [9] C. Platella, C. Riccardi, D. Montesarchio, G.N. Roviello, D. Musumeci, G-quadruplex-based aptamers against protein targets in therapy and diagnostics, *Biochim. Biophys. Acta Gen. Subj.* 1861 (2017) 1429–1447, <https://doi.org/10.1016/j.bbagen.2016.11.027>.
- [10] D. Musumeci, C. Platella, C. Riccardi, F. Moccia, D. Montesarchio, Fluorescence sensing using DNA aptamers in cancer research and clinical diagnostics, *Cancers* 9 (2017) 174–217, <https://doi.org/10.3390/cancers9120174>.
- [11] R. Hänsel-Hertsch, M. Di Antonio, S. Balasubramanian, DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential, *Nat. Rev. Mol. Cell Biol.* 18 (2017) 279–284, <https://doi.org/10.1038/nrm.2017.3>.
- [12] S. Balasubramanian, L.H. Hurley, S. Neidle, Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discov.* 10 (2011) 261–275, <https://doi.org/10.1038/nrd3428>.
- [13] G. Biffi, D. Tannahill, J. McCafferty, S. Balasubramanian, Quantitative visualization of DNA G-quadruplex structures in human cells, *Nat. Chem.* 5 (2013) 182–186, <https://doi.org/10.1038/nchem.1548>.
- [14] J.L. Huppert, S. Balasubramanian, Prevalence of quadruplexes in the human genome, *Nucleic Acids Res.* 33 (2005) 2908–2916, <https://doi.org/10.1093/nar/gki609>.
- [15] T.A. Brooks, S. Kendrick, L. Hurley, Making sense of G-quadruplex and i-motif functions in oncogene promoters, *FEBS J.* 277 (2010) 3459–3469, <https://doi.org/10.1111/j.1742-4658.2010.07759.x>.
- [16] C. Marchetti, K.G. Zyner, S.A. Ohnmacht, M. Robson, S.M. Haider, J.P. Morton, G. Marsico, T. Vo, S. Laughlin-Toth, A.A. Ahmed, G. Di Vita, I. Pazitna, M. Gunaratnam, R.J. Besser, A.C.G. Andrade, S. Diocou, J.A. Pike, D. Tannahill, R.B. Pedley, T.R.J. Evans, W.D. Wilson, S. Balasubramanian, S. Neidle, Targeting multiple effector pathways in pancreatic ductal adenocarcinoma with a G-quadruplex-binding small molecule, *J. Med. Chem.* 61 (2018) 2500–2517, <https://doi.org/10.1021/acs.jmedchem.7b01781>.
- [17] D.L. Ma, V.P.Y. Ma, D.S.H. Chan, K.H. Leung, H.J. Zhong, C.H. Leung, In silico screening of quadruplex-binding ligands, *Methods* 57 (2012) 106–114, <https://doi.org/10.1016/j.ymeth.2012.02.001>.
- [18] S. Cosconati, L. Marinelli, R. Trotta, A. Virno, L. Mayol, E. Novellino, A.J. Olson, A. Randazzo, Tandem application of virtual screening and NMR experiments in the discovery of brand new DNA quadruplex groove binders, *J. Am. Chem. Soc.* 131 (2009) 16336–16337, <https://doi.org/10.1021/ja9063662>.
- [19] B. Pagano, J. Amato, N. Iaccarino, C. Cingolani, P. Zizza, A. Biroccio, E. Novellino, A. Randazzo, Looking for efficient G-quadruplex ligands: evidence for selective stabilizing properties and telomere damage by drug-like molecules, *ChemMedChem* 10 (2015) 640–649, <https://doi.org/10.1002/cmdc.201402552>.
- [20] C. Platella, D. Musumeci, A. Arciello, F. Doria, M. Freccero, A. Randazzo, J. Amato, B. Pagano, D. Montesarchio, Controlled pore glass-based oligonucleotide affinity support: towards high throughput screening methods for the identification of conformation-selective G-quadruplex ligands, *Anal. Chim. Acta* 1030 (2018) 133–141, <https://doi.org/10.1016/j.aca.2018.04.071>.
- [21] C. Platella, D. Musumeci, J. Amato, A. Randazzo, B. Pagano, D. Montesarchio, The Content Is Object of the Italian Patent Application N. 102017000030473, 2017 filed on March 20.
- [22] D. Musumeci, J. Amato, P. Zizza, C. Platella, S. Cosconati, C. Cingolani, A. Biroccio, E. Novellino, A. Randazzo, C. Giancola, B. Pagano, D. Montesarchio, Tandem application of ligand-based virtual screening and G4-OAS assay to identify novel G-quadruplex-targeting chemotypes, *Biochim. Biophys. Acta Gen. Subj.* 1861 (2017) 1341–1352, <https://doi.org/10.1016/j.bbagen.2017.01.024>.
- [23] D. Musumeci, J. Amato, A. Randazzo, E. Novellino, C. Giancola, D. Montesarchio, B. Pagano, G-quadruplex on oligo affinity support (G4-OAS): an easy affinity chromatography-based assay for the screening of G-quadruplex ligands, *Anal. Chem.* 86 (2014) 4126–4130, <https://doi.org/10.1021/ac500444m>.
- [24] http://www.glenresearch.com/Technical/TB_OAS.pdf.

- [25] M. Zuffo, F. Doria, S. Botti, G. Bergamaschi, M. Freccero, G-quadruplex fluorescence sensing by core-extended naphthalene diimides, *Biochim. Biophys. Acta Gen. Subj.* 1861 (2017) 1303–1311, <https://doi.org/10.1016/j.bbagen.2016.11.034>.
- [26] J. Dai, M. Carver, C. PUNCHIHEWA, R.A. Jones, D. Yang, 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 (2007) 4927–4940, <https://doi.org/10.1093/nar/gkm522>.
- [27] T. Simonsson, M. Pribylova, M. Vorlickova, A nuclease hypersensitive element in the human c-myc promoter adopts several distinct i-tetraplex structures, *Biochem. Biophys. Res. Commun.* 278 (2000) 158–166, <https://doi.org/10.1006/bbrc.2000.3783>.
- [28] D. Sun, L.H. Hurley, The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression, *J. Med. Chem.* 52 (2009) 2863–2874, <https://doi.org/10.1021/jm900055s>.
- [29] R.I. Mathad, E. Hatzakis, J. Dai, D. Yang, c-MYC promoter G-quadruplex formed at the 5'-end of NHE III element: insights into biological relevance and parallel-stranded G-quadruplex stability, *Nucleic Acids Res.* 39 (2011) 9023–9033, <https://doi.org/10.1093/nar/gkr612>.
- [30] A.T. Phan, V. Kuryavyi, S. Burge, S. Neidle, D.J. Patel, Structure of an unprecedented G-quadruplex scaffold in the human c-kit promoter, *J. Am. Chem. Soc.* 129 (2007) 4386–4392, <https://doi.org/10.1021/ja068739h>.
- [31] D. Wei, G.N. Parkinson, A.P. Reszka, S. Neidle, Crystal structure of a c-kit promoter quadruplex reveals the structural role of metal ions and water molecules in maintaining loop conformation, *Nucleic Acids Res.* 40 (2012) 4691–4700, <https://doi.org/10.1093/nar/gks023>.
- [32] D. Wei, J. Husby, S. Neidle, Flexibility and structural conservation in a c-KIT G-quadruplex, *Nucleic Acids Res.* 43 (2015) 629–644, <https://doi.org/10.1093/nar/gku1282>.
- [33] V. Kuryavyi, A.T. Phan, D.J. Patel, Solution structures of all parallel-stranded monomeric and dimeric G-quadruplex scaffolds of the human c-kit2 promoter, *Nucleic Acids Res.* 38 (2010) 6757–6773, <https://doi.org/10.1093/nar/gkq558>.
- [34] S.M.L. Palumbo, S.W. Ebbinghaus, L.H. Hurley, Formation of a unique end-to-end stacked pair of G-quadruplexes in the hTERT core promoter with implications for inhibition of telomerase by G-quadruplex-interactive ligands, *J. Am. Chem. Soc.* 131 (2009) 10878–10891, <https://doi.org/10.1021/ja902281d>.
- [35] J.B. Chaires, J.O. Trent, R.D. Gray, W.L. Dean, R. Buscaglia, S.D. Thomas, D.M. Miller, An improved model for the hTERT promoter quadruplex, *PLoS One* 9 (2014), <https://doi.org/10.1371/journal.pone.0115580>.
- [36] H.R. Drew, R.M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, R.E. Dickerson, Structure of a B-DNA dodecamer: conformation and dynamics, *Proc. Natl. Acad. Sci. Unit. States Am.* 78 (1981) 2179–2183, <https://doi.org/10.1073/pnas.78.4.2179>.
- [37] D. Monchard, C. Allain, H. Bertrand, N. Smargiasso, F. Rosu, V. Gabelica, A. De Cian, J.L. Mergny, M.P. Teulade-Fichou, Ligands playing musical chairs with G-quadruplex DNA: a rapid and simple displacement assay for identifying selective G-quadruplex binders, *Biochimie* 90 (2008) 1207–1223, <https://doi.org/10.1016/j.biochi.2008.02.019>.
- [38] C. Platella, S. Guida, L. Bonmassar, A. Aquino, E. Bonmassar, G. Ravagnan, D. Montesarchio, G.N. Roviello, D. Musumeci, M.P. Fuggetta, Antitumour activity of resveratrol on human melanoma cells: a possible mechanism related to its interaction with malignant cell telomerase, *Biochim. Biophys. Acta Gen. Subj.* 1861 (2017) 2843–2851, <https://doi.org/10.1016/j.bbagen.2017.08.001>.
- [39] A. Cerrato, F. Morra, A. Celetti, Use of poly ADP-ribose polymerase [PARP] inhibitors in cancer cells bearing DDR defects: the rationale for their inclusion in the clinic, *J. Exp. Clin. Oncol.* 35 (2016) 179, <https://doi.org/10.1186/s13046-016-0456-2>.
- [40] L. Pompili, C. Leonetti, A. Biroccio, E. Salvati, Diagnosis and treatment of ALT tumors: is Trabectedin a new therapeutic option? *J. Exp. Clin. Oncol.* 36 (2017) 189, <https://doi.org/10.1186/s13046-017-0657-3>.
- [41] C. Thiriet, J.J. Hayes, Chromatin in need of a fix: phosphorylation of H2AX connects chromatin to DNA repair, *Mol. Cell.* 18 (2005) 617–622, <https://doi.org/10.1016/j.molcel.2005.05.008>.
- [42] H. Takai, A. Smogorzewska, T. De Lange, DNA damage foci at dysfunctional telomeres, *Curr. Biol.* 13 (2003) 1549–1556, [https://doi.org/10.1016/S0960-9822\(03\)00542-6](https://doi.org/10.1016/S0960-9822(03)00542-6).
- [43] A.I. Karsisiotis, N.M.A. Hessari, E. Novellino, G.P. Spada, A. Randazzo, M. Webba da Silva, Topological characterization of nucleic acid G-Quadruplexes by UV absorption and circular dichroism, *Angew. Chem. Int. Ed.* 50 (2011) 10645–10648, <https://doi.org/10.1002/anie.201105193>.
- [44] J. Kypr, I. Kejnovská, D. Renčuk, M. Vorlíčková, Circular dichroism and conformational polymorphism of DNA, *Nucleic Acids Res.* 37 (2009) 1713–1725, <https://doi.org/10.1093/nar/gkp026>.
- [45] B. Pagano, S. Cosconati, V. Gabelica, L. Petraccone, S. De Tito, L. Marinelli, V. La Pietra, F.S. di Leva, I. Lauri, R. Trotta, E. Novellino, C. Giancola, A. Randazzo, State-of-the-art methodologies for the discovery and characterization of DNA G-quadruplex binders, *Curr. Pharmaceut. Des.* 18 (2012) 1880–1899, <https://doi.org/10.1080/1381-6128.2012.712812>.
- [46] K.N. Luu, A.T. Phan, V. Kuryavyi, L. Lacroix, D.J. Patel, Structure of the human telomere in K⁺ solution: an intramolecular (3 + 1) G-quadruplex scaffold, *J. Am. Chem. Soc.* 128 (2006) 9963–9970, <https://doi.org/10.1021/ja062791w>.
- [47] A. Ambrus, D. Chen, J. Dai, R.A. Jones, D. Yang, Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization, *Biochemistry* 44 (2005) 2048–2058, <https://doi.org/10.1021/bi048242p>.
- [48] J.W. Jaroszewski, V. Clausen, J.S. Cohen, O. Dahl, NMR investigations of duplex stability of phosphorothioate and phosphorodithioate DNA analogues modified in both strands, *Nucleic Acids Res.* 24 (1996) 829–834, <https://doi.org/10.1093/nar/24.5.829>.
- [49] C. Entzian, T. Schubert, Studying small molecule-aptamer interactions using MicroScale Thermophoresis (MST), *Methods* 97 (2016) 27–34, <https://doi.org/10.1016/j.ymeth.2015.08.023>.
- [50] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A.J. Olson, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, *J. Comput. Chem.* 19 (1998) 1639–1662, doi:10.1002/(SICI)1096-987X(19981115)19:14<1639::AID-JCC10>3.0.CO;2-B.
- [51] S. Cosconati, S. Forli, A.L. Perryman, R. Harris, D.S. Goodsell, A.J. Olson, Virtual screening with AutoDock: theory and practice, *Expert Opin. Drug Discov.* 5 (2010) 597–607, <https://doi.org/10.1517/17460441.2010.484460>.
- [52] E. Salvati, M. Scarsella, M. Porru, A. Rizzo, S. Iachettini, L. Tentori, G. Graziani, M. D'Incalci, M.F.G. Stevens, A. Orlandi, D. Passeri, E. Gilson, G. Zupi, C. Leonetti, A. Biroccio, PARP1 is activated at telomeres upon G4 stabilization: possible target for telomere-based therapy, *Oncogene* 29 (2010) 6280–6293, <https://doi.org/10.1038/onc.2010.344>.
- [53] T.L. Hwang, A.J. Shaka, Water suppression that works. excitation sculpting using arbitrary wave-forms and pulsed-field gradients, *J. Magn. Reson., Ser. A* 112 (1995) 275–279, <https://doi.org/10.1006/jmra.1995.1047>.
- [54] G. Morris, R. Huey, AutoDock4 and autoDockTools4: automated docking with selective receptor flexibility, *J. Comput. Chem.* 30 (2009) 2785–2791, <https://doi.org/10.1002/jcc.21256>.
- [55] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera: a visualization system for exploratory research and analysis, *J. Comput. Chem.* 25 (2004) 1605–1612, <https://doi.org/10.1002/jcc.20084>.



Discovery of the first dual G-triplex/G-quadruplex stabilizing compound: a new opportunity in the targeting of G-rich DNA structures?☆



Jussara Amato^a, Alessia Pagano^a, Sandro Cosconati^b, Giorgio Amendola^b, Iolanda Fotticchia^a, Nunzia Iaccarino^a, Jessica Marinello^c, Alessio De Magis^c, Giovanni Capranico^c, Ettore Novellino^a, Bruno Pagano^a, Antonio Randazzo^{a,*}

^a Department of Pharmacy, University of Naples Federico II, via D. Montesano 49, 80131 Naples, Italy

^b DISTABIF, Second University of Naples, via Vivaldi 43, 81100 Caserta, Italy

^c Department of Pharmacy and Biotechnology, University of Bologna, 40126 Bologna, Italy

ARTICLE INFO

Article history:

Received 12 October 2016

Received in revised form 3 November 2016

Accepted 5 November 2016

Available online 9 November 2016

Keywords:

G-triplex

G-quadruplex

Ligand

Anticancer agents

Drug discovery

ABSTRACT

Background: Guanine-rich DNA motifs can form non-canonical structures known as G-quadruplexes, whose role in tumorigenic processes makes them attractive drug-target candidates for cancer therapy. Recent studies revealed that the folding and unfolding pathways of G-quadruplexes proceed through a quite stable intermediate named G-triplex.

Methods: Virtual screening was employed to identify a small set of putative G-triplex ligands. The G-triplex stabilizing properties of these compounds were analyzed by CD melting assay. DSC, non-denaturing gel electrophoresis, NMR and molecular modeling studies were performed to investigate the interaction between the selected compound **1** and G-rich DNA structures. Cytotoxic activity of **1** was evaluated by MTT cell proliferation assay.

Results: The experiments led to the identification of a promising hit that was shown to bind preferentially to G-triplex and parallel-stranded G-quadruplexes over duplex and antiparallel G-quadruplexes. Molecular modeling results suggested a partial end-stacking of **1** to the external G-triad/G-tetrads as a binding mode. Biological assays showed that **1** is endowed with cytotoxic effect on human osteosarcoma cells.

Conclusions: A tandem application of virtual screening along with the experimental investigation was employed to discover a G-triplex-targeting ligand. Experiments revealed that the selected compound actually acts as a dual G-triplex/G-quadruplex stabilizer, thus stimulating further studies aimed at its optimization.

General significance: The discovery of molecules able to bind and stabilize G-triplex structures is highly appealing, but their transient state makes challenging their recognition. These findings suggest that the identification of ligands with dual G-triplex/G-quadruplex stabilizing properties may represent a new route for the design of anticancer agents targeting the G-rich DNA structures. This article is part of a Special Issue entitled "G-quadruplex" Guest Editor: Dr. Concetta Giancola and Dr. Daniela Montesarchio.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

DNA represents one of the most important macromolecules in living systems since it encodes the genetic information in most organisms [1]. One important property of DNA is its structural polymorphism that plays key roles in many cellular processes and human diseases [2]. Indeed, in addition to the Watson–Crick duplex structure, various secondary structures could exist *in vitro* and *in vivo*, such as cruciforms [2], Z-DNA [3], triplexes (H-DNA) [4,5], G-quadruplexes [6,7], and i-motifs [8]. Among these structures, G-quadruplexes (G4s) have attracted a

great interest in the recent years because of their involvement in a number of critical cancer-related genomic aberrations and pathways [9]. Consequently, they emerged as promising drug targets and much efforts are currently underway to discover effective ligands for them [10]. G4s are four-stranded structures composed of stacked guanine quartets (G-tetrads) in which four guanine bases are Hoogsteen hydrogen bonded to each other. G4s can form within G-rich sequences located at different critical positions of the human genome, notably at telomeres and oncogene promoters [11]. Because of the biological significance of this peculiar DNA arrangement and the related biological implications, studies on its folding pathway are of topical interest. Several spectroscopic studies show that some G-rich sequences fold into intramolecular G4 structures by following a multistep pathway that proceeds through a quite stable intermediate named G-triplex (G3) [12–16]. In addition, recent studies have revealed that G3 plays a role also in the unfolding process of some G4 structures [14].

☆ This article is part of a Special Issue entitled "G-quadruplex" Guest Editor: Dr. Concetta Giancola and Dr. Daniela Montesarchio.

* Corresponding author.

E-mail address: antonio.randazzo@unina.it (A. Randazzo).

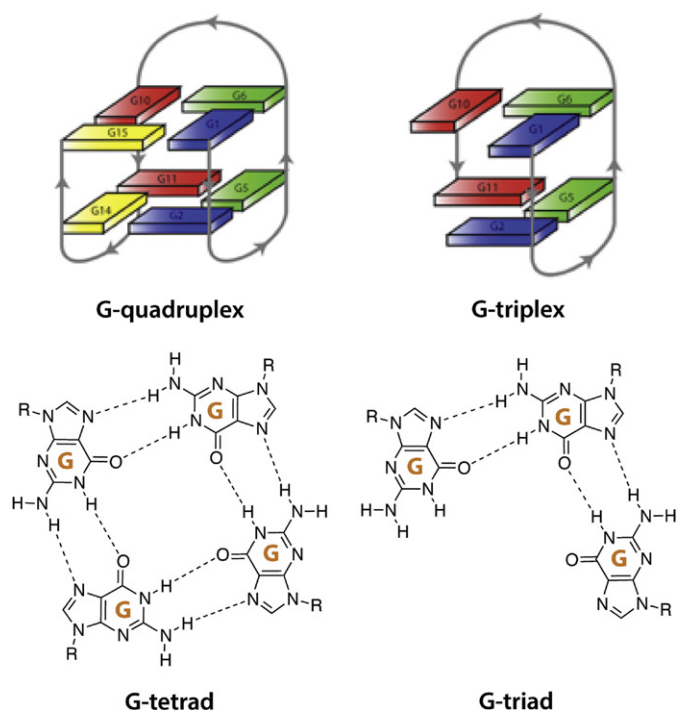


Fig. 1. Comparison between G-quadruplex and G-triplex structures formed by the 15-mer- and 11-mer-TBA sequences.

Similarly to the G4, a G3 structure is stabilized by Hoogsteen-like hydrogen bonds between guanine residues involved in G:C:G triad planes (Fig. 1) [15,16]. However, the lower number of hydrogen bonds, together with a smaller stacking surface of the triads with respect to the G-tetrads, make G3 much less stable than G4 and, therefore, even more challenging to isolate both *in vitro* and *in vivo*. For all these reasons, it is clear that the discovery of molecules able to interact and stabilize G3s is highly appealing, also for the understanding of the putative biological and therapeutic importance of these intermediates.

Here, we attempted at searching for brand new molecular scaffolds able to target a very simple G3 structure, namely the 11-mer-TBA (PDB ID: 2MKM) [16] by means of a structure-based virtual screening (VS) approach and a combination of experimental techniques. Particularly, we report a high-throughput *in silico* screening of a large number of compounds from a commercially available database by combining a structure-based pharmacophore model approach with docking simulations. This allowed us to select 15 small organic molecules (Table S1), characterized by hydrogen-bond acceptor/donor groups, (hetero)aromatic and/or non-aromatic heterocyclic rings, as best hits for the G3 formed by the 11-mer-TBA DNA sequence. The selected hits were evaluated for their binding properties by biophysical methodologies, and one of them, a dihydropyrimidin-4-one derivative (6-(4-fluorophenyl)-2-[(8-methoxy-4-methylquinazolin-2-yl)amino]-3,4-dihydropyrimidin-4-one, herein referred to as compound **1**, Fig. 2) turned out to be a promising lead compound able to stabilize the G3

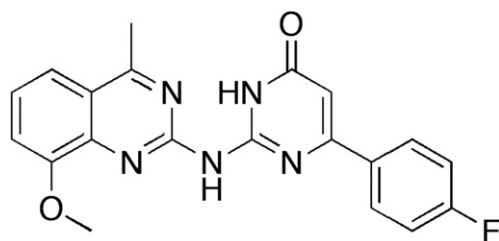


Fig. 2. Chemical structure of compound **1**.

structure. Interestingly, the physicochemical characterization of the binding profile of this ligand, led to the discovery that **1** is also able to interact and stabilize some G4 structures, thus acting as a dual G3 and G4 stabilizer, with no affinity for the duplex DNA. The subsequent investigation of the biological activity of the selected compound demonstrated that it is endowed with cytotoxic effect on human U2OS cells.

2. Materials and methods

2.1. Virtual screening

VS calculations on the G3 structure were performed employing the software Autodock 4.2 (AD4.2) [17] which is effective in detecting nucleic-acid-binding small molecules [18]. Thus, AD4.2 was used to dock a diversity set of 58,870 compounds from the commercially available Molecule chemical database (www.molecule.com). Prior to docking calculations, the library was processed according to the ZINC protocol [19] to generate all the tautomeric and protomeric states for each compound. Given the presence of multiple druggable sites on G3 (PDB ID: 2MKM), the docking search area was chosen to enclose the entire DNA structure. VS results were sorted on the basis of the predicted ligand binding free energies (ΔG_{AD4}). All the solutions with a ΔG_{AD4} greater than -6.00 kcal/mol and a cluster size population lower than 10 out of 100 individuals were discarded. The remaining solutions were visually inspected to discard all the compounds that were not predicted to establish tight contacts with G3 (i.e., coulombic interactions with the phosphate backbone atoms or stacking interactions with the guanine residues). Subsequently, the remaining solutions were grouped based on their structural similarity to avoid structural redundancy, and the individuals with the lowest ΔG_{AD4} value within each group were finally selected. Finally, these compounds were carefully inspected for good geometries in the predicted binding pose. At the end of the visual inspection, 15 compounds were considered for further experimental investigations.

2.2. Oligonucleotide synthesis and sample preparation

The DNA sequences were synthesized using standard β -cyanoethylphosphoramidite solid phase chemistry on an ABI 394 DNA/RNA synthesizer (Applied Biosystem) at the 5- μ mol scale. DNA detachment from support and deprotection were performed by treatment with concentrated ammonia aqueous solution at 55 °C for 12 h. The combined filtrates and washings were concentrated under reduced pressure, dissolved in water, and purified by high-performance liquid chromatography (HPLC) on a Nucleogel SAX column (Macherey-Nagel, 1000-8/46), using buffer A consisting of 20 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ aqueous solution (pH 7.0), containing 20% (v/v) CH_3CN , buffer B consisting of 1 M KCl, 20 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ aqueous solution (pH 7.0), containing 20% (v/v) CH_3CN , and a linear gradient from 0% to 100% B for 30 min with a flow rate 1 mL/min. The fractions of the oligomers were collected and successively desalted by Sep-pak cartridges (C-18). The isolated oligomer was proved to be >98% pure by NMR. In particular, the following oligonucleotides have been used for the experiments: the G3-forming oligonucleotide corresponding to the 3'-end truncation of the thrombin binding aptamer (TBA) sequence (d[GGTTGGTGTGG], 11-mer-TBA); two G4-forming sequences from the promoter regions of *c-KIT* (d[AGGGAGGGCGCTGGGAGGAGGG], *c-kit1*) and *c-MYC* (d[TGAGGGGTAGGGTGGGTA], *myc22*) oncogenes; and the self-complementary duplex-forming dodecamer (d[CGCGAATTCGCG], *ds12*). The concentration of the oligonucleotides was determined by UV absorption measurements at 90 °C using appropriate molar extinction coefficient values ϵ ($\lambda = 260$ nm) calculated by the nearest neighbor model [20]. G3 and G4 s were prepared in the appropriate buffer (10 mM potassium phosphate, 70 mM KCl, pH 7.0 for 11-mer-TBA [16]; 10 mM potassium phosphate, 100 mM KCl, pH 7.0 for *c-kit1* [21] and 5 mM potassium phosphate, 20 mM KCl, pH 7.0 for *myc22* [22]) at 10–30 μ M oligonucleotide concentration, unless otherwise stated.

Samples were heated at 90 °C for 5 min, then gradually cooled to room temperature overnight and stored at 4 °C for 24 h before experiments. Finally, duplex DNA was prepared at 30 μ M single strand concentration in 10 mM Li_3PO_4 , 100 mM KCl, pH 7.0 and annealed as reported above.

2.3. Circular dichroism spectroscopy

Circular dichroism (CD) experiments were recorded on a Jasco J-815 spectropolarimeter equipped with a PTC-423S/15 Peltier temperature controller. CD spectra were recorded at 5 °C (for G3) and 20 °C (for G4s and ds12) in the wavelength range of 230–360 nm and averaged over 3 scans. The scan rate was set to 100 nm/min, with a 1 s response time, and 1 nm bandwidth. CD spectra were baseline corrected and analyzed using Origin 7.0 software. 10 μ M of G4 and 30 μ M of G3 and duplex DNA were used. DNA/ligand mixtures were obtained by adding 4 M equiv. of ligands (stock solutions of ligands were 6 mM in DMSO). CD melting experiments were carried out in the 5–80 °C or 20–105 °C temperature range, at 1 °C/min heating rate by following changes of CD signal at the wavelengths of the maximum CD intensity. CD melting experiments were recorded in the absence and presence of ligands added to the folded DNA structures. The melting temperatures were determined from curve fit using Origin 7.0 software. ΔT_m values were determined as the difference in melting temperature between the DNA with and without ligands [23]. Finally, CD melting experiments were also performed by varying concentration of **1** added to DNA. In particular, four different [ligand]/[DNA] ratio were examined (1:1, 2:1, 4:1 and 8:1).

2.4. Gel electrophoresis

Native gel electrophoresis analysis was carried out on 20% polyacrylamide gel at 5 °C, which was run in 1 \times TB (pH 7.5) buffer supplemented with 80 mM KCl. An oligonucleotide concentration of 50 μ M was used for each sample. Various amounts (1–10 equiv) of **1** were incubated with DNA at 5 °C for 1 h before loading. Prior to loading the mixture onto the gel, 1 μ L of glycerol solution (60% v/v) was added. The total volume loaded in each well was 10 μ L. Gel was imaged by UV-shadowing, using the exposure wavelength of 254 nm (for nucleobase detection).

2.5. Differential scanning calorimetry

Differential scanning calorimetry (DSC) measurements were carried out on a nano-DSC (TA Instruments). The 11-mer-TBA DNA sequence was annealed in a buffer solution containing 10 mM potassium phosphate, 70 mM KCl (pH 7.0) and 10% (v/v) DMSO. DNA/**1** complex was obtained by adding ligand solution (4 M equiv.) to the folded G3 structure. The 11-mer-TBA concentration was fixed for all DSC experiments to 1.1 mg mL⁻¹. Scans were performed at 0.5 °C/min in the 2–95 °C temperature range. A buffer-buffer scan, under the same conditions, was subtracted from the corresponding buffer-sample endotherm scans and linear-polynomial baselines were drawn for each scan. The corrected thermograms were then normalized per mole of DNA to obtain the corresponding molar heat capacity curves. Model-free enthalpy ($\Delta H^0(T_m)$) for the overall unfolding of DNA structure was estimated by integrating the area under the heat capacity versus temperature curves and represent the average of at least three different heating experiments. T_m values correspond to the maximum of each thermogram peak. DSC curves provided the van't Hoff enthalpy ($\Delta H^0_{v,H}$), calculated assuming a simple two-state transition [24]. The reported errors on thermodynamic parameters are the deviations of the mean from the multiple determinations.

2.6. Nuclear Magnetic Resonance experiments

NMR experiments were performed on a 700 MHz Varian UnityINOVA spectrometer. One-dimensional proton spectra of the sample in H₂O were recorded at 5 °C using pulsed-field gradient DPGSE for H₂O suppression. DNA samples were prepared at 0.2–0.5 mM strand concentration in

0.6 mL (H₂O/D₂O 9:1) buffer solution. The solution was either 10 mM KH_2PO_4 , 70 mM KCl, 0.2 mM EDTA, pH 7.0 (for 11-mer-TBA) or 25 mM KH_2PO_4 , 70 mM KCl, 0.2 mM EDTA, pH 6.8 (for myc22). Aliquots of a stock solution of **1** in DMSO-d₆ were added directly to the DNA solution inside the NMR tube. The NMR data were processed on an iMAC running iNMR software (www.inmr.net).

2.7. Additional docking calculations and Molecular Dynamics (MD) simulations

By following the above-described protocol, compound **1** was also docked on the G4 target structures from the promoter regions of c-KIT (c-kit1, PDB ID: 4WO2) and c-MYC (myc22, PDB ID: 1XAV) oncogenes. Accurate analysis of the docking results revealed that two possible solutions were achieved for the dockings attained on the G3 and c-kit1 DNA structures, while a single well-clustered solution was achieved for the dockings attained on the myc22 structure. Thus five [**1**]/[DNA] complexes were created, two for the G3 and c-kit1 and one for the myc22 DNA. These binary complexes were further investigated by means of MD simulations carried out with the AMBER 14.0 package software (AMBER 14, University of California, San Francisco). Ligand force-field parameters were derived using the Antechamber program [25], and partial charges for the substrates were derived using the AM1-BCC procedure in Antechamber. MD simulations on the aforementioned five complexes were carried out in explicit solvent and periodic boundary conditions. 11, 21, and 22 K⁺ ions were added to the solvent bulk of the G3, c-KIT, and c-MYC complexes, respectively, to maintain neutrality in the systems. First, water shells and counterions were minimized using steepest descent and conjugate gradient algorithms. Then, a minimization of the entire ensemble was performed setting a convergence criterion on the gradient of 0.001 kcal mol⁻¹ Å⁻¹. Equilibration runs were carried out by heating the system to 300 K in 1 ns followed by 1 ns of constant pressure equilibration at 300 K. This was followed by 40 ns MD simulations in the NPT ensemble (constant temperature and pressure). The parm99 Amber force field for nucleic acids, modified using the recently developed parmbcsc0 parameters [26], was used for the DNA and the counterion, whereas the TIP3P model [27] was employed to explicitly represent water molecules. van der Waals and short-range electrostatic interactions were estimated within a 10 Å cut-off, whereas the long-range electrostatic interactions were assessed by using the particle mesh Ewald (PME) method [28] with a 1 Å charge grid spacing interpolated by a fourth-order B-spline, and by setting the direct sum tolerance to 10⁻⁵. Bonds involving hydrogen atoms were constrained by using the SHAKE algorithm [29] with a relative geometric tolerance for coordinate resetting of 0.00001 Å. Berendsen's coupling algorithms were employed to maintain constant temperature and pressure with the same scaling factor for both solvent and solutes and with the time constant for heat bath coupling maintained at 1.5 ps [30]. The pressure for the isothermal-isobaric ensemble was regulated by using a pressure relaxation time of 1 ps in Berendsen's algorithm. The simulations of the solvated [**1**]/[DNA] models were performed using a constant pressure of 1 atm and a constant temperature of 300 K. Analysis of MD trajectories was attained using the ptraj software [31]. This software was also used to calculate the average structures of the four complexes that were energy minimized by employing the same geometrical optimization protocol mentioned above. Visual inspections of the calculated complexes were attained using the MGL Tools [32] and the UCSF Chimera packages [33].

2.8. Cell cultures and immunofluorescence

Human osteosarcoma U2OS were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Carlo Erba), supplemented with 2 mM L-glutamine and 10% FBS (fetal bovine serum). U2OS cells were seeded in 35 mm dishes at a concentration of 100,000 cell/mL. Twenty-four hours after seeding, cells were treated with compound **1**. Cells were then fixed

in methanol:acetic acid (3:1), permeabilized with 0.1% triton- \times 100/PBS, and blocking was performed in 2% milk/PBS. Immunofluorescence was performed using standard methods with BG4, anti-FLAG (Cell Signaling Technology), and anti-rabbit Alexa 488-conjugated (Invitrogen) antibodies. Nuclei were stained with DAPI (Sigma-Aldrich), and coverslips were mounted with Mowiol 4–88 (Sigma-Aldrich). Fluorescence signal was determined using ImageJ software with the following formula: corrected total cell fluorescence (CTCF) = integrated density – (area of selected cell \times mean fluorescence of background readings). Significance was determined by parametric tests “Student’s *t*-test”. Plasmid for BG4 antibody purification was kindly obtained by S. Balasubramanian. BG4 protein expression was induced by the autoinduction method as described by Studier [34]. BG4 was purified using silica-based resin (Protino Ni-IDA, Macherey Nagel) pre-charged with Ni²⁺ ions, eluted with 250 mM Imidazole/PBS pH 8.0. The eluted antibody was concentrated and purified from imidazole using Amicon Ultra-15 centrifugal filter units (Millipore).

2.9. MTT cell proliferation assay

U2OS cells were seeded in 24-wells plate at a concentration of 20,000 cell/mL. Twenty-four hours after seeding, the compound was added to the cells. After 1 h or 24 h of treatment, the compound was removed and cells were further cultured in complete drug-free medium for 2 days. Then, thiazolyl blue tetrazolium bromide (MTT) (Sigma-Aldrich) was added to each well and incubated for 1 h at 37 °C. Next, the medium was removed and 300 μ L of dimethyl sulfoxide (Sigma-Aldrich) were added to cells and incubated for 1 h at room temperature. Finally, 100 μ L of the solution were transferred to a 96-well plate, and the absorbance at 595 nm was measured using a multiplate reader (Perkin Elmer). The linear regression parameters were determined to calculate the IC₅₀ (GraphPad Prism 4.0, Graph Pad Software Inc.).

3. Results and discussion

3.1. Compound selection

There are several studies using VS calculations for the identification of G4-targeting ligands [18,35], however none for the targeting of the G3 DNA structure. To this end, the intramolecular G3 structure formed by the d(GGTGGTGTGG) sequence (11-mer-TBA) was used as a target model (Fig. 1). This structure consists of two G-triads linked with two parallel G-strands (G1–G2 and G10–G11) interleaved with an antiparallel G-strand (G5–G6), which are connected together by two side loops (T3–T4, T7–G8–T9) [15,16]. Overall, the G3 formed by the 11-mer-TBA sequence maintains a structure that recalls the chair-like, antiparallel folding of TBA also in the absence of the fourth missing strand [36]. Starting from this structure, the docking software AD4.2 was used to dock a diversity set of a commercially available molecular database (see [Material and methods](#) for details) [37]. This VS process resulted in 15 compounds that were selected and purchased after the visual inspection of the theoretical results.

3.2. Circular dichroism studies

In order to identify true hits, the 15 computationally selected small molecules were first investigated by CD spectra and CD-melting analysis [38]. CD experiments were performed to examine the potential of compounds **1–15** to alter the native folding topology of the G3. The CD spectrum of the structure adopted by the 11-mer-TBA sequence alone showed two positive bands at 289 nm and 253 nm, and a small negative one below 240 nm (Fig. 3). These bands are characteristic of a homopolar stacking of the nucleobases, as expected for the antiparallel-stranded G3 folding adopted by the 11-mer-TBA sequence [15,16]. Then, DNA/ligand mixtures were obtained by adding ligands (4 M equiv.) to the folded 11-mer-TBA G3 structure. No significant variations of CD signal were observed upon addition of any compound (**1–15**) (data not

shown), clearly suggesting an overall preservation of the G3 architecture in the presence of each ligand.

The stabilizing properties of the compounds were evaluated by CD-melting experiments measuring the ligand-induced change in the melting temperature (ΔT_m) of G3 structure. CD melting experiments in the absence and presence of each compound (4:1 [compound]/[DNA] ratio) were recorded by following CD changes at the wavelength of maximum of intensity (289 nm) (Table S1 and Figs. S1–15). These experiments showed that 14 out of 15 compounds do not have the ability to significantly increase the T_m of the 11-mer-TBA structure ($\Delta T_m < 2$ °C), thus proving that they are nonbinders or weak binders of G3. On the other hand, these CD experiments allowed for the identification of a compound (**1**) that significantly increased the thermal stability of the 11-mer-TBA folded form and that could be considered as a promising G3 binder. Indeed, this compound turned out to have a good stabilizing effect with a ΔT_m of 8.5 °C at the investigated ligand/DNA ratio (Table S1 and Fig. 3). Therefore, we performed additional CD melting experiments for compound **1** by using a wider range of ligand concentrations to evaluate the effect of increasing concentrations of ligand on the stability of the G3 structure. Four different [ligand]/[DNA] ratios were examined (1:1, 2:1, 4:1, 8:1) (Fig. 3). Under the experimental conditions used, biphasic behavior was observed at low [ligand]/[DNA] ratios ($\leq 2:1$), which prevented the obtaining of accurate melting temperature data from those curves. The first melting transition most likely results from the fraction of the G3 that does not bind **1** and the second one from the fraction of DNA binding the compound. Overall, a dose-dependent increase of G3 thermal stability was observed up to 4:1 [ligand]/[DNA] ratio. Unfortunately, at 8:1 [ligand]/[DNA] ratio, a decrease of thermal stability was observed. This was probably an artifact generated by the progressive formation of a suspension inside the cuvette during the melting experiment.

3.3. Non-denaturing gel electrophoresis

To better clarify what happens to the 11-mer-TBA G3 upon addition of large amounts of compound **1**, ligand-DNA complexes were also studied by non-denaturing gel electrophoresis (PAGE). As shown in Fig. S16, the 11-mer-TBA G3 moves essentially as a single band in the gel, thus suggesting the presence of the unimolecular G3 structure as unique conformation in solution and the absence of high-order DNA structures [39]. Compound **1** addition to the G3 did not affect the G3 mobility until 4:1 [ligand]/[DNA] ratio, unless the appearance of some faint lower mobility bands. Conversely, in the presence of a large excess of **1** (i.e. 8:1 and 10:1 [ligand]/[DNA] ratio), a continuous dark background of unresolved polymeric structures was observed in the corresponding lanes. Overall, PAGE results are in perfect agreement with the observation from CD studies, suggesting that large amounts of ligand induce DNA aggregation.

3.4. Differential scanning calorimetry analysis

DSC experiments were carried out to study the thermodynamic stability of the 11-mer-TBA G3 in the absence and presence of **1**. The obtained DSC melting profiles are shown in Fig. 3c. The 11-mer-TBA G3 in the absence of ligand showed a highly reversible unfolding process, as the heating profiles of subsequent scans are superimposable, thus indicating that the observed unfolding/folding process is an equilibrium process [15]. The DSC curves show a maximum at T_m of 34.5 (± 0.5) °C and a $\Delta H^0(T_m)$ of 130 (± 5) kJ mol⁻¹ (very close to the calculated van't Hoff enthalpy) (Table 1). These values are in good agreement with what reported in the literature [15], and with the results obtained by CD reported herein. On the other hand, the 11-mer-TBA G3 in the presence of ligand results in non-reversible DSC profiles. The first DSC curve transition shows a maximum at T_m of 47.5 (± 0.5) °C and a $\Delta H^0(T_m)$ of 170 (± 5) kJ mol⁻¹. The total enthalpy measured in the DSC heating scans, which directly represents the amount of all the interactions existing in the folded DNA structure, suggests the presence of additional stabilizing interactions in the

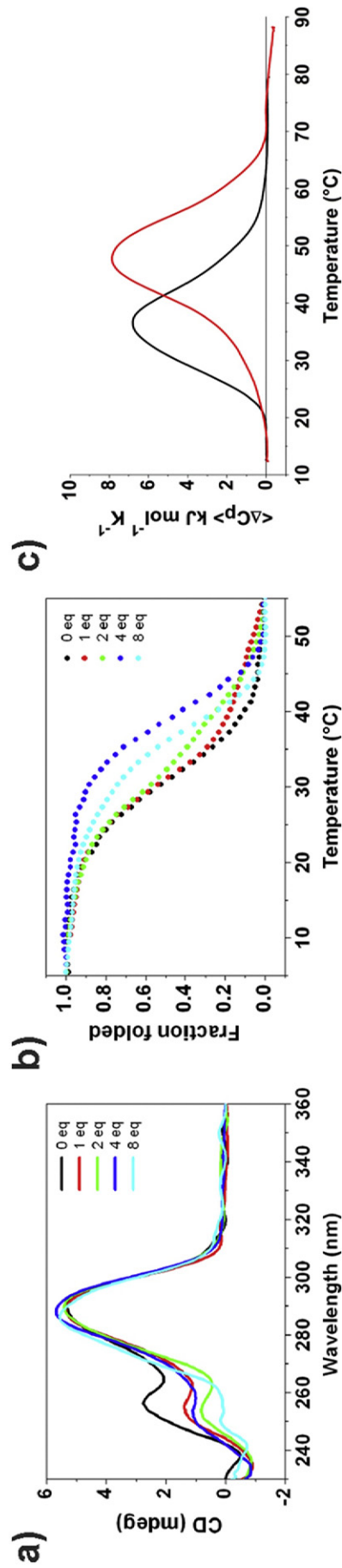


Fig. 3. (a) CD spectra and (b) CD melting experiments of 11-mer-TBA G3 in the absence and presence of increasing amounts of compound **1**. (c) DSC profiles for 11-mer-TBA G3 in the absence (black line) and presence (red line) of 4 M equivalents of compound **1**.

Table 1
Thermodynamic parameters for the unfolding of 11-mer-TBA in the absence and presence of **1** determined by DSC.

	T_m (°C)	$\Delta_{cal}H^\circ$ (kJ/mol)	$\Delta_{v,H}H^\circ$ (kJ/mol)
11-mer-TBA	34.5 ± 0.5	130 ± 5	128 ± 10
11-mer-TBA/ 1	47.5 ± 0.5	170 ± 5	–

presence of **1**. Overall, DSC data are indicative of the propensity of **1** to form stable interactions with G3-forming sequence resulting in the formation of a bound complex whose thermal stability is significantly enhanced compared to the native state.

3.5. Compound **1** interacts with parallel G4 structures

Since G3 is an intermediate in the folding process of G4-forming sequences, we decided to verify the ability of **1** to bind and stabilize different G4 structures in solution. To this purpose, several diverse G4-forming sequences able to form parallel, antiparallel and hybrid G4 structures were selected for this study. Particularly, two parallel G4-forming sequences from the *c-KIT* (*c-kit1*) and the *c-MYC* (*myc22*) promoter regions were used, in addition to TBA and telomeric Tel_{23} sequences, which are able to fold into antiparallel and hybrid G4 structures, respectively. The structures adopted by each G4 sample were verified by CD spectroscopy. *c-kit1* (Fig. S17) and *myc22* (Fig. S18) displayed the characteristic CD signature of parallel-stranded G4 topologies (positive and negative bands at 264 nm and around 240 nm, respectively) [38]. On the other hand, TBA showed a positive band at 295 nm and a negative band at around 240 nm (Fig. S19), characteristics of an antiparallel G4 conformation, while Tel_{23} showed a positive band at 289 nm with a shoulder at ca. 268 nm, and a weak negative band at around 240 nm, in agreement with the formation of a hybrid G4 structure (Fig. S20) [38]. These bands were not significantly modified upon addition of **1** (4 M equiv.) to any of the structures analyzed, clearly suggesting an overall preservation of their G4 architectures in the presence of ligand. Hence, the stabilizing properties of **1** on those structures were evaluated by recording CD melting experiments of DNAs in the absence and presence of ligand (4:1 ligand/DNA ratio), following the variations of the intensities of CD signals at the wavelength of 264, 295 and 289 nm for parallel, antiparallel and hybrid G4 s, respectively (Fig. S17–20). Results of these experiments, reported in Table 2, reveal interesting differences in the binding behavior of **1** to the different G4 topologies. In fact, whereas **1** demonstrated no increase in melting temperature (T_m) for the antiparallel and hybrid G4 structures, it exhibited good G4-stabilizing properties toward the parallel G4 structures adopted by *myc22* and *c-kit1* promoter sequences. Actually, the addition of this ligand exerted a more prominent thermal stabilization of *myc22* rather than *c-kit1* displaying $\Delta T_m > 15$ °C for *myc22* and of 4.5 °C for *c-kit1* (Table 2). We hypothesize that the more exposed terminal G-tetrad surface of the *myc22* G4 might be more conducive of favorable stacking interactions rather than in *c-kit1* [10b,21,40]. Overall, these results suggest a selective preference of **1** for parallel-type G4 folds over antiparallel ones, and reveal that the parallel fold of the *myc22* G4 is the preferred target of **1** among the investigated G4s.

Table 2
Ligand **1**-induced thermal stabilization of G4 and duplex DNAs measured by CD melting experiments.

Compound	ΔT_m (°C) ^a				
	<i>c-kit1</i>	<i>myc22</i>	Tel_{23}	TBA	ds12
1	4.5	>15.0	0.0	0.0	0.0

^a ΔT_m represents the difference in melting temperature [$\Delta T_m = T_m(DNA + 4 \text{ ligand equivalents}) - T_m(DNA)$]. All experiments were duplicated and the values reported are the average of two measurements.

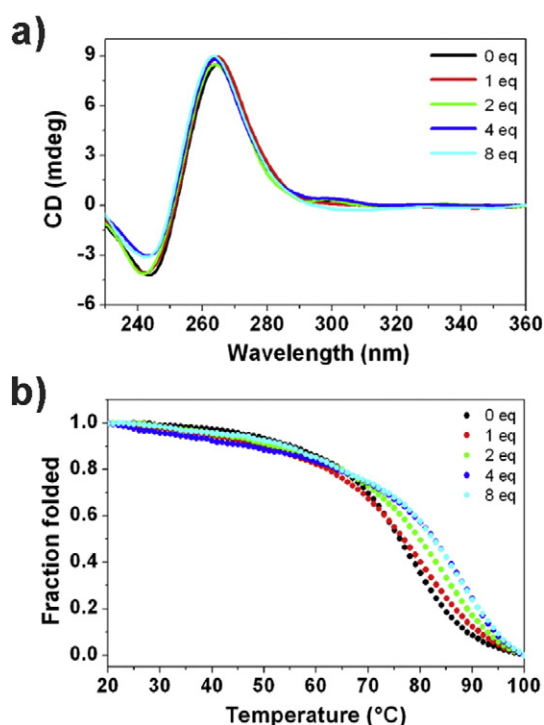


Fig. 4. (a) CD spectra and (b) CD melting experiments of *myc22* G4 in the absence and presence of increasing amounts of compound **1**.

Finally, the effect of increasing concentrations of **1** on the stability of *myc22* G4 structure was also investigated by carrying out CD melting experiments (Fig. 4). An increase of the T_m of this G4 was observed with every addition of ligand up to the 4:1 ratio, indicating that **1** could bind and stabilize the G4 in a specific manner (the addition of more than four equivalents of ligand does not lead to a further enhancement of G4 stability).

3.6. Compound **1** does not stabilize duplex DNA

The selectivity of **1** toward G3 and G4 structures over duplex DNA, was determined by comparing the above results with those obtained using a duplex DNA (ds12) [41], which is formed by a self-complementary 12-nt sequence. In the presence of K^+ , the CD spectrum of ds12 was characterized by a positive band at around 280 nm and a negative one at 250 nm, typical of values observed for duplex DNA (Fig. S21). These bands were not modified upon ligand addition. Furthermore, CD melting studies show that **1** was unable to induce any change in the melting temperature of the duplex structure, clearly indicating no affinity of this compound for the duplex DNA (Fig. S21).

3.7. Nuclear Magnetic Resonance (NMR) studies

NMR studies were carried out to obtain structural information about the binding mode of compound **1**. As stated above, this compound is able to stabilize the 11-mer-TBA G3 structure. In addition, it was also found to stabilize G4 structures in the parallel conformation, thus acting as a dual G3/G4 binder, with no affinity for the duplex DNA.

The spectral regions of the imino and aromatic protons of 11-mer-TBA G3 and *myc22* in the absence and presence of increasing amounts of compound **1** are shown in Fig. 5. The imino protons of either 11-mer-TBA G3 [15,16] and *myc22* [22,40] were previously assigned and these were used to monitor the shift of the signals and thus grasping the mode of interaction. Considering 11-mer-TBA G3, no appreciable shift of any signal was observed until 2:1 [ligand]/[DNA] ratio (Fig. 5a). However, at 2 M eq. of ligand, a severe broadening of all

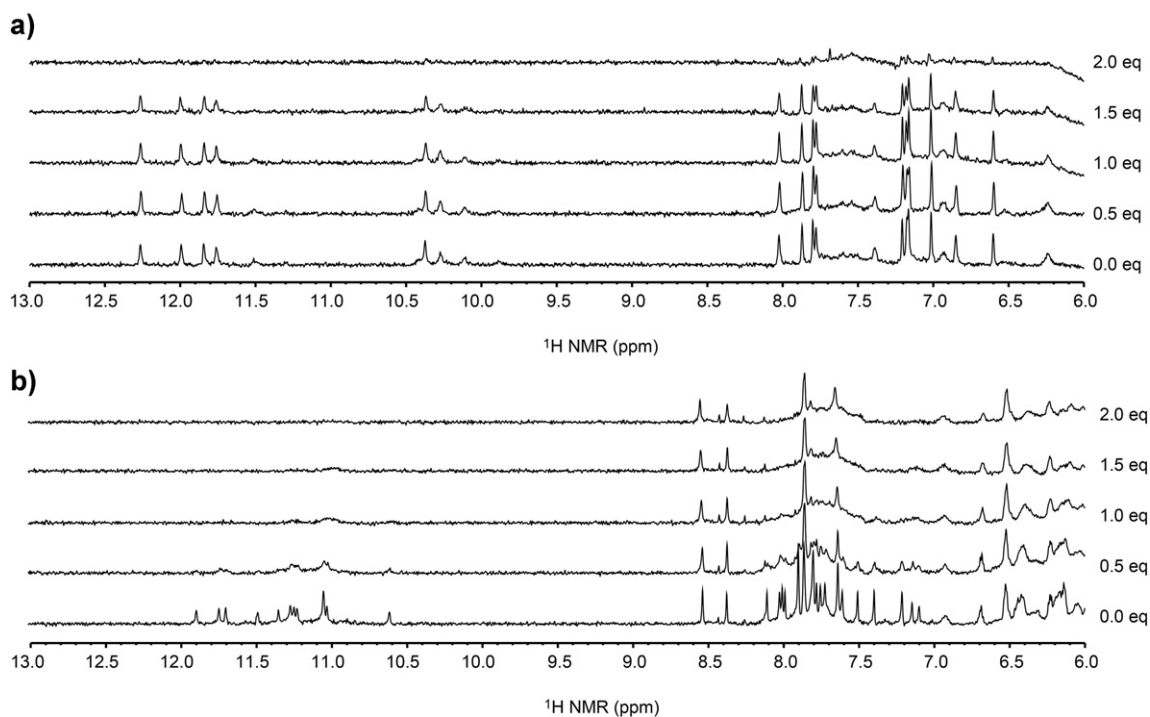


Fig. 5. Imino and aromatic proton regions of (a) the 11-mer-TBA G3 and (b) the myc22 G4 titrated with compound **1**.

DNA signals occurred, suggesting that i) the binding of **1** to the G3 is a dynamic process with a fast exchange rate on the NMR time-scale, and ii) probably additional non-specific ligand-DNA interactions occur.

Concerning myc22, imino protons became already broad upon addition of just 0.5 M eq. of **1** (Fig. 5b). Major perturbation was observed for the imino protons belonging to the external 5' G-tetrad (G7–G11–G16–G20) and 3' G-tetrad (G9–G13–G18–G22), while the remaining imino protons of the middle tetrad (G8–G12–G17–G21) were less affected. In contrast to G3, aromatic protons of myc22 G4 were poorly affected

by the binding. Altogether, these results suggest that end-stacking is the preferred binding mode of the ligand on this molecule.

3.8. Molecular Modeling studies

To better characterize the binding of **1** to the investigated G3 and G4 structures, theoretical studies were attained by performing docking experiments coupled with molecular dynamics simulations. In particular the ligand was docked into the published X-ray structure of *c-KIT* G4

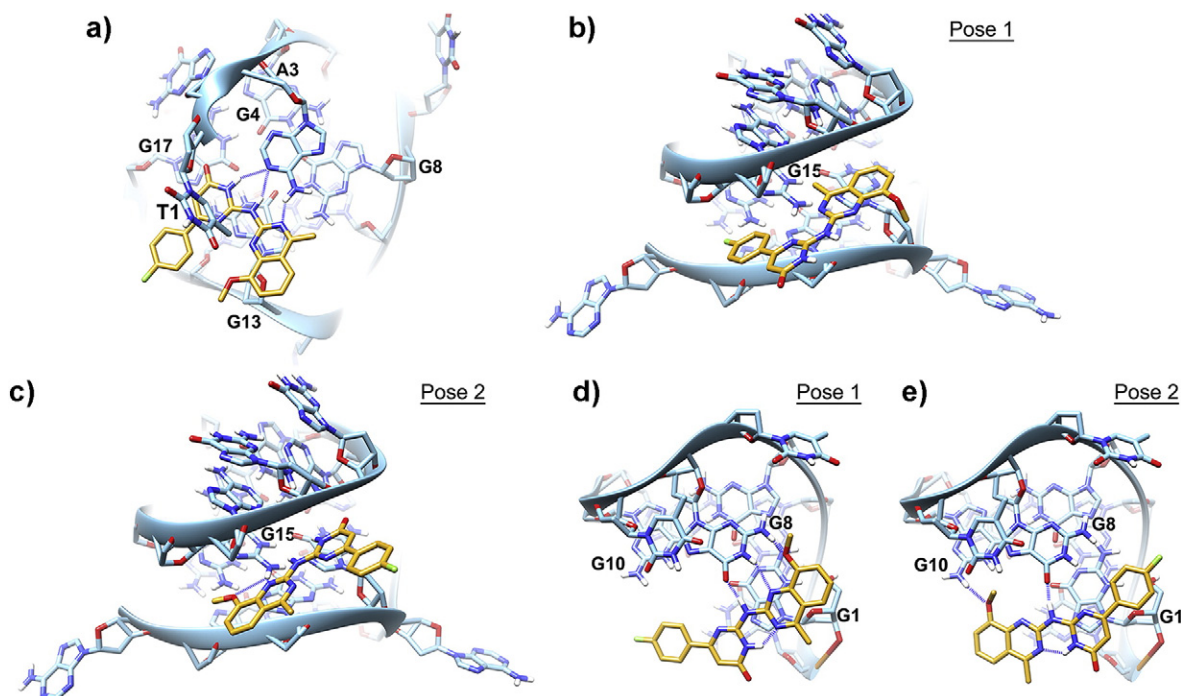


Fig. 6. Binding mode of **1** as calculated by the docking program in the structures of (a) *c-MYC*, (b and c) *c-KYT* G4s, and (d and e) G3. DNA is represented as blue sticks and ribbons while the ligand as yellow sticks. H-bond interactions are indicated as dashed blue lines while only interacting residues are labeled for clarity reasons.

(PDB ID: 4WO2), into the solution structures of *c*-MYC G4 (PDB ID: 1XAV) and into the structure of G3 (PDB ID: 2MKM). These calculations (also attained with the AD4.2 software) suggest that, in the *c*-MYC G4 structure, **1** is able to engage stacking interactions at the G4:G8:G13:G17 plane. Moreover, this pose is stabilized by H-bonds with the adjacent A3 residue while the pendant 4-fluorophenyl moiety seems to be solvent exposed (Fig. 6a). When docking simulations were attained on the *c*-KIT G4, two different solutions were suggested in which the ligand is proposed to occupy one of the G4 available grooves. In the first one, (herein referred to as pose 1) the ligand spans the entire groove by maximizing the van der Waals interactions between its hydrophobic side with the DNA and exposing its polar atoms towards the solvent. Conversely, in the second calculated binding pose (herein referred to as pose 2) the ligand occupies the same groove by exposing its polar groups towards the DNA. This allows the ligand to establish a double-bridged H-bond with the NH₂ group of G15 while the pendant hydrophobic *p*-fluorophenyl group is exposed to the solvent (Fig. 6b and c).

The presence of two possible binding poses was also suggested when **1** was docked in the G3 solution structure. As happened for the calculations achieved for the *c*-MYC G4, also the **1**-G3 recognition seems to be mediated by stacking interactions between the ligand and the DNA G1 residue with the two predicted binding poses differing for the ligand orientation with respect to the DNA structure. In both poses, the ligand seems to be able to establish H-bonds with the G8 residue, (Fig. 8d and e) while its methoxy substituent is alternatively able to form an additional H-bond with G8 and G10 (pose 1 and 2 in Fig. 8d and e, respectively). From the data reported so far, a preference for a single binding interaction pattern between the ligand and the target DNAs could not be given. In this respect, MD simulations allow to further study the viability of the calculated binding poses from a dynamical point of view. Thus, 40 ns explicit solvent MD simulations were performed for the 5 supposed complexes of compound **1** (one for the *c*-MYC G4 and two for the *c*-KIT G4 and G3), starting from the binding poses calculated by the docking software and taking advantage of the Amber 14 program. In regards to the **1**/*c*-MYC complex, plotting of the root mean square deviation (rmsd) of the ligand heavy atoms over the 40 ns production run reveals that the ligand adopts a stable binding

conformation after 6 ns which is kept for the remainder of the simulation time (Fig. 7a). In this binding mode, the ligand is sandwiched between the G4:G8:G13:G17 and T1 while additional H-bonds with the A3 residue are formed (Fig. 7b). Interestingly, plotting of the rmsd calculated for the DNA backbone atoms over time demonstrates that the starting DNA topology is preserved for the entire simulation time (with an average rmsd value of 2.1 Å).

Similar calculations were very helpful to probe the stability of the two possible binding modes calculated for the ligand in the *c*-KIT G4. In this complex, MD simulations for pose 1 suggest that the ligand is able to adopt a fairly stable conformation after around 9 ns in which, while still occupying the G4 groove, engages an additional H-bond interaction with the G15 NH₂ and a T-shaped charged transfer interaction with A1 residue. Analysis of the rmsd calculated considering the DNA backbone atoms allowed to infer that also in this case the DNA G4 topology is preserved during the entire simulation with an average value of 1.7 Å. On the other hand, plotting of the rmsd values over time of the ligand heavy atoms in pose 2 of the *c*-KIT G4 complex demonstrated that the ligand is unable to find a stable position within the G4 groove, thereby suggesting that pose 1 should give a more viable picture of the recognition between **1** and *c*-KIT G4.

Analysis of MD simulations attained starting from the two possible binding mode of the inspected ligand on the G3 structure indicate that both the binding poses are stable during the simulation time. Nevertheless, visualization of the trajectory that the ligand follows during the simulation time reveals that in pose 1 the ligand slightly begins to leave the binding site. To this end, plotting of the distances between the centers of mass calculated for the DNA atoms and the ligand ones outline that the ligand and the DNA move apart during the simulation time (Fig. 7c). The same analysis attained for pose 2 of the **1**/G3 complex demonstrates that this interaction is indeed more stable. Further confirmation is also given by the analysis of the rmsd plots calculated considering the DNA backbone atoms that would indicate a certain instability of the G3 topology when the ligand adopts the binding pose 1. In turn, these considerations would allow to give preference to pose 2 in which the ligand remains fairly stable in the G3 site during the MD production run and engages H-bond interactions with the T9 residue and charge-transfer contacts with G1.

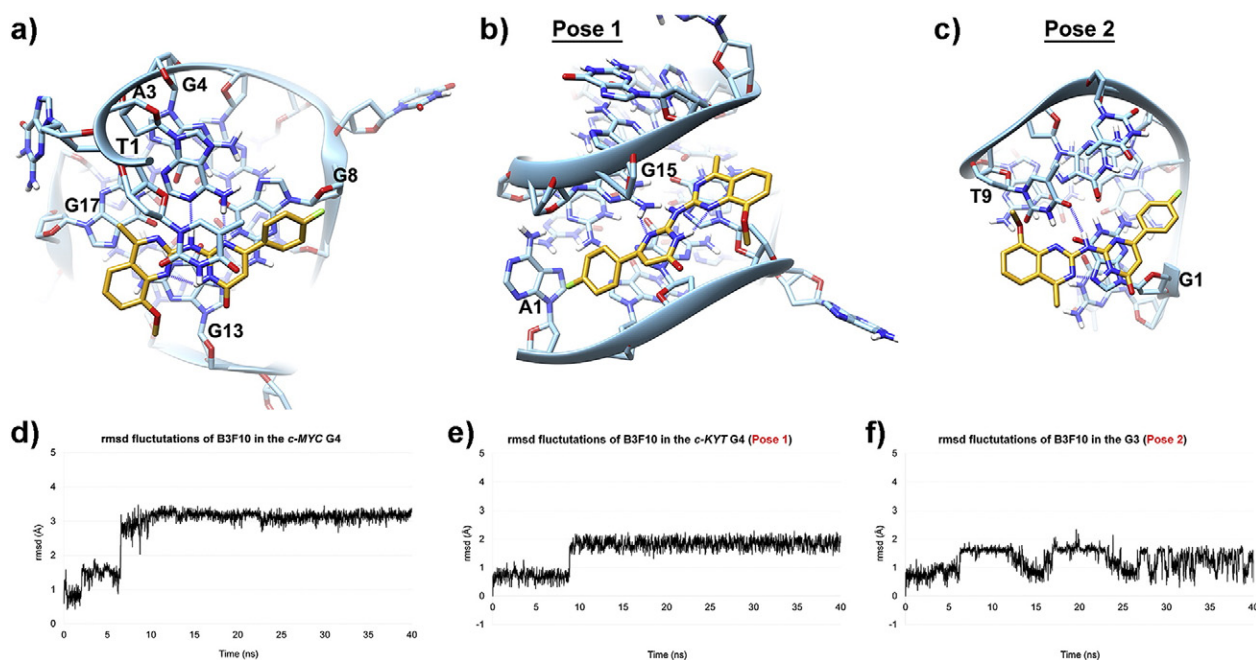


Fig. 7. Energy-minimized average structures of the complex between **1** and (a) *c*-MYC, (pose 1, b) *c*-KIT G4s, and (pose 2, c) G3. DNA is represented as blue sticks and ribbons while the ligand as yellow sticks. H-bond interactions are indicated as dashed blue lines while only interacting residues are labeled for clarity reasons. Panels d, e and f represent the plots of rmsd (Å) of the ligand heavy atoms over the 40 ns production run for the **1**/*c*-MYC G4 (d), **1**/*c*-KIT G4 (pose 1, e) and **1**/G3 (pose 2, f) complexes.

Table 3
Cytotoxic activity of compound **1** in human U2OS cells.^a

	IC ₅₀ (μM) - 1 h	IC ₅₀ (μM) - 24 h
1	>50	5.9 ± 1.88

^a The concentrations killing 50% of cells (IC₅₀) are means ± SE of three independent experiments performed in triplicates.

3.9. Cytotoxic activity of **1** and stabilization of G4 structures in human cancer cells

The cytotoxic effect of **1** was evaluated in human U2OS cells by MTT assay after 1 and 24 h of treatment. Short drug treatments did not show any ability to inhibit cells proliferation at concentrations below 50 μM (Table 3). Differently, after 24 h of treatment, the compound demonstrated cell killing activity with IC₅₀ = 5.9 μM.

We next asked if **1** was able to stabilize G4 structures in living cells. Nuclear G4 structures were visualized in human U2OS cells by

immunofluorescence microscopy using the BG4 antibody, a known specific antibody against G4 s [42]. A dose-response analysis was performed treating cells with **1** for 24 h at concentrations below the IC₅₀ (0.2 μM, 0.5 μM, 2 μM). Optical microscopy visualization before the immunostaining, revealed a minimal cell death only at 2 μM concentration. The fluorescence signal generated by incubation with a fluorochrome-labeled secondary antibody was visualized and analyzed as described in details in Experimental Section. As it is possible to see in Fig. 8, BG4 nuclear staining in control untreated cells showed a typical dotted signal, indicating that G4 structures were present at discrete chromatin sites in the nucleus, in agreement with previous studies [10b]. The drug treatment determines a slight but significant (*p*-value < 0.01) increase in the fluorescence intensity in the nucleus at all tested concentrations, demonstrating that **1** is a stabilizer of G4 structures in human U2OS cells.

Taken together the results suggest that **1** is able to stabilize G4 structures in cultured human cancer cells even if to a lesser extent compared to other molecules developed for being specific G4 stabilizers [10b]. This outcome is also in agreement with the CD-melting data indicating that this molecule is not able to bind and stabilize indiscriminately all the G4 structures.

4. Conclusions

In the present study, a tandem application of VS calculations along with the experimental investigation (CD, NMR, and DSC) was employed to discover a novel G3-targeting ligand as a potential anticancer drug. CD experiments provided evidence that, among the selected chemotypes, compound **1** was able to interact with and stabilize the G3 structure, which essentially conserved unaltered its native conformation upon interaction. However, the results of CD experiments revealed that the selected compound is actually a dual G3/G4 stabilizer, since it was also able to stabilize the parallel G4 DNA structures, while no duplex stabilizing properties were observed. Molecular modeling calculations were performed to provide a possible molecular basis for the binding of **1** to the G3 and G4 structures, highlighting the possible structural elements involved in the interaction. The analysis of the biological properties of the selected compound demonstrated that it is endowed with cytotoxic effect on human osteosarcoma cells.

In summary, these findings suggest that the identification of ligands with dual G3/G4 stabilizing properties may represent a new route for the design of G-rich DNA-targeting compounds, thus stimulating further studies aimed at the development of novel anticancer drugs.

Transparency document

The Transparency document associated with this article can be found, in online version.

Acknowledgements

This work was supported by the Italian Association for Cancer Research (AIRC) (IG-14150 to A.R., IG-16730 to B.P. and IG-15886 to G.C.), and by “Programma STAR” 2014 of University of Naples “Federico II” to B.P. (14-CSP3-C03-141).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.bbagen.2016.11.008>.

References

- [1] J.D. Watson, F.H. Crick, The structure of DNA, *Cold Spring Harb. Symp. Quant. Biol.* 18 (1953) 123–131.
- [2] (a) A. Bacolla, R.D. Wells, Non-B DNA conformations as determinants of mutagenesis and human disease, *Mol. Carcinog.* 48 (2009) 273–285;

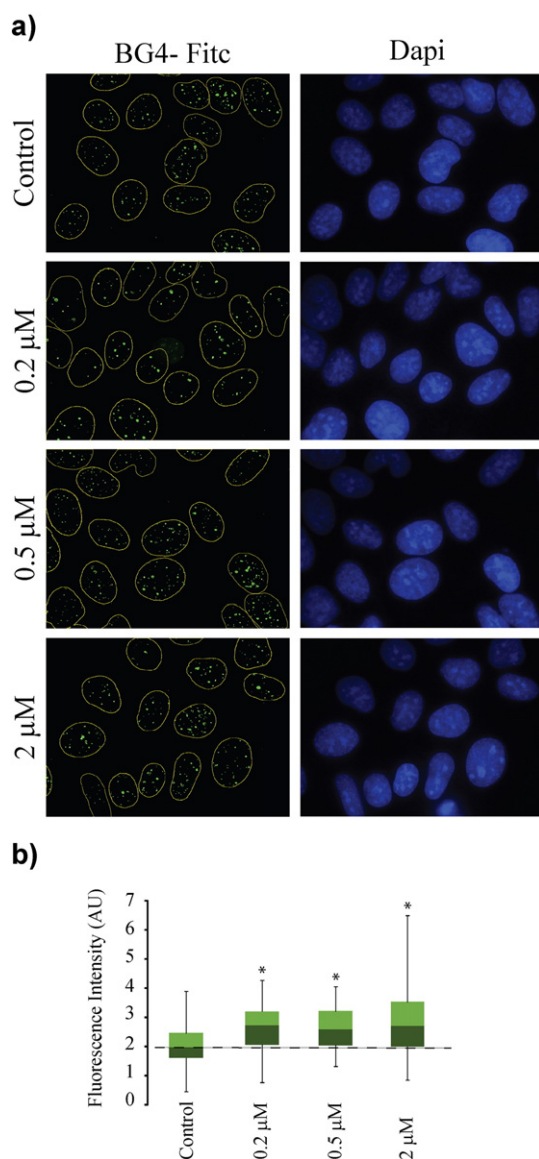


Fig. 8. (a) Immunofluorescence showing BG4 foci (green) in human osteosarcoma U2OS cell nuclei (blue). Cells were treated with compound **1** for 24 h and then fixed. G4s and DNA were stained with BG4 antibody and DAPI, respectively. (b) Box Plot analysis of G4 signal in the nucleus of control and treated cells. Statistical significance was determined with the Student's *t*-test (**P* < 0.05).

- (b) M. Kaushika, S. Kaushikb, K. Royb, A. Singhb, S. Mahendrub, M. Kumarb, S. Chaudharyb, S. Ahmedb, S. Kukretib, A bouquet of DNA structures: emerging diversity, *Biochem. Biophys. Rep.* 5 (2016) 388–395.
- [3] A.H. Wang, G.J. Quigley, F.J. Kolpak, J.L. Crawford, J.H. van Boom, G. van der Marel, A. Rich, Molecular structure of a left-handed double helical DNA fragment at atomic resolution, *Nature* 282 (1979) 680–686.
- [4] Y.M. Agazie, G.D. Burkholder, J.S. Lee, Triplex DNA in the nucleus: direct binding of triplex-specific antibodies and their effect on transcription, replication and cell growth, *Biochem. J.* 316 (1996) 461–466.
- [5] A. Jain, G. Wang, K.M. Vasquez, DNA Triple Helices: biological consequences and therapeutic potential, *Biochimie* 90 (2008) 1117–1130.
- [6] D. Rhodes, H.J. Lipps, G-quadruplexes and their regulatory roles in biology, *Nucleic Acids Res.* 43 (2015) 8627–8637.
- [7] M.L. Bochman, K. Paeschke, V.A. Zakian, DNA secondary structures: stability and function of G-quadruplex structures, *Nat. Rev. Genet.* 13 (2012) 770–780.
- [8] J. Amato, N. Iaccarino, A. Randazzo, E. Novellino, B. Pagano, Noncanonical DNA secondary structures as drug targets: the prospect of the i-motif, *ChemMedChem* 9 (2014) 2026–2030.
- [9] (a) S. Balasubramanian, L.H. Hurley, S. Neidle, Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discov.* 10 (2011) 261–275;
- (b) R. Simone, P. Fratta, S. Neidle, G.N. Parkinson, A.M. Isaacs, G-quadruplexes: Emerging roles in neurodegenerative diseases and the non-coding transcriptome, *FEBS Lett.* 589 (2015) 1653–1668.
- [10] (a) S. Neidle, Quadruplex nucleic acids as novel therapeutic targets, *J. Med. Chem.* 59 (2016) 5987–6011;
- (b) J. Amato, R. Morigi, B. Pagano, A. Pagano, S. Ohnmacht, A. De Magis, Y.P. Tiang, G. Capranico, A. Locatelli, A. Graziadio, A. Leoni, M. Rambaldi, E. Novellino, S. Neidle, A. Randazzo, Toward the development of specific G-quadruplex binders: synthesis, biophysical, and biological studies of new hydrazone derivatives, *J. Med. Chem.* 59 (2016) 5706–5720;
- (c) D. Musumeci, J. Amato, A. Randazzo, E. Novellino, C. Giancola, D. Montesarchio, B. Pagano, G-quadruplex on oligo affinity support (G4-OAS): an easy affinity chromatography-based assay for the screening of G-quadruplex ligands, *Anal. Chem.* 86 (2014) 4126–4130;
- (d) J. Amato, N. Iaccarino, B. Pagano, R. Morigi, A. Locatelli, A. Leoni, M. Rambaldi, P. Zizza, A. Biroccio, E. Novellino, A. Randazzo, Bis-indole derivatives with anti-tumor activity turn out to be specific ligands of human telomeric G-quadruplex, 2014. *Front. Chem.* 2 54, <http://dx.doi.org/10.3389/fchem.2014.00054>.
- [11] (a) J.L. Huppert, S. Balasubramanian, G-quadruplexes in promoters throughout the human genome, *Nucleic Acids Res.* 35 (2007) 406–413;
- (b) A.L. Moye, K.C. Porter, S.B. Cohen, T. Phan, K.G. Zyner, N. Sasaki, G.O. Lovrecz, J.L. Beck, T.M. Bryan, Telomeric G-quadruplexes are a substrate and site of localization for human telomerase, *Nat. Commun.* 6 (2015) 7643.
- [12] A. Rajendran, M. Endo, K. Hidaka, H. Sugiyama, Direct and single-molecule visualization of the solution-state structures of G-hairpin and G-triplex intermediates, *Angew. Chem. Int. Ed. Eng.* 53 (2014) 4107–4112.
- [13] X. Wu, P. Xu, J. Wang, Y. Xu, T. Fu, M. Zhao, D. Zhang, J. Liu, H. Shen, Z. Xiu, G. Li, Theoretical studies on the folding mechanisms for different DNA G-quadruplexes, *Adv. Exp. Med. Biol.* 827 (2015) 123–141.
- [14] X.M. Hou, W.Q. Wu, X.L. Duan, N.N. Liu, H.H. Li, J. Fu, S.X. Dou, M. Li, X.G. Xi, Molecular mechanism of G-quadruplex unwinding helicase: sequential and repetitive unfolding of G-quadruplex by Pif1 helicase, *Biochem. J.* 466 (2015) 189–199.
- [15] V. Limongelli, S. De Tito, L. Cerofolini, M. Fragai, B. Pagano, R. Trotta, S. Cosconati, L. Marinelli, E. Novellino, I. Bertini, A. Randazzo, C. Luchinat, M. Parrinello, The G-triplex DNA, *Angew. Chem. Int. Ed. Eng.* 52 (2013) 2269–2273.
- [16] L. Cerofolini, J. Amato, A. Giachetti, V. Limongelli, E. Novellino, M. Parrinello, M. Fragai, A. Randazzo, C. Luchinat, G-triplex structure and formation propensity, *Nucleic Acids Res.* 42 (2014) 13393–13404.
- [17] (a) G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Har, R.K. Belew, A.J. Olson, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, *J. Comput. Chem.* 19 (1998) 1639–1662;
- (b) R. Huey, G.M. Morris, A.J. Olson, D.S. Goodsell, A semiempirical free energy force field with charge-based desolvation, *J. Comput. Chem.* 28 (2007) 1145–1152;
- (c) S. Cosconati, A.L. Perryman, R. Harris, D.S. Goodsell, A.J. Olson, Virtual screening with AutoDock: theory and practice, *Expert Opin. Drug Discovery* 5 (2010) 597–607.
- [18] (a) S. Cosconati, L. Marinelli, R. Trotta, A. Virno, L. Mayol, E. Novellino, A.J. Olson, A. Randazzo, Tandem application of virtual screening and NMR experiments in the discovery of brand new DNA quadruplex groove binders, *J. Am. Chem. Soc.* 131 (2009) 16336–16337;
- (b) S. Cosconati, A. Rizzo, R. Trotta, B. Pagano, S. Iachettini, S. De Tito, I. Lauri, I. Fotticchia, M. Giustiniano, L. Marinelli, C. Giancola, E. Novellino, A. Biroccio, A. Randazzo, Shooting for selective drug-like G-quadruplex binders: evidence for telomeric DNA damage and tumor cell death, *J. Med. Chem.* 55 (2012) 9785–9792;
- (c) F.S. Di Leva, P. Zizza, C. Cingolani, C. D'Angelo, B. Pagano, J. Amato, E. Salvati, C. Sissi, O. Pinato, L. Marinelli, A. Cavalli, S. Cosconati, E. Novellino, A. Randazzo, A. Biroccio, Exploring the chemical space of G-quadruplex binders: discovery of a novel chemotype targeting the human telomeric sequence, *J. Med. Chem.* 56 (2013) 9646–9654.
- [19] J.J. Irwin, B.K. Shoichet, ZINC – a free database of commercially available compounds for virtual screening, *J. Chem. Inf. Model.* 45 (2005) 177–182.
- [20] C.R. Cantor, M.M. Warshaw, H. Shapiro, Oligonucleotide interactions. III. Circular dichroism studies of the conformation of deoxyoligonucleotides, *Biopolymers* 9 (1970) 1059–1077.
- [21] D. Wei, J. Husby, S. Neidle, Flexibility and structural conservation in a c-KIT G-quadruplex, *Nucleic Acids Res.* 43 (2015) 629–644.
- [22] A. Ambrus, D. Chen, J. Dai, R.A. Jones, D. Yang, Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization, *Biochemistry* 44 (2005) 2048–2058.
- [23] B. Pagano, J. Amato, N. Iaccarino, C. Cingolani, P. Zizza, A. Biroccio, E. Novellino, A. Randazzo, Looking for efficient G-quadruplex ligands: evidence for selective stabilizing properties and telomere damage by drug-like molecules, *ChemMedChem* 10 (2015) 640–649.
- [24] B. Pagano, A. Randazzo, I. Fotticchia, E. Novellino, L. Petraccone, C. Giancola, Differential scanning calorimetry to investigate G-quadruplexes structural stability, *Methods* 64 (2013) 43–51.
- [25] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, *J. Comput. Chem.* 25 (2004) 1157–1174.
- [26] (a) W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz Jr., D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, A second generation force field for the simulation of proteins, nucleic acids, and organic molecules, *J. Am. Chem. Soc.* 117 (1995) 5179–5197;
- (b) T.E. Cheatham III, P. Cieplak, P.A. Kollman, A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat, *J. Biomol. Struct. Dyn.* 16 (1999) 845–862;
- (c) A. Pérez, I. Marchán, D. Svozil, J. Sponer, T.E. Cheatham III, C.A. Lughton, M. Orozco, Refinement of the AMBER force field for nucleic acids: improving the description of α/γ conformers, *Biophys. J.* 92 (2007) 3817–3829.
- [27] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, L.M. Klein, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.* 79 (1983) 926–935.
- [28] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, *J. Chem. Phys.* 103 (1995) 8577–8593.
- [29] J.P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J. Comput. Phys.* 23 (1977) 327–341.
- [30] H.J.C. Berendsen, J.P.M. Postma, W.F. Van Gunsteren, A. Di Nola, J.R. Haak, Molecular dynamics with coupling to an external bath, *J. Chem. Phys.* 81 (1984) 3684–3690.
- [31] J. Shao, S.W. Tanner, N. Thompson, T.E. Cheatham, Clustering molecular dynamics trajectories: I. Characterizing the performance of different clustering algorithms, *J. Chem. Theory Comput.* 3 (2007) 2312–2334.
- [32] M.F. Sanner, Python: a programming language for software integration and development, *J. Mol. Graph. Model.* 17 (1999) 57–61.
- [33] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera – a visualization system for exploratory research and analysis, *J. Comput. Chem.* 25 (2004) 1605–1612.
- [34] F.W. Studier, Protein production by auto-induction in high density shaking cultures, *Protein Expr. Purif.* 41 (2005) 207–234.
- [35] (a) S. Alcaro, C. Musetti, S. Distinto, M. Casati, G. Zagotto, A. Artese, L. Parrotta, F. Moraca, G. Costa, F. Ortuso, E. Maccioni, C. Sissi, 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 (2013) 843–855;
- (b) R. Rocca, G. Costa, A. Artese, L. Parrotta, F. Ortuso, E. Maccioni, O. Pinato, M.L. Greco, C. Sissi, S. Alcaro, S. Distinto, F. Moraca, Hit identification of a novel dual binder for h-telo/c-myc G-quadruplex by a combination of pharmacophore structure-based virtual screening and docking refinement, *ChemMedChem* 11 (2016) 1721–1733.
- [36] (a) P. Schultze, R.F. Macaya, J. Feigon, Three-dimensional solution structure of the thrombin-binding DNA aptamer d(GGTTGGTGGTGG), *J. Mol. Biol.* 235 (1994) 1532–1547;
- (b) S. De Tito, F. Morvan, A. Meyer, J.J. Vasseur, A. Cummaro, L. Petraccone, B. Pagano, E. Novellino, A. Randazzo, C. Giancola, D. Montesarchio, Fluorescence enhancement upon G-quadruplex folding: synthesis, structure, and biophysical characterization of a dansyl/cyclodextrin-tagged thrombin binding aptamer, *Bioconjug. Chem.* 24 (2013) 1917–1927.
- [37] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson, AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility, *J. Comput. Chem.* 30 (2009) 2785–2791.
- [38] (a) S. Masiero, R. Trotta, S. Pieraccini, S. De Tito, R. Perone, A. Randazzo, G.P. Spada, A non-empirical chromophoric interpretation of CD spectra of DNA G-quadruplex structures, *Org. Biomol. Chem.* 8 (2010) 2683–2692;
- (b) A.I. Karsisiotis, N.M. Hessari, E. Novellino, G.P. Spada, A. Randazzo, M. Webba Da Silva, Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism, *Angew. Chem. Int. Ed.* 50 (2011) 10645–10648;
- (c) A. Randazzo, G.P. Spada, M. Webba Da Silva, Circular dichroism of quadruplex structures, *Top. Curr. Chem.* 330 (2013) 67–86.
- [39] L. Petraccone, A. Malafrente, J. Amato, C. Giancola, G-quadruplexes from human telomeric DNA: how many conformations in PEG containing solutions? *J. Phys. Chem. B* 116 (2012) 2294–2305.
- [40] J. Dai, M. Carver, L.H. Hurley, D. Yang, 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 (2011) 17673–17680.
- [41] (a) R.E. Dickerson, H.R. Drew, Structure of a B-DNA dodecamer. II. Influence of base sequence on helix structure, *J. Mol. Biol.* 149 (1981) 761–768;
- (b) H.R. Drew, R.E. Dickerson, Structure of a B-DNA dodecamer. III. Geometry of hydration, *J. Mol. Biol.* 151 (1981) 535–556.
- [42] G. Biffi, D. Tannahill, J. McCafferty, S. Balasubramanian, Quantitative visualization of DNA G-quadruplex structures in human cells, *Nat. Chem.* 5 (2013) 182–186.



Common G-Quadruplex Binding Agents Found to Interact With i-Motif-Forming DNA: Unexpected Multi-Target-Directed Compounds

Alessia Pagano^{1†}, Nunzia Iaccarino^{1†}, Mahmoud A. S. Abdelhamid², Diego Brancaccio¹, Emanuele U. Garzarella¹, Anna Di Porzio¹, Ettore Novellino¹, Zoë A. E. Waller^{2,3}, Bruno Pagano¹, Jussara Amato^{1*} and Antonio Randazzo^{1*}

OPEN ACCESS

Edited by:

Stefano Alcaro,
Università degli studi Magna Græcia di
Catanzaro, Italy

Reviewed by:

Juan Carlos Morales,
Consejo Superior de Investigaciones
Científicas (CSIC), Spain
Janez Plavec,
National Institute of Chemistry
Slovenia, Slovenia
Claudia Sissi,
Università degli Studi di Padova, Italy

*Correspondence:

Jussara Amato
jussara.amato@unina.it
Antonio Randazzo
antonio.randazzo@unina.it

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Medicinal and Pharmaceutical
Chemistry,
a section of the journal
Frontiers in Chemistry

Received: 27 April 2018

Accepted: 22 June 2018

Published: 24 July 2018

Citation:

Pagano A, Iaccarino N,
Abdelhamid MAS, Brancaccio D,
Garzarella EU, Di Porzio A, Novellino E,
Waller ZAE, Pagano B, Amato J and
Randazzo A (2018) Common
G-Quadruplex Binding Agents Found
to Interact With i-Motif-Forming DNA:
Unexpected Multi-Target-Directed
Compounds. *Front. Chem.* 6:281.
doi: 10.3389/fchem.2018.00281

¹ Department of Pharmacy, University of Naples Federico II, Naples, Italy, ² School of Pharmacy, University of East Anglia, Norwich Research Park, Norwich, United Kingdom, ³ Centre for Molecular and Structural Biochemistry, University of East Anglia, Norwich Research Park, Norwich, United Kingdom

G-quadruplex (G4) and i-motif (iM) are four-stranded non-canonical nucleic acid structural arrangements. Recent evidences suggest that these DNA structures exist in living cells and could be involved in several cancer-related processes, thus representing an attractive target for anticancer drug discovery. Efforts toward the development of G4 targeting compounds have led to a number of effective bioactive ligands. Herein, employing several biophysical methodologies, we studied the ability of some well-known G4 ligands to interact with iM-forming DNA. The data showed that the investigated compounds are actually able to interact with both DNA *in vitro*, thus acting *de facto* as multi-target-directed agents. Interestingly, while all the compounds stabilize the G4, some of them significantly reduce the stability of the iM. The present study highlights the importance, when studying G4-targeting compounds, of evaluating also their behavior toward the i-motif counterpart.

Keywords: G-quadruplex, i-motif, Berberine, BRACO-19, Mitoxantrone, Phen-DC3, Pyridostatin, RHPS4

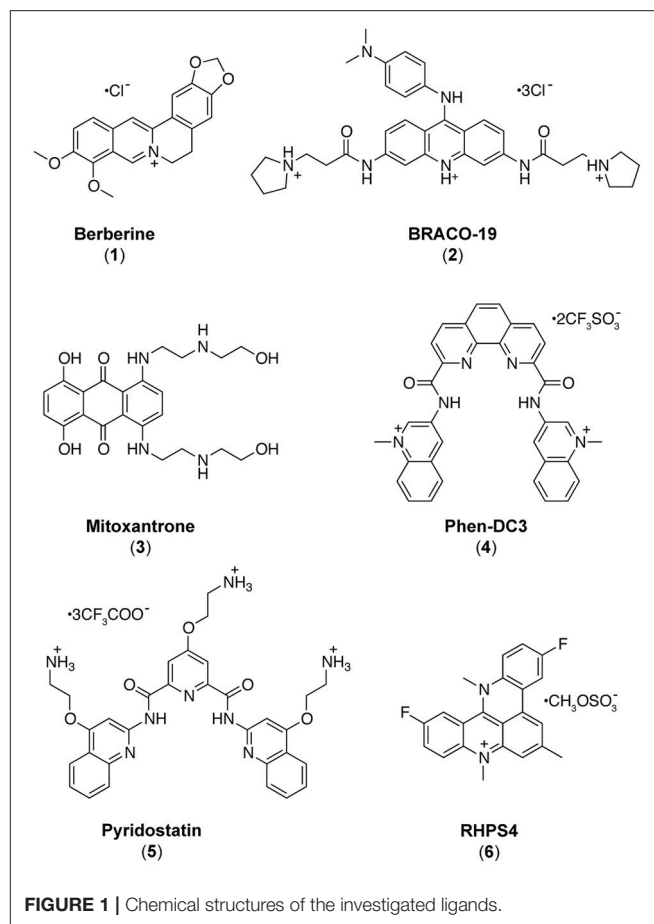
INTRODUCTION

GC-rich nucleic acids are able to form a variety of non-canonical secondary structures (Zhao et al., 2010; Cerofolini et al., 2014). The best studied of these are G-quadruplexes (G4s), four stranded alternative nucleic acid secondary structures formed from guanine-rich DNA or RNA composed of stacked tetrads of guanines formed by Hoogsteen hydrogen bonding (Burge et al., 2006). Sequences which can form G4s are prevalent within regulatory regions of the genome, particularly within the promoter region of genes (Huppert and Balasubramanian, 2007; Chambers et al., 2015; Bedrat et al., 2016). Good evidence has been provided to support the hypothesis that G4s exist in human cells (Biffi et al., 2013), play a role in human diseases (Haeusler et al., 2014; Maizels, 2015) and can be targeted with ligands to modulate biological functions (Siddiqui-Jain et al., 2002; Lam et al., 2013; Zizza et al., 2016).

More recently, increasing interest is being paid to the i-motif (iM) structure, another four stranded structure which can form in sequences rich in cytosine, composed of two intercalated hairpins, stabilized by hemi-protonated cytosine-cytosine⁺ (C·C⁺) base pairs (Gehring et al., 1993). Putative iM forming sequences also occur throughout the genome (Wright et al., 2016b;

Fleming et al., 2017), typically opposing regions which can form G4s, though the sequence requirements for stable formation are somewhat different. Studies on the iM were previously limited based on the assumption that because they are stabilized in slightly acidic conditions they were not physiologically relevant, despite a solid foundation of data indicating that these structures are detectable at neutral pH *in vitro* (Mergny et al., 1995). Over the years, this assumption has been challenged with examples of iMs which can form at neutral pH, at low temperature (Zhou et al., 2010), under conditions of negative superhelicity (Sun and Hurley, 2009), and molecular crowding (Rajendran et al., 2010) conditions. Further examples of sequences which are naturally stable at neutral pH have been found in the genome, initially by investigating sequences which oppose G4s (Brazier et al., 2012), but multiple other examples have followed (Wright et al., 2016b; Fleming et al., 2017; Mir et al., 2017). Just recently further evidence for i-motif formation *in vivo* has been provided by in cell NMR experiments (Dzatko et al., 2018) and the discovery of an antibody that binds i-motif specifically in the nuclei of human cells (Zeraati et al., 2018). The C-rich regions of genomes are of particular interest because cytosine forms part of the basis for epigenetic regulation, and epigenetic modification of cytosine has been found to alter the stability of iM (Bhavsar-Jog et al., 2014; Xu et al., 2015; Wright et al., 2017). Moreover, ligands which bind and stabilize iM have been shown to modulate biological functions (Amato et al., 2014b). For example, stabilization of the iM forming sequence in the human telomere was found to inhibit telomerase activity and interfere with telomere biology (Li et al., 2006; Chen et al., 2012); stabilization of iM forming sequence in the promoter region of BCL2 was found to cause an increase in gene expression (Kang et al., 2014; Kendrick et al., 2014) and an iM interacting compound was found to downregulate PDGFR- β promoter activity (Brown et al., 2017).

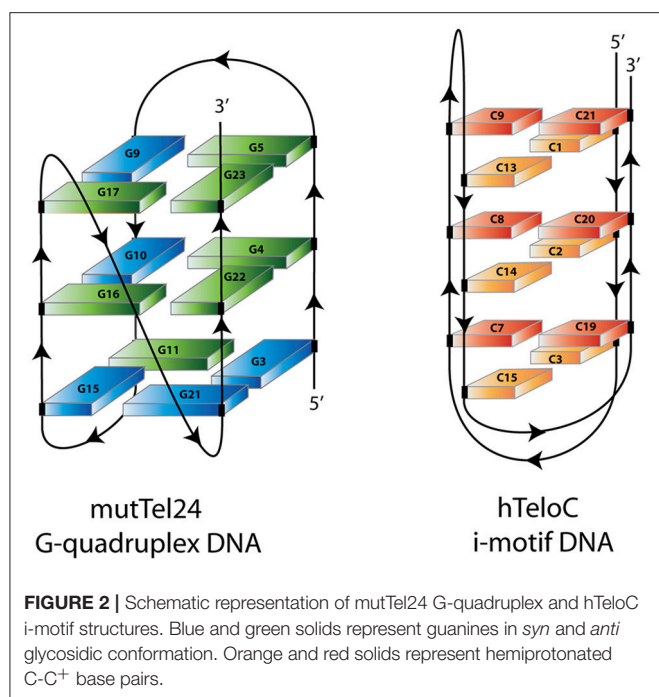
As it appears that formation of iM and/or G4 structures could incite different biological outcomes, it is important to understand the potential structures a compound is able to interact with. In contrast to the hundreds of G4 binding ligands (Pagano et al., 2007, 2010, 2015; Di Leva et al., 2013; Li et al., 2013; Amato et al., 2014a, 2018), there are comparatively very few iM binding compounds reported in the literature (Day et al., 2014). Some ligands which were described to bind G4 have also been found to bind iM (Fedoroff et al., 2000; Wright et al., 2016a; Xu et al., 2016), so we decided to assess and compare the capability to interact with iM-forming DNA of several known bioactive G4 binding agents: Berberine (1) (Franceschin et al., 2006), BRACO-19 (2) (Gowan et al., 2002), Mitoxantrone (3) (Huang et al., 2007), Phen-DC3 (4) (De Cian et al., 2007a), Pyridostatin (5) (Rodriguez et al., 2008), and RHPS4 (6) (Izbicka et al., 1999) (Figure 1). The interaction of these compounds with G4- and iM-forming sequences were investigated *in vitro* in different experimental conditions employing several biophysical methodologies (Pagano et al., 2012) (Figure 2). The data unequivocally demonstrate that, even if in different ways, actually these molecules interact with both DNA, thus acting *de facto* as multi-target-directed compounds.



MATERIALS AND METHODS

Oligonucleotide Synthesis and Sample Preparation

An ABI 394 DNA/RNA synthesizer (Applied Biosystem) was employed to prepare DNA sequences at 5- μ mol scale by using standard β -cyanoethylphosphoramidite solid phase chemistry. The subsequent detachment of DNA from support and its deprotection were carried out by means of an aqueous solution of concentrated ammonia at 55°C for 12 h. The filtrates and the washings were combined and concentrated under reduced pressure, solubilized in water, and then purified by high-performance liquid chromatography (HPLC) equipped with a Nucleogel SAX column (Macherey-Nagel, 1000-8/46). Two buffers were employed for the purification: buffer A, consisting of a 20 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ aqueous solution (pH 7.0) and containing 20% (v/v) CH_3CN , and buffer B, consisting of 1 M KCl, 20 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ aqueous solution (pH 7.0), containing 20% (v/v) CH_3CN , combined in a 30 min linear gradient going from 0 to 100% B with a flow rate of 1 mL/min. The purified fractions of the oligomers were then desalted by using C-18 cartridges (Sep-pak). The purity of the isolated oligomer was evaluated by NMR and it turned out to be higher than 98%. In particular, the



following oligonucleotides were employed for the experiments: d[CCCT(AACCCT)₃] (hTeloC) and d[(TT(GGGTTA)₃GGA)] (mutTel24). The oligonucleotide concentrations were established by measuring the UV absorption at 90°C taking into account the molar extinction coefficient values ϵ ($\lambda = 260$ nm) determined by the nearest neighbor model (Cantor et al., 1970). hTeloC and mutTel24 were dissolved in 10 mM sodium phosphate buffer and 10 mM potassium phosphate buffer, respectively, at different pH values before establishing the experimental conditions to be used (pH 4.3 and pH 5.7). DNA samples were heated at 90°C for 5 min, and then gradually cooled to room temperature overnight.

Circular Dichroism Spectroscopy

Circular dichroism (CD) experiments were recorded on a Jasco J-815 spectropolarimeter equipped with a PTC-423S/15 Peltier temperature controller. Each spectrum was recorded in the 220–360 nm wavelength range, averaged over three scans and subtracted of the buffer baseline. The scan rate was set to 100 nm/min, with a 1 s response time, and 1 nm bandwidth. Spectra were analyzed using Origin 7.0 software. CD experiments (spectra and melting) were performed using 10–15 μ M oligonucleotide concentration, in the absence and presence of 5 molar equivalents of ligands (10 mM in DMSO). CD melting were performed at 1°C/min heating rate in the 5–90 and 20–100°C temperature range for hTeloC and mutTel24, respectively. Changes of CD signal were followed at the wavelengths of the maximum CD intensity, 288 and 290 nm for hTeloC and mutTel24, respectively. The melting temperatures ($T_{1/2}$) were mathematically calculated by using the curve fitting function in Origin 7.0 software. $\Delta T_{1/2}$ values represent the difference between the melting temperature of the DNA with and without ligands.

UV-Melting

A JASCO V-730 UV-visible spectrophotometer equipped with a Peltier temperature controller was employed to perform the UV thermal denaturation experiments. The oligonucleotide concentrations were 10 μ M for both hTeloC and mutTel24 DNA in the appropriate buffer, as indicated above. Experiments were performed by following changes of UV signal at 295 nm, at a heating rate of 1°C/min, in the temperature ranges of 5–100 and 20–100°C for hTeloC and mutTel24, respectively. The melting temperatures ($T_{1/2}$) were mathematically calculated by using the curve fitting function in Origin 7.0 software. $\Delta T_{1/2}$ values represent the difference between the melting temperature of the DNA with and without ligands.

Nuclear Magnetic Resonance Experiments

A 700 MHz Varian Unity INOVA spectrometer was employed to perform the NMR experiments. One-dimensional proton spectra were recorded at 7°C using pulsed-field gradient DPGFSE for water suppression. All DNA samples were prepared at 0.2 mM strand concentration in 0.22 mL (H₂O/D₂O 9:1) buffer solution. DNA/ligand mixtures were obtained by adding aliquots of a stock solution of the six ligands in DMSO-d₆ directly to the DNA solution inside the NMR tube (Randazzo et al., 2002; Amato et al., 2017). NMR data were processed using the iNMR software (www.inmr.net).

FRET and FRET-Melting

A FP-8300 spectrofluorometer (Jasco) equipped with a Peltier temperature controller accessory (Jasco PCT-818) was employed to carry out FRET experiments. The dual-labeled oligonucleotides corresponding to the G4 forming sequence 5'-FAM-d(GGG[TTAGGG]₃)-TAMRA-3' (G4-F21T) (Amato et al., 2016; Salvati et al., 2017) and the iM forming sequence 5'-FAM-d(TAACCC)₄-TAMRA-3' (iM-F24T) were used. Such sequences are characterized by the presence of the donor fluorophore FAM (6-carboxyfluorescein) and the acceptor fluorophore TAMRA (6-carboxytetramethylrhodamine) that are covalently bound at 5'- and 3'-ends, respectively. Labeled oligonucleotides were purchased from Biomers (Germany). G4-F21T and iM-F24T were prepared at 1 μ M concentration in 10 mM potassium phosphate buffer and 10 mM sodium phosphate buffer, respectively. Samples were annealed in a hot water bath at 90°C for 2 min, and then cooled to room temperature overnight. FRET measurements were performed both in the absence and presence of 5 molar equivalents of compounds 1–6. The final concentration of G4-F21T and iM-F24T was 0.1 μ M. A sealed quartz cuvette with a path length of 1 cm was used. FRET spectra were acquired before (at 5 and 20°C for iM and G4, respectively) and after (at 90°C) melting assay. The dual-labeled oligonucleotides were excited at 492 nm, and emission spectra were recorded between 500 and 650 nm using 100 nm/s scan speed. Excitation and emission slit widths were both set at 5 nm. FRET-melting experiments were performed by setting the excitation wavelength at 492 nm and the detection wavelength at 522 nm. The emission intensity of FAM was then normalized between 0 and 1. Data analysis was carried out using Origin 7.0 software.

Fluorescent Intercalator Displacement (FID) Assay

For the FID experiments, oligonucleotides d[(TAACCC)₄] (hTeloC_{FID}) and d[(TTTGGG(TTAGGG)₃A)] (mutTel24_{FID}) were purchased from Eurogentec and then purified via HPLC. Solid DNA samples were initially dissolved as a 1 mM stock solution in MilliQ water. 10 mM stock solutions of the candidate ligands were prepared in DMSO. Further dilutions were carried out in buffer: 10 mM NaH₂PO₄ for hTeloC_{FID} and 10 mM KH₂PO₄ for mutTel24_{FID}. DNA samples were thermally annealed at 90 μM in the respective buffers in an Applied Biosystems Veriti 96 well thermal cycler by holding at 95°C for 5 min and cooling at a rate of 1°C/min to 20°C. FID experiments were performed on a BMG CLARIOstar plate reader using Corning 96-Well Solid Black Flat Bottom plates. A 10 mM stock solution of thiazole orange (TO) was prepared in DMSO and diluted to 2 μM in the appropriate buffer for either hTeloC_{FID} or mutTel24_{FID}. Ninety microliters of the 2 μM TO solution were added to each well and fluorescence emission at 450 nm measured with excitation at 430 nm; this was normalized to 0% representing background fluorescence. One microliter of DNA was then added, shaken using double orbital shaking at 700 rpm in the plate reader for 15 s, and allowed to equilibrate for 15 min. After equilibration, fluorescence emission was measured as before, and normalized to 100% representing maximal fluorescence enhancement from the TO probe binding to the DNA secondary structure. 0.9 μL aliquots of ligand were titrated into each well (in triplicate) and measured as before. Fluorescence measurements after ligand addition were normalized between the 0 and 100% levels determined per the respective well. Percentage TO displacement was calculated as the difference between the normalized 100% fluorescence level and the normalized fluorescence measured after each ligand addition. The concentration for each ligand at which 50% of the TO was displaced (DC₅₀) was calculated by using Origin data analysis software to plot percentage TO displacement against ligand concentration. These data were fitted with dose-response curves and the equations of the curves were solved for $y = 50$ to give the DC₅₀ values.

RESULTS

Most of the investigations reported in the literature dealing with the determination of iM structures in solution have been accomplished in sodium buffer and under acidic conditions, generally at pH values down to 4.3 (Gehring et al., 1993; Gallego et al., 1997; Malliavin et al., 2003). This is because the cytidine pK_a is about 4.2, and the use of low pH values guarantees to obtain stable hemi-protonated C-C⁺ pairs. However, these conditions are far away from physiological. This may also have consequences on the study of the interaction between these target molecules and potential ligands, which may be differently protonated with respect to their state under physiological pH. Therefore, in order to find experimental conditions as close to physiological as possible, the behavior of the telomeric iM-forming sequence (hTeloC) in 10 mM sodium buffer at different pH values was

investigated by 1D ¹H-NMR and CD spectroscopies (Figures S1, S2, Supplementary Material). In particular, the pH range 4.3–7.0 was explored. NMR and CD spectra clearly indicated that the hTeloC sequence turned out to be folded into an iM structure only at pH values lower than 5.7, while above this value the iM structure is in equilibrium with the unfolded species. Therefore, we decided to perform our studies at the two boundaries pH values, namely 4.3 and 5.7. Since the main aim of this investigation was to evaluate the interaction of some known G4 ligands with an iM-forming DNA, experiments on the human telomeric G4 (mutTel24) were performed in parallel for comparison.

Circular Dichroism Studies

The structures adopted by hTeloC and mutTel24 were first investigated by circular dichroism (CD) spectroscopy in the absence of ligands. At both pH 4.3 and 5.7, hTeloC showed almost superimposable CD spectra having a positive band at 288 nm and a negative one at around 260 nm (Figure 3A). These bands are characteristic of an iM folding topology (Guo et al., 2008). mutTel24 also displayed almost superimposable CD spectra at both pH values. These spectra were characterized by two positive bands at around 290 and 270 nm, and a negative one at around 240 nm (Figure 3B). These bands are perfectly superimposable to those observed for the same molecule at pH 7.0, thus indicating the presence of the expected hybrid [3+1] G4 structure (hybrid-1) as the major conformation under acidic conditions (Karsisiotis et al., 2011; Gray et al., 2014).

CD experiments were also performed to examine the potential of compounds 1–6 to alter the native folding topology of the two investigated DNA structures both at pH 4.3 and 5.7. DNA/ligand mixtures were obtained by adding 5 molar equivalents of compound to the folded G4 and iM structures so as to have an excess with respect to potential binding sites. In the case of mutTel24, regardless of the pH, Berberine (1), BRACO-19 (2), Phen-DC3 (4) and RHPS4 (6) induced a significant change in the CD spectrum of the G4 structure (Figures 4A,B and Figures S3, S4, Supplementary Material). In particular, the loss of the band at 270 nm followed by an intensity's increase of the band at 290 nm suggested a conformational change of the G4 topology from the hybrid to the antiparallel conformation (Masiero et al., 2010; Randazzo et al., 2013). Conversely, Mitoxantrone (3) and Pyridostatin (5) did not produce any measurable conformational change of the G4 structure, even if a decrease of the band at 290 nm is observed upon their addition.

As for the iM structure, at pH 4.3, Mitoxantrone (3), Phen-DC3 (4), and RHPS4 (6) induced a significant hypochromic shift of the positive band at 288 nm, with more marked effects observed for 3 and 4. Conversely, Berberine (1), BRACO-19 (2), and Pyridostatin (5) caused a hyperchromic shift of the band at 288 nm at this pH (Figure 4C and Figure S5, Supplementary Material). On the other hand, at pH 5.7 all the compounds have been shown to induce a hypochromic effect of the band at 288 nm that turned out to be particularly marked in the case of 2, 3, and 4 (Figure 4D and Figure S6, Supplementary Material). These data suggest that some interaction takes place and that, in some cases, the molecules seem to induce the unfolding of the structure.

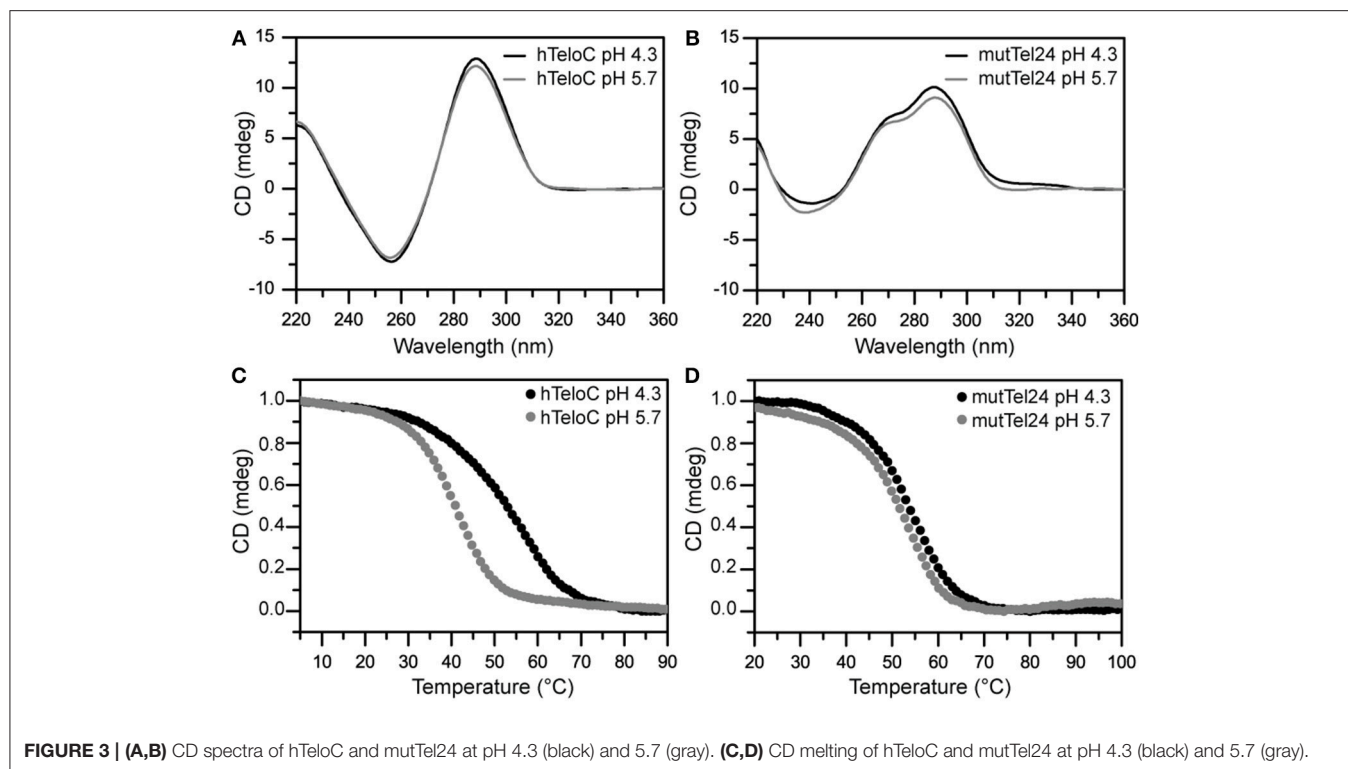


FIGURE 3 | (A,B) CD spectra of hTeloC and mutTel24 at pH 4.3 (black) and 5.7 (gray). **(C,D)** CD melting of hTeloC and mutTel24 at pH 4.3 (black) and 5.7 (gray).

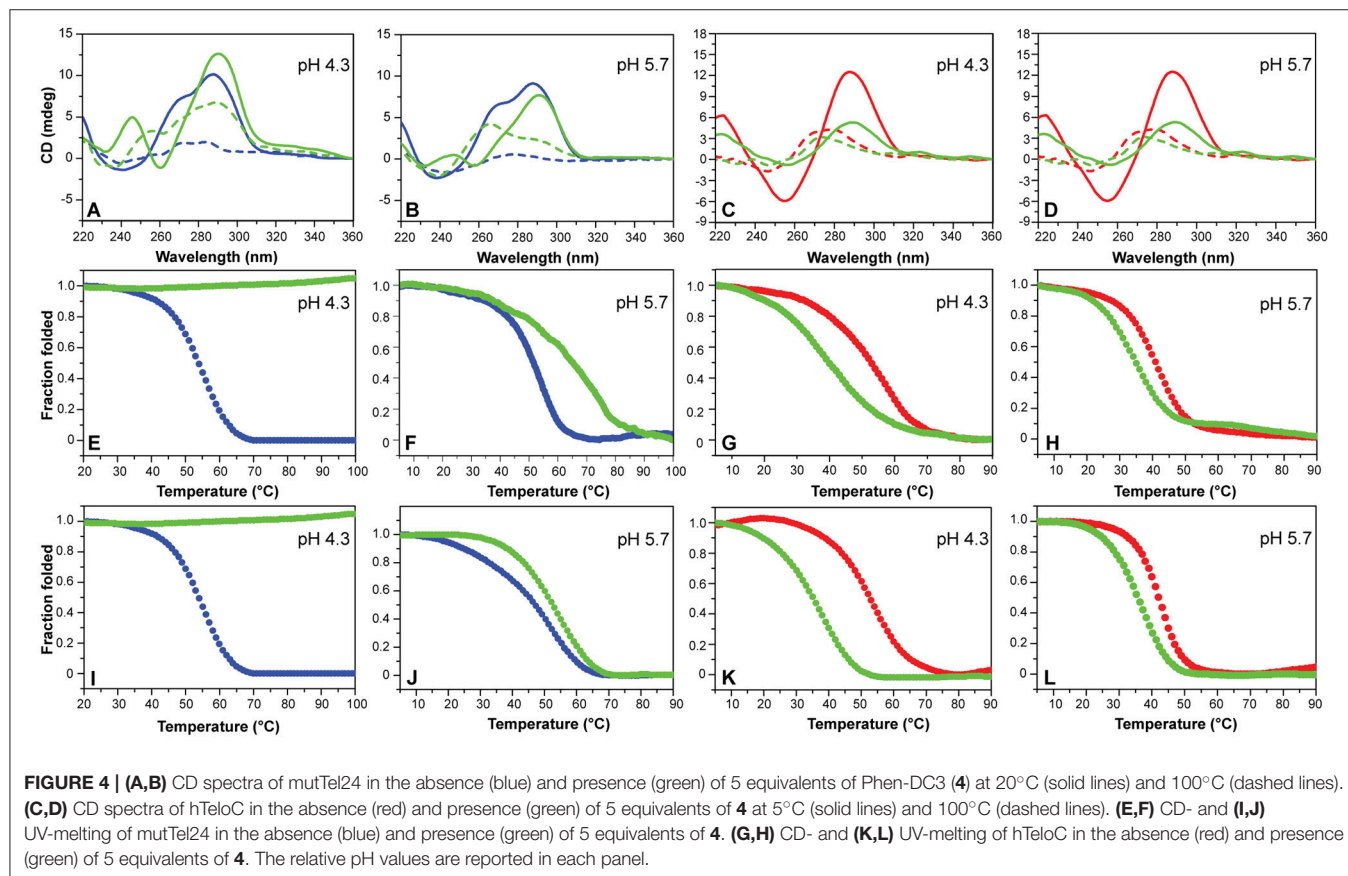


FIGURE 4 | (A,B) CD spectra of mutTel24 in the absence (blue) and presence (green) of 5 equivalents of Phen-DC3 (**4**) at 20°C (solid lines) and 100°C (dashed lines). **(C,D)** CD spectra of hTeloC in the absence (red) and presence (green) of 5 equivalents of **4** at 5°C (solid lines) and 100°C (dashed lines). **(E,F)** CD- and **(I,J)** UV-melting of mutTel24 in the absence (blue) and presence (green) of 5 equivalents of **4**. **(G,H)** CD- and **(K,L)** UV-melting of hTeloC in the absence (red) and presence (green) of 5 equivalents of **4**. The relative pH values are reported in each panel.

CD-Melting Analysis

CD-melting experiments were employed to evaluate the thermal stability of iM and G4 structures adopted by hTeloC and mutTel24 sequences, respectively, under the two experimental conditions used (pH 4.3 and 5.7). The melting temperatures ($T_{1/2}$) of the iM structure were found to be 52.8 and 40.6°C at pH 4.3 and 5.7, respectively (Figure 3C). The lower thermal stability of the iM structure observed at pH 5.7 can be ascribed to the lower extent of protonation of the cytosines. On the other hand, only a very small variation in $T_{1/2}$ values was observed for the G4 structure between pH 4.3 and 5.7 (53.5 and 51.1°C, respectively, Figure 3D).

Then, the effect of the investigated compounds on the stability of the DNA secondary structures was evaluated by measuring the ligand-induced change in the melting temperature ($\Delta T_{1/2}$) of G4 and iM at pH 4.3 and 5.7. CD-melting curves of DNA in the absence and presence of each compound were obtained by following the variations of the intensity of CD signal at the wavelengths of 290 and 288 nm for mutTel24 and hTeloC, respectively (Figure 4 and Figures S7–S10, Supplementary Material). Very intriguingly, results of these experiments (Table 1) clearly indicated that 1–6 exert a different effect on iM compared to G4. As expected, all the tested compounds were able to thermally stabilize the G4 structure at both pH values, even if to a different extent (Figures 4E,F, and Figures S7, S8, Supplementary Material). On the contrary, regardless of pH, Berberine (1), Pyridostatin (5), and RHPS4 (6) did not show a remarkable influence on the iM thermal stability, while BRACO-19 (2), Mitoxantrone (3), and Phen-DC3 (4) significantly decreased it (Figures 4G,H and Figures S9, S10, Supplementary Material). These results are in agreement with those obtained from the CD spectra analysis.

UV-Melting Analysis

The effect of compounds 1–6 on the stability of the G4 and iM structures was also investigated by UV-melting experiments. As for CD-melting studies, the ligand-induced changes in the melting temperature ($\Delta T_{1/2}$) of the two DNA structures were obtained by recording UV-melting experiments in the absence

and presence of each compound at both pH 4.3 and 5.7. UV-melting curves were acquired by following the change in UV signal intensity at 295 nm for both mutTel24 and hTeloC (Figures 4I–L, and Figures S11–S14, Supplementary Material). Results of these experiments (Table 1) are consistent with CD-melting ones, and denote, once again, a different behavior for the investigated compounds toward the iM and G4 DNA structures.

Nuclear Magnetic Resonance Studies

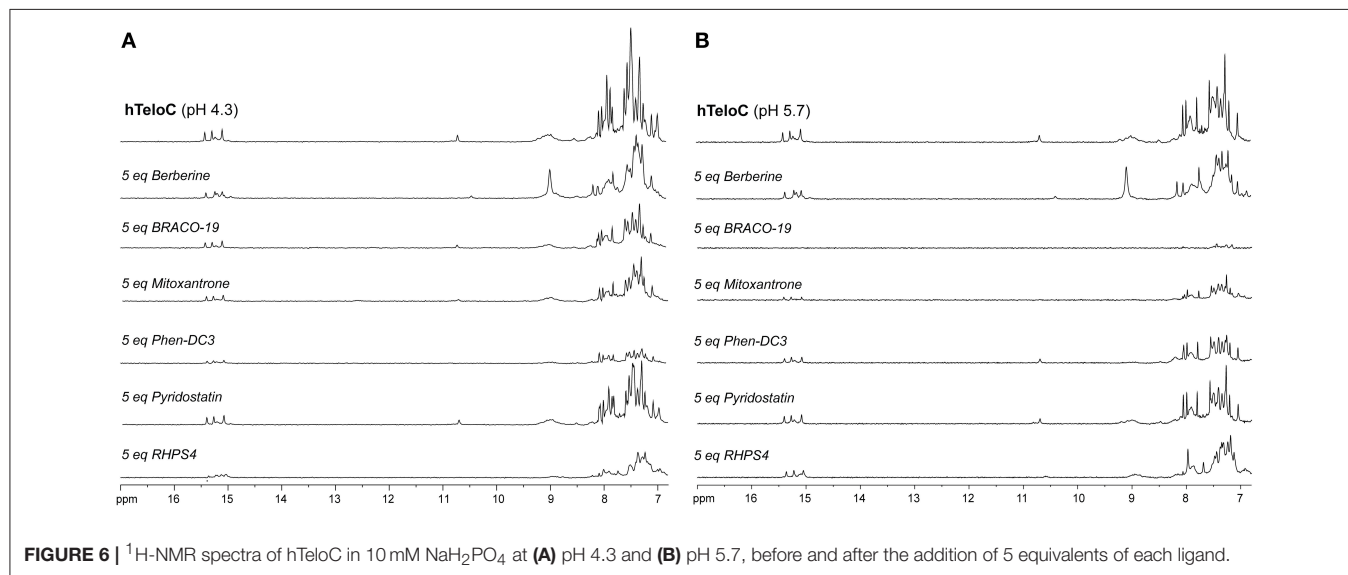
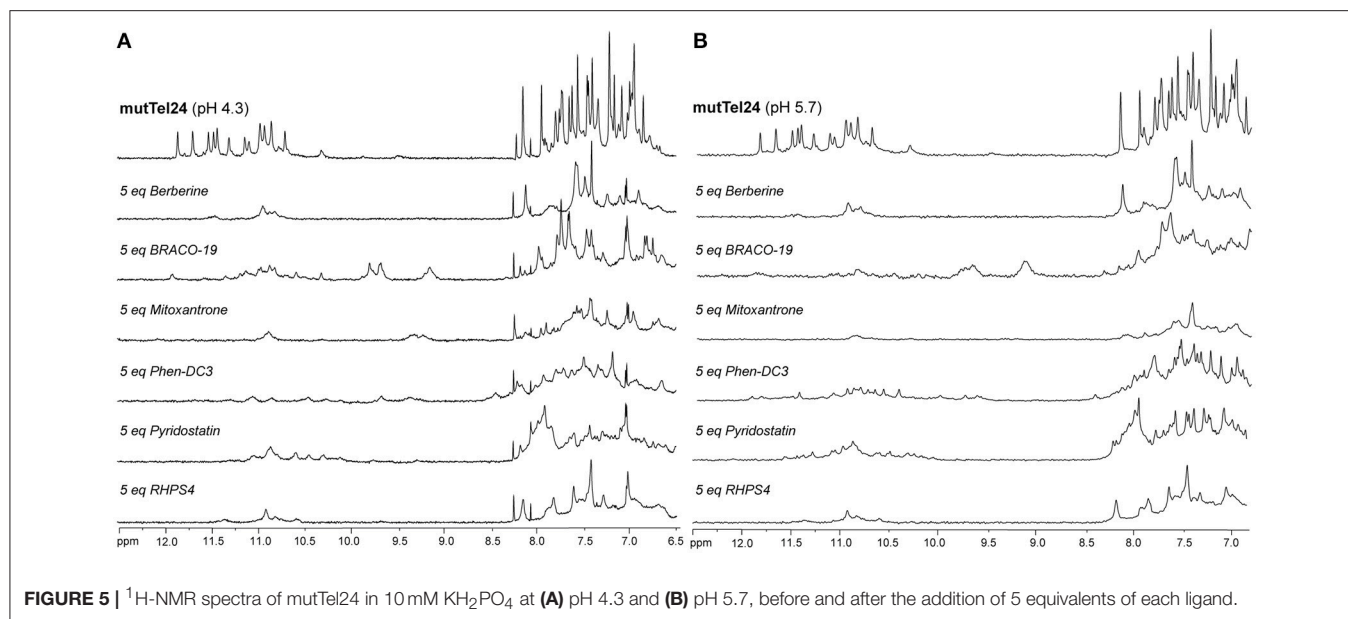
NMR spectroscopy was employed in order to obtain structural information about the DNA interaction of the six compounds. Also in this case, 1D ^1H -NMR spectra were recorded at pH 4.3 and 5.7. Under the experimental conditions used, the mutTel24 sequence forms a single G4 conformation characterized by 12 well-resolved imino proton peaks, corresponding to the 12 guanines involved in the three G-tetrad planes (Luu et al., 2006). On the other hand, hTeloC folds in an i-motif structure characterized by 3 well-resolved imino proton peaks that correspond to the 6 intercalated C-C⁺ pairs (Phan et al., 2000).

The imino and aromatic proton regions of mutTel24 and hTeloC in the absence and presence of 5 equivalents of each compound are shown in Figures 5, 6, respectively. Regardless of the pH, both imino and aromatic proton signals of mutTel24 turned out to be significantly affected by the addition of the ligands (Figure 5). On the other hand, addition of compounds to the iM structure led to different results. At pH 4.3, the main changes were observed for Berberine (1) and RHPS4 (6) in both aromatic and imino regions, while little changes could be observed for BRACO-19 (2) and Mitoxantrone (3). Instead, Phen-DC3 (4) caused a general decrease of signal intensities (Figure 6A). Very little changes could be observed in the spectrum of iM upon addition of Pyridostatin (5). At pH 5.7, changes were observed for Berberine (1) BRACO-19 (2), Mitoxantrone (3), and RHPS4 (6) (Figure 6B), with BRACO-19 affecting the most the NMR spectrum of the iM structure. Interestingly, as for the experiment at pH 4.3, PhenDC3 (4) and Pyridostatin (5) caused a general decrease of the signal intensities.

TABLE 1 | Ligand-induced thermal stabilization of hTeloC and mutTel24 DNA measured by CD and UV melting experiments.

	$\Delta T_{1/2}$ (°C)* CD melting				$\Delta T_{1/2}$ (°C)* UV melting			
	hTeloC		mutTel24		hTeloC		mutTel24	
	pH 4.3	pH 5.7	pH 4.3	pH 5.7	pH 4.3	pH 5.7	pH 4.3	pH 5.7
Berberine	−2.5	−0.8	+13.4	+12.4	+1.2	+0.9	+11.7	+15.3
BRACO-19	−13.4	−9.2	+12.4	+8.9	−17.6	−6.5	ND	ND
Mitoxantrone	−4.8	−9.9	+7.7	+4.5	−16.6	−1.0	+4.1	+0.6
Phen-DC3	−13.4	−6.8	ND	+14.4	−17.1	−6.3	ND	+5.3
Pyridostatin	−2.8	+0.8	+12.9	+8.8	+1.1	+0.2	+9.3	+5.5
RHPS4	−1.0	−0.3	+22.0	+20.6	+0.7	+0.1	ND	ND

* $\Delta T_{1/2} = T_{1/2}(\text{DNA} + \text{ligand}) - T_{1/2}(\text{DNA})$. All experiments were performed in duplicate, and $\Delta T_{1/2}$ values are reported as the mean. Errors were $\pm 0.5^\circ\text{C}$. ND, not determined.



FRET and FRET-Melting Studies

Dual labeled (FAM/TAMRA) human telomeric sequences G4-F21T and iM-F24T, which are able to form G4 and iM structures, respectively, were used (see Experimental section). Experiments were performed only at pH 5.7, because the fluorescence of the FAM is not stable at pH 4.3. In order to verify that **1-6** did not interfere in the FAM emission spectrum, the fluorescence spectrum of each compound was recorded by exciting at 492 nm and collecting its emission spectrum between 500 and 650 nm. Unfortunately, the emission spectrum of RHPS4 (**6**) was found to overlap with FAM, and, therefore, it was not used in these experiments.

FRET-melting assays were then performed to further investigate the compound-induced effects on the iM and G4

thermal stabilities. In agreement with CD- and UV-melting experiments, all the tested compounds induced a thermal stabilization of G4-F21T (Figure S15, Supplementary Material). Conversely, in the case of iM-F24T, FRET melting data did not agree at all with CD and UV melting results. In fact, all ligands, except Berberine, showed a significant thermal stabilization of the iM structure (Figure S16, Supplementary Material).

FRET spectra of FAM/TAMRA-modified oligonucleotides (G4-F21T and iM-F24T) in the absence and presence of each compound (at 1:5 DNA/ligand ratio) were also analyzed (Figure 7). The results show that, in both cases, BRACO-19 (**2**), Mitoxantrone (**3**), Phen-DC3 (**4**), and Pyridostatin (**5**) caused a significant decrease of the band intensity at 580 nm, thus suggesting that actually they could interact with the probes.

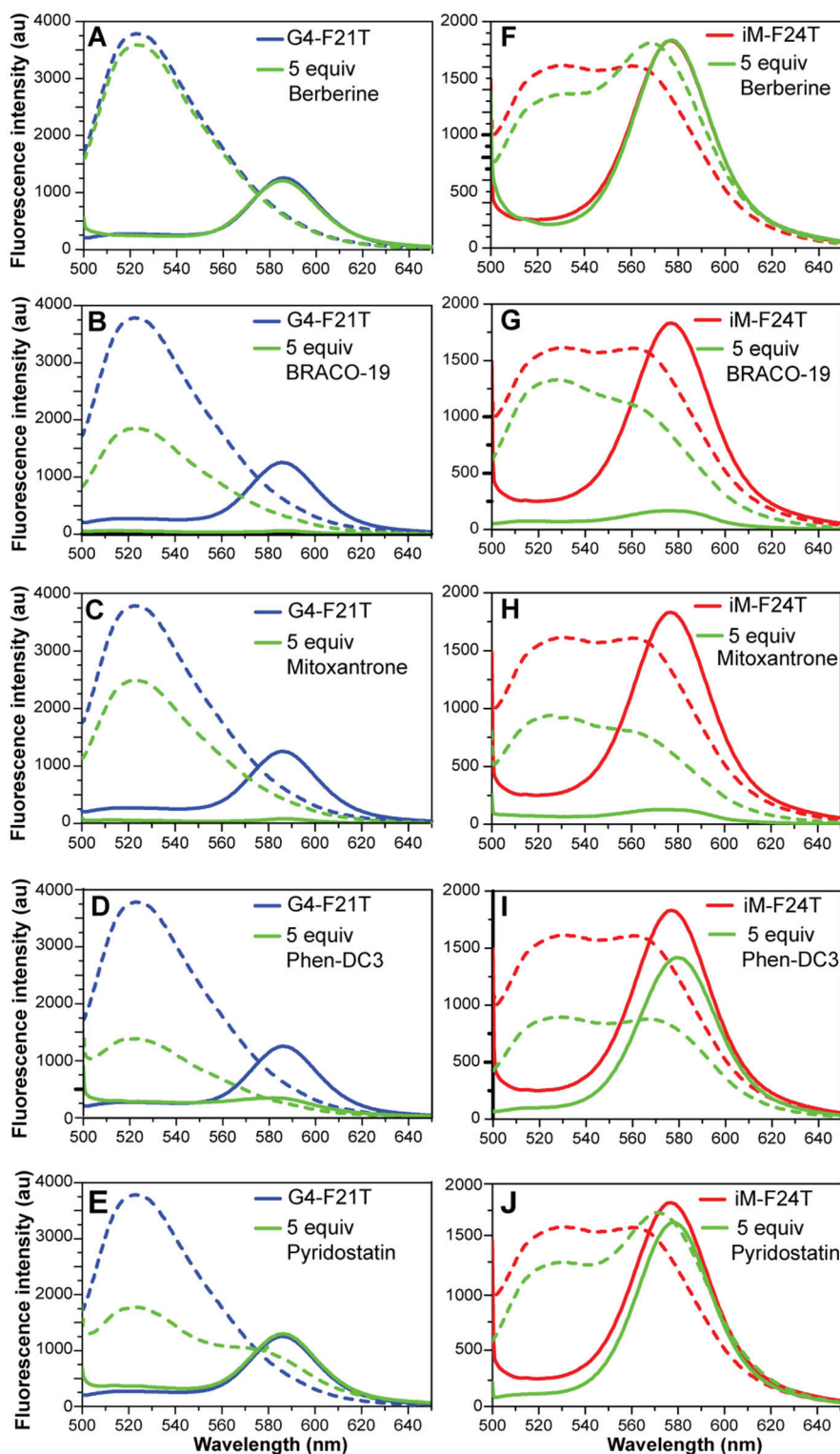


FIGURE 7 | (A–E) Fluorescence spectra of G4-F21T alone (in 10 mM KH_2PO_4 buffer) at 5°C (blue line) and at 90°C (dashed blue line), and after the addition of 5 equivalents of each ligand at 5°C (green line) and at 90°C (dashed green line). **(F–J)** Fluorescence spectra of iM-F24T alone (in 10 mM NaH_2PO_4 buffer) at 5°C (red line) and at 90°C (red dashed line), and after the addition of 5 equivalents of each ligand at 5°C (green line) and at 90°C (dashed green line). All the experiments were performed at pH 5.7.

Fluorescent Intercalator Displacement (FID) Assay

FID assays are long established and have been used extensively to determine relative binding affinities for ligands with duplex DNA (Boger et al., 2000). They have recently been adapted and validated for use with both G4 and iM structures (Monchaud et al., 2006; Sheng et al., 2017). The assay relies on a light-up fluorescent probe, in this case thiazole orange (TO), which binds to the structure of interest and can be competitively displaced by candidate ligands thus enabling the determination of their relative binding affinity to the structure under examination. Here, aliquots of ligands **1-5** were titrated in triplicate against both of the G4 and iM structures and the concentrations at which 50% displacement (DC_{50}) was achieved were calculated from dose-response curves fitted to this data (Figures S17, S18, Supplementary Material). Unfortunately, as with the FRET, RHPS4 (**6**) was excluded from analysis due to the fluorescence profile of the ligand overlapping with the assay parameters. The results of the FID assay showed that all ligands (**1-5**) bound to both the G4 and the iM DNA. Unsurprisingly, they also all showed a slightly higher affinity for the G4 formed by mutTel24_{FID} over the iM formed by hTeloC_{FID} (Table 2). Nevertheless, all ligands tested were found to displace TO at both pH 4.3 and 5.7, indicative of an interaction with the iM-forming DNA.

DISCUSSION

DNA has a well-known propensity to adopt various alternative non-canonical conformations *in vitro*, including G4 and iM. Recent investigations have demonstrated the formation of such structures in regulatory regions of the human genome, including gene promoters and telomeres, also providing evidences for the key role that G4s and iMs can play in several biological pathways (Kang et al., 2014; Salvati et al., 2014; Maizels, 2015).

Human telomeric DNA consists of a 2–20 kb double-stranded region composed by (TTAGGG)/(CCCTAA) repeats, and of a single-stranded 3'-end G-rich sequence. The G-rich strand can adopt G4 conformations, while the opposite C-rich strand can fold into the iM structure. Of the two structures, telomeric G4 is by far the most studied. This disparity is mainly due to acidic pH required for the protonation of cytosine, since the parallel duplexes, the basic component of iM, are stabilized by hemiprotonated C-C⁺ base pairs. In this investigation, the behavior of the human telomeric iM-forming sequence (hTeloC) was studied in the pH range of 4.3–7.0 by 1D ¹H-NMR and CD spectroscopies. Results clearly indicate that under the experimental conditions used, the iM structure is still well-preserved at pH 5.7, while, just above this pH value, it turns out to be in equilibrium with the random coil (Figures S1, S2, Supplementary Material). Moreover, the lower stability of the iM structure observed at pH 5.7 well reflects the lower extent of cytosine protonation compared to pH 4.3, thus confirming, once again, that iM structures are very sensitive to pH (Kovanda et al., 2015). On the other hand, results of NMR and CD experiments on the G-rich telomeric sequence clearly showed that it retains

the hybrid [3+1] G4 structure as major conformation at both pH 4.3 and 5.7, and that its thermal stability is basically not affected by the pH.

Once the working conditions were established, the interaction of the well-known G4 ligands Berberine (**1**), BRACO-19 (**2**), Mitoxantrone (**3**), Phen-DC3 (**4**), Pyridostatin (**5**), and RHPS4 (**6**), with the iM-forming sequence was explored in comparison with the G4, by using a combination of spectroscopic techniques. Experiments were performed at the two boundary pH values. Results of the different experiments undoubtedly showed that all the investigated compounds actually are able to interact, even if in a different way, with both G4- and iM-forming DNA.

Fluorescent intercalator displacement (FID) assay clearly showed that the thiazole orange probe, which binds to the investigated structures, is competitively displaced by the compounds **1-5** (compound **6** could not be used). The consequent determination of their relative affinity for the DNA under examination reveals that, regardless of the pH (4.3 or 5.7), the compounds exhibit only a slightly higher affinity for mutTel24 over hTeloC.

Ligand-induced effects on both G4 and iM structures were examined by means of UV, CD, and NMR. As far as mutTel24 is concerned, results of CD experiments show that four (Berberine, BRACO-19, Phen-DC3, and RHPS4) out of the six ligands induce a conformational change from the hybrid [3+1] to an antiparallel structure (Figures S3, S4). All the ligands also indiscriminately affect the NMR spectrum of mutTel24 (Figure 5). Results of CD- and UV-melting experiments agree in indicating that the six compounds are able to stabilize the G4 at both pH values (Figures S7, S8, S11, S12, Supplementary Material). Interestingly, the two ligands not inducing the G4 conformational change (Mitoxantrone and Pyridostatin) seem also to be the less effective in terms of thermal stabilization under the experimental conditions used. Moreover, some of the investigated ligands induced a higher increase of G4 melting temperature ($\Delta T_{1/2}$) at pH 4.3 rather than 5.7 (Table 1), that may be ascribed in part to the different protonation states of such molecules at the different pHs. This was, for example, the case of BRACO-19 (**2**) and Pyridostatin (**5**). For these molecules the protonation state was theoretically determined in the 3.5–7.5 pH range, by using the pK_a prediction program (ChemAxon, www.chemaxon.com) based on the calculation of partial charge of atoms in the molecule. The computed distribution of the microspecies for BRACO-19 and Pyridostatin, reported in Figure S19, shows that actually there is a variation between pH 4.3 and 5.7 for these molecules.

Concerning the iM, some differences were observed between pH 4.3 and 5.7. In particular, the addition of **1-6** to the iM caused major effects at pH 5.7, in which the iM structure is less stable *per se*.

Interestingly, CD- and UV-melting results clearly indicate that the molecules which were found to mainly decrease the CD and NMR signals (namely BRACO-19, Mitoxantrone, and Phen-DC3) were also those that significantly decreased iM thermal stability at both pH values (Figures S9, S10, S13, S14, Supplementary Material).

TABLE 2 | Ligand DC₅₀ values for hTeloC_{FID} and mutTel24_{FID} determined using the FID assay*.

	hTeloC _{FID}				mutTel24 _{FID}			
	pH 4.3		pH 5.7		pH 4.3		pH 5.7	
	DC ₅₀ (μ M)	SE (μ M)	DC ₅₀ (μ M)	SE (μ M)	DC ₅₀ (μ M)	SE (μ M)	DC ₅₀ (μ M)	SE (μ M)
Berberine	30.38	1.58E-02	1.46	5.88E-03	3.32	2.63E-02	1.26	7.22E-03
BRACO-19	0.66	1.61E-03	0.87	8.08E-04	0.26	5.14E-03	0.50	1.95E-03
Mitoxantrone	0.70	2.87E-03	1.34	6.99E-03	0.54	3.03E-04	0.95	6.07E-03
Phen-DC3	0.97	1.32E-02	0.95	3.79E-03	0.26	1.58E-03	0.39	1.42E-03
Pyridostatin	9.09	6.18E-03	18.02	5.89E-02	3.15	2.33E-03	9.42	3.90E-02

* All experiments were performed in triplicate and DC₅₀ values are reported as the 50% displacement value calculated from fitted dose response curves. Standard errors are calculated using R-square values from the statistics on the data fit.

Overall, these results show that compounds **1-6** are able to interact with the telomeric iM-forming DNA. However, three of them (Berberine, Pyridostatin, and RHPS4) do not have relevant effects on the thermal stability of iM, while the others (BRACO-19, Mitoxantrone, and Phen-DC3) are able to destabilize it. We speculate that these compounds could reasonably make some non-specific interactions with the single-stranded (unfolded) C-rich DNA, resulting in a shift of folded-unfolded equilibrium toward the unfolded form, especially during the melting experiment, which in turn results in a decrease of $T_{1/2}$.

The FRET methodology, used here to further characterize the interaction of G4 and iM DNA with the investigated ligands, deserves a separate discussion. With respect of other spectroscopic techniques (such as UV, CD, and NMR), FRET has a higher sensitivity and it can explore a large range of ligand concentrations (Monchaud et al., 2006; Sheng et al., 2017). Additionally, it turns out to be the main methodology when the UV absorbance of a ligand overlaps with that of the DNA (Guédin et al., 2010). However, some artifacts may occur when compounds are inherently fluorescent and/or interact with the fluorescent probes rather than the DNA itself (De Cian et al., 2007b), and this is what probably happened in this case. Indeed, FRET melting results did not agree at all with both CD and UV melting data, especially in the case of iM-F24T (Figure S16, Supplementary Material). To understand the reasons for this different behavior, FRET spectra of labeled DNA (G4-F21T and iM-F24T) in the absence and presence of compounds were analyzed. Typically, when DNA is folded, the two dyes are in close proximity so FAM fluorescence peak at 522 nm (upon excitation at 492 nm) is quenched and its energy is transferred to TAMRA, which then emits light at 580 nm. On the other hand, FAM's fluorescence is no longer quenched when sufficient spatial separation of the two dyes occurs (for example upon unfolding of the DNA structure), therefore its fluorescence signal at 522 nm is observable. In principle, compounds that are able to interact with the fluorophores may affect the emission properties of the probes and decrease the intensity of the bands at 580 nm (if the DNA is structured) or at 522 nm (if the DNA is unstructured). Interestingly, four compounds (BRACO-19,

Mitoxantrone, Phen-DC3, and Pyridostatin) caused a significant decrease of the band intensity at 580 nm (for both G4-F21T and iM-F24T), clearly suggesting that they actually interact with the fluorophores. Therefore, the different stabilizing effects observed for the iM structure across the different spectroscopic techniques could be ascribed to ligand interaction with the FRET probes. This hypothesis is further corroborated by the fact that Berberine (**1**), which did not cause any observable change in the FRET spectrum, showed no variation in the DNA's thermal stability by FRET, in agreement with CD and UV experiments. For the same reasons, this suggests that the thermal stabilizations of G4-F21T measured by FRET are not accurate, being also potentially affected by the ligands' interaction with the probes. Therefore, a clear message came out from a careful examination of FRET data, meaning that false-positive responses can be obtained due to ligands ability to bind end-labeling DNA probes. This could occur particularly when the investigated compounds have an extended aromatic core for which π - π stacking interactions with the large aromatic surface of the probes could be favored. This is especially the case of G4-interacting molecules. Overall, this study emphasizes the need of using a combination of techniques when examining DNA targeting ligands, in order to avoid an inaccurate evaluation of their binding/stabilizing properties.

CONCLUSIONS

Herein, a combination of spectroscopic techniques was employed to determine whether well-known bioactive G4 ligands, namely Berberine (**1**), BRACO-19 (**2**), Mitoxantrone (**3**), Phen-DC3 (**4**), Pyridostatin (**5**), and RHPS4 (**6**) are able to interact with an iM-forming DNA. Two human telomeric DNA sequences able to form iM and G4 structures were studied and the experiments performed at two different pH values. The experimental results showed that all the investigated G4 ligands were also able to interact with the telomeric iM-forming DNA. Very interestingly, BRACO-19, Mitoxantrone, and Phen-DC3 have been shown to destabilize the iM structure.

The majority of iM forming sequences are generally less stable than G4s under physiological conditions. The

delicate equilibrium between the folded and unfolded DNA forms is highly sensitive to the environmental conditions (such as pH and ionic strength), and ligands may affect this fine equilibrium, shifting it toward different/less stable forms.

The here reported results are even more interesting if viewed in the context of regulation of gene expression. Indeed, recent investigations have suggested that G4 and iM structures may have opposing functions in the control of oncogene transcription: while G4 formation and its ligand-induced stabilization generally inhibits gene expression, stabilization of iM seems to have transcription activating capabilities (Kang et al., 2014; Kendrick et al., 2014). Therefore, a molecule that is able to stabilize a G4 structure and to destabilize an iM structure may exert a synergistic effect on the inhibition of transcription. These are the cases of BRACO-19 (2), Mitoxantrone (3), and Phen-DC3 (4), whose biological activity may be ascribed to both mechanisms.

Overall, the present study highlights the necessity, when studying G4-targeting compounds, of evaluating also their effects on the i-motif counterparts, especially if one is looking for a “specific” drug.

REFERENCES

- Amato, J., Iaccarino, N., Pagano, B., Morigi, R., Locatelli, A., Leoni, A., et al. (2014a). Bis-indole derivatives with antitumor activity turn out to be specific ligands of human telomeric G-quadruplex. *Front. Chem.* 2:54. doi: 10.3389/fchem.2014.00054
- Amato, J., Iaccarino, N., Randazzo, A., Novellino, E., and Pagano, B. (2014b). Noncanonical DNA secondary structures as drug targets: the prospect of the i-motif. *ChemMedChem* 9, 2026–2030. doi: 10.1002/cmdc.201402153
- Amato, J., Morigi, R., Pagano, B., Pagano, A., Ohnmacht, S., De Magis, A., et al. (2016). Toward the development of specific g-quadruplex binders: synthesis, biophysical, and biological studies of new hydrazone derivatives. *J. Med. Chem.* 59, 5706–5720. doi: 10.1021/acs.jmedchem.6b00129
- Amato, J., Pagano, A., Capasso, D., Di Gaetano, S., Giustiniano, M., Novellino, E., et al. (2018). Targeting the BCL2 gene promoter G-Quadruplex with a new class of furopyridazinone-based molecules. *ChemMedChem* 13, 406–410. doi: 10.1002/cmdc.201700749
- Amato, J., Pagano, A., Cosconati, S., Amendola, G., Fotticchia, I., Iaccarino, N., et al. (2017). Discovery of the first dual G-triplex/G-quadruplex stabilizing compound: a new opportunity in the targeting of G-rich DNA structures? *Biochim. Biophys. Acta* 1861, 1271–1280. doi: 10.1016/j.bbagen.2016.11.008
- Bedrat, A., Lacroix, L., and Mergny, J. L. (2016). Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res.* 44, 1746–1759. doi: 10.1093/nar/gkw006
- Bhavsar-Jog, Y. P., Van Dornshuld, E., Brooks, T. A., Tschumper, G. S., and Wadkins, R. M. (2014). Epigenetic modification, dehydration, and molecular crowding effects on the thermodynamics of i-motif structure formation from C-Rich, DNA. *Biochemistry* 53, 1586–1594. doi: 10.1021/bi401523b
- Biffi, G., Tannahill, D., McCafferty, J., and Balasubramanian, S. (2013). Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* 5, 182–186. doi: 10.1038/nchem.1548
- Boger, D. L., Fink, B. E., and Hedrick, M. P. (2000). Total synthesis of distamycin, A., and 2640 analogues: a solution-phase combinatorial approach to the discovery of new, bioactive dna binding agents and development of a rapid, high-throughput screen for determining relative DNA binding affinity or DNA bind. *J. Am. Chem. Soc.* 122, 6382–6394. doi: 10.1021/ja994192d

AUTHOR CONTRIBUTIONS

AR, JA, BP, and ZW conceived and designed the experiments. AP, NI, EG, and AD performed the CD, UV, and FRET experiments. DB performed the NMR experiments. MA performed the FID experiments. AP and JA carried out the synthesis of oligonucleotides. AP, NI, MA, EN, ZW, BP, JA, and AR analyzed the results. AR, JA, and ZW wrote the paper. All authors verified the data, contributed to the manuscript, and approved the final version.

ACKNOWLEDGMENTS

This work was supported by the Italian Association for Cancer Research (AIRC) (IG-18695 to AR, IG-16730 to BP). MA was supported by a Ph.D. studentship from the Eastern Academic Research Consortium.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fchem.2018.00281/full#supplementary-material>

- Brazier, J. A., Shah, A., and Brown, G. D. (2012). I-Motif formation in gene promoters: unusually stable formation in sequences complementary to known G-quadruplexes. *Chem. Commun.* 48:10739. doi: 10.1039/c2cc30863k
- Brown, R. V., Wang, T., Chappeta, V. R., Wu, G., Onel, B., Chawla, R., et al. (2017). The consequences of overlapping G-Quadruplexes and i-Motifs in the platelet-derived growth factor receptor β core promoter nucleosome hypersensitive element can explain the unexpected effects of mutations and provide opportunities for selective targeting of both structures by small molecules to downregulate gene expression. *J. Am. Chem. Soc.* 139, 7456–7475. doi: 10.1021/jacs.6b10028
- Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K., and Neidle, S. (2006). Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.* 34, 5402–5415. doi: 10.1093/nar/gkl655
- Cantor, C. R., Warshaw, M. M., and Shapiro, H. (1970). Oligonucleotide interactions. III. Circular dichroism studies of the conformation of deoxyoligonucleotides. *Biopolymers* 9, 1059–1077. doi: 10.1002/bip.1970.360090909
- Cerofolini, L., Amato, J., Giachetti, A., Limongelli, V., Novellino, E., Parrinello, M., et al. (2014). G-triplex structure and formation propensity. *Nucleic Acids Res.* 42, 13393–13404. doi: 10.1093/nar/gku1084
- Chambers, V. S., Marsico, G., Boutell, J. M., Di Antonio, M., Smith, G. P., and Balasubramanian, S. (2015). High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.* 33, 877–881. doi: 10.1038/nbt.3295
- Chen, Y., Qu, K., Zhao, C., Wu, L., Ren, J., Wang, J., et al. (2012). Insights into the biomedical effects of carboxylated single-wall carbon nanotubes on telomerase and telomeres. *Nat. Commun.* 3:1074. doi: 10.1038/ncomms2091
- Day, H. A., Pavlou, P., and Waller, Z. A. E. (2014). I-Motif DNA: structure, stability and targeting with ligands. *Bioorg. Med. Chem.* 22, 4407–4418. doi: 10.1016/j.bmc.2014.05.047
- De Cian, A., DeLemos, E., Mergny, J. L., Teulade-Fichou, M. P., and Monchaud, D. (2007a). Highly efficient G-quadruplex recognition by bisquinolinium compounds. *J. Am. Chem. Soc.* 129, 1856–1857. doi: 10.1021/ja067352b
- De Cian, A., Guittat, L., Kaiser, M., Saccà, B., Amrane, S., Bourdoncle, A., et al. (2007b). Fluorescence-based melting assays for studying quadruplex ligands. *Methods* 42, 183–195. doi: 10.1016/j.ymeth.2006.10.004

- Di Leva, F. S., Zizza, P., Cingolani, C., D'Angelo, C., Pagano, B., Amato, J., et al. (2013). Exploring the chemical space of G-quadruplex binders: discovery of a novel chemotype targeting the human telomeric sequence. *J. Med. Chem.* 56, 9646–9654. doi: 10.1021/jm401185b
- Dzatkó, S., Krafčíková, M., Hänsel-Hertsch, R., Fessl, T., Fiala, R., Loja, T., et al. (2018). Evaluation of the stability of DNA i-Motifs in the Nuclei of living mammalian cells. *Angew. Chem. Int. Ed. Engl.* 57, 2165–2169. doi: 10.1002/anie.201712284
- Fedoroff, O. Y., Rangan, A., Chemeris, V. V., and Hurley, L. H. (2000). Cationic porphyrins promote the formation of i-motif DNA and bind peripherally by a nonintercalative mechanism. *Biochemistry* 39, 15083–15090. doi: 10.1021/bi001528j
- Fleming, A. M., Ding, Y., Rogers, R. A., Zhu, J., Zhu, J., Burton, A. D., et al. (2017). 4n-1 Is a “Sweet Spot” in DNA i-Motif folding of 2'-Deoxycytidine homopolymers. *J. Am. Chem. Soc.* 139, 4682–4689. doi: 10.1021/jacs.6b10117
- Franceschin, M., Rossetti, L., D'Ambrosio, A., Schirripa, S., Bianco, A., Ortaggi, G., et al. (2006). Natural and synthetic G-quadruplex interactive berberine derivatives. *Bioorg. Med. Chem. Lett.* 16, 1707–1711. doi: 10.1016/j.bmcl.2005.12.001
- Gallego, J., Chou, S. H., and Reid, B. R. (1997). Centromeric pyrimidine strands fold into an intercalated motif by forming a double hairpin with a novel T:G:C:T tetrad: solution structure of the d(TCCCGTTTCCA) dimer. *J. Mol. Biol.* 273, 840–856. doi: 10.1006/jmbi.1997.1361
- Gehring, K., Leroy, J.-L., and Guéron, M. (1993). A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature* 363, 561–565. doi: 10.1038/363561a0
- Gowan, S. M., Harrison, J. R., Patterson, L., Valenti, M., Read, M., a, Neidle, S., et al. (2002). A G-quadruplex-interactive potent small-molecule inhibitor of telomerase exhibiting *in vitro* and *in vivo* antitumor activity. *Mol. Pharmacol.* 61, 1154–1162. doi: 10.1124/mol.61.5.1154
- Gray, R. D., Trent, J. O., and Chaires, J. B. (2014). Folding and unfolding pathways of the human telomeric G-quadruplex. *J. Mol. Biol.* 426, 1629–1650. doi: 10.1016/j.jmb.2014.01.009
- Guédin, A., Lacroix, L., and Mergny, J.-L. (2010). Thermal melting studies of ligand DNA interactions. *Methods Mol. Biol.* 613, 25–35. doi: 10.1007/978-1-60327-418-0_2
- Guo, K., Gokhale, V., Hurley, L. H., and Sun, D. (2008). Intramolecularly folded G-quadruplex and i-motif structures in the proximal promoter of the vascular endothelial growth factor gene. *Nucleic Acids Res.* 36, 4598–4608. doi: 10.1093/nar/gkn380
- Hausler, A. R., Donnelly, C. J., Periz, G., Simko, E. A. J., Shaw, P. G., Kim, M. S., et al. (2014). C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* 507, 195–200. doi: 10.1038/nature13124
- Huang, H.-S., Chen, I.-B., Huang, K.-F., Lu, W.-C., Shieh, F.-Y., Huang, Y.-Y., et al. (2007). Synthesis and human telomerase inhibition of a series of regioisomeric disubstituted amidoanthraquinones. *Chem. Pharm. Bull.* 55, 284–292. doi: 10.1248/cpb.55.284
- Huppert, J. L., and Balasubramanian, S. (2007). G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* 35, 406–413. doi: 10.1093/nar/gkl1057
- Izbicka, E., Wheelhouse, R. T., Raymond, E., Davidson, K. K., Lawrence, R. A., Sun, D., et al. (1999). Effects of cationic porphyrins as G-quadruplex interactive agents in human tumor cells. *Cancer Res.* 59, 639–644.
- Kang, H. J., Kendrick, S., Hecht, S. M., and Hurley, L. H. (2014). The transcriptional complex between the BCL2 i-motif and hnRNP LL is a molecular switch for control of gene expression that can be modulated by small molecules. *J. Am. Chem. Soc.* 136, 4172–4185. doi: 10.1021/ja4109352
- Karsisiotis, A. I., Hessari, N. M. A., Novellino, E., Spada, G. P., Randazzo, A., and Webba Da Silva, M. (2011). Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism. *Angew. Chem. Int. Ed. Engl.* 50, 10645–10648. doi: 10.1002/anie.201105193
- Kendrick, S., Kang, H. J., Alam, M. P., Madathil, M. M., Agrawal, P., Gokhale, V., et al. (2014). The dynamic character of the BCL2 promoter i-motif provides a mechanism for modulation of gene expression by compounds that bind selectively to the alternative DNA hairpin structure. *J. Am. Chem. Soc.* 136, 4161–4171. doi: 10.1021/ja410934b
- Kovanda, A., Zalar, M., Šket, P., Plavec, J., and Rogelj, B. (2015). Anti-sense DNA d(GGCCCC)nexpansions in C9ORF72 form i-motifs and protonated hairpins. *Sci. Rep.* 5:17944. doi: 10.1038/srep17944
- Lam, E. Y. N., Beraldi, D., Tannahill, D., and Balasubramanian, S. (2013). G-quadruplex structures are stable and detectable in human genomic DNA. *Nat. Commun.* 4, 1796. doi: 10.1038/ncomms2792
- Li, Q., Xiang, J. F., Yang, Q. F., Sun, H. X., Guan, A. J., and Tang, Y. L. (2013). G4LDB: A database for discovering and studying G-quadruplex ligands. *Nucleic Acids Res.* 41, D1115–D1123. doi: 10.1093/nar/gks1101
- Li, X., Peng, Y., Ren, J., and Qu, X. (2006). Carboxyl-modified single-walled carbon nanotubes selectively induce human telomeric i-motif formation. *Proc. Natl. Acad. Sci. U.S.A.* 103, 19658–19663. doi: 10.1073/pnas.0607245103
- Luu, K. N., Phan, A. T., Kuryavii, V., Lacroix, L., and Patel, D. J. (2006). Structure of the human telomere in K+ solution: an intramolecular (3 + 1) G-quadruplex scaffold. *J. Am. Chem. Soc.* 128, 9963–9970. doi: 10.1021/ja062791w
- Maizels, N. (2015). G4-associated human diseases. *EMBO Rep.* 16, 910–922. doi: 10.15252/embr.201540607
- Malliavin, T. E., Snoussi, K., and Leroy, J. L. (2003). The NMR structure of [Xd(C2)]4 investigated by molecular dynamics simulations. *Magn. Reson. Chem.* 41, 18–25. doi: 10.1002/mrc.1109
- Masiero, S., Trotta, R., Pieraccini, S., De Tito, S., Perone, R., Randazzo, A., et al. (2010). A non-empirical chromophoric interpretation of CD spectra of DNA G-quadruplex structures. *Org. Biomol. Chem.* 8, 2683–2692. doi: 10.1039/c003428b
- Mergny, J. L., Lacroix, L., Hélène, C., Han, X., and Leroy, J. L. (1995). Intramolecular Folding of Pyrimidine Oligodeoxynucleotides into an i-DNA Motif. *J. Am. Chem. Soc.* 117, 8887–8898. doi: 10.1021/ja00140a001
- Mir, B., Serrano, I., Buitrago, D., Orozco, M., Escaja, N., and González, C. (2017). Prevalent sequences in the human genome can form mini i-Motif structures at physiological pH. *J. Am. Chem. Soc.* 139, 13985–13988. doi: 10.1021/jacs.7b07383
- Monchard, D., Allain, C., and Teulade-Fichou, M.-P. (2006). Development of a fluorescent intercalator displacement assay (G4-FID) for establishing quadruplex-DNA affinity and selectivity of putative ligands. *Bioorg. Med. Chem. Lett.* 16, 4842–4845. doi: 10.1016/j.bmcl.2006.06.067
- Pagano, B., Amato, J., Iaccarino, N., Cingolani, C., Zizza, P., Biroccio, A., et al. (2015). Looking for efficient G-quadruplex ligands: evidence for selective stabilizing properties and telomere damage by drug-like molecules. *ChemMedChem* 10, 640–649. doi: 10.1002/cmdc.201402552
- Pagano, B., Cosconati, S., Gabelica, V., Petraccone, L., De Tito, S., Marinelli, L., et al. (2012). State-of-the-art methodologies for the discovery and characterization of DNA G-quadruplex binders. *Curr. Pharm. Des.* 18, 1880–1899. doi: 10.2174/138161212799958332
- Pagano, B., Fotticchia, I., De Tito, S., Mattia, C., a, Mayol, L., Novellino, E., et al. (2010). Selective Binding of Distamycin A Derivative to G-Quadruplex Structure [d(TGGGGT)]4. *J. Nucleic Acids* 2010, 1–7. doi: 10.4061/2010/247137
- Pagano, B., Mattia, C. A., Virno, A., Randazzo, A., Mayol, L., and Giancola, C. (2007). Thermodynamic analysis of quadruplex DNA-drug interaction. *Nucleosides Nucleotides Nucleic Acids* 26, 761–765. doi: 10.1080/15257770701499069
- Phan, A. T., Guéron, M., and Leroy, J. L. (2000). The solution structure and internal motions of a fragment of the cytidine-rich strand of the human telomere. *J. Mol. Biol.* 299, 123–144. doi: 10.1006/jmbi.2000.3613
- Rajendran, A., Nakano, S., and Sugimoto, N. (2010). Molecular crowding of the cosolutes induces an intramolecular i-motif structure of triplet repeat DNA oligomers at neutral pH. *Chem. Commun.* 46, 1299–1301. doi: 10.1039/b922050j
- Randazzo, A., Galeone, A., Esposito, V., Varra, M., and Mayol, L. (2002). Interaction of distamycin a and netropsin with quadruplex and duplex structures: A comparative 1H-NMR study. *Nucleosides, Nucleotides Nucleic Acids* 21, 535–545. doi: 10.1081/NCN-120015067
- Randazzo, A., Spada, G. P., and Da Silva, M. W. (2013). Circular dichroism of quadruplex structures. *Top. Curr. Chem.* 330, 67–86. doi: 10.1007/128_2012_331
- Rodriguez, R., Müller, S., Yeoman, J. A., Trentesaux, C., Riou, J. F., and Balasubramanian, S. (2008). A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. *J. Am. Chem. Soc.* 130, 15758–15759. doi: 10.1021/ja805615w

- Salvati, E., Botta, L., Amato, J., Di Leva, F. S., Zizza, P., Gioiello, A., et al. (2017). Lead Discovery of Dual G-Quadruplex Stabilizers and Poly(ADP-ribose) Polymerases (PARPs) Inhibitors: a new avenue in anticancer treatment. *J. Med. Chem.* 60, 3626–3635. doi: 10.1021/acs.jmedchem.6b01563
- Salvati, E., Zizza, P., Rizzo, A., Iachettini, S., Cingolani, C., D'angelo, C., et al. (2014). Evidence for G-quadruplex in the promoter of vegfr-2 and its targeting to inhibit tumor angiogenesis. *Nucleic Acids Res.* 42, 2945–2957. doi: 10.1093/nar/gkt1289
- Sheng, Q., Neaverson, J. C., Mahmoud, T., Stevenson, C. E. M., Matthews, S. E., and Waller, Z. A. E. (2017). Identification of new DNA i-motif binding ligands through a fluorescent intercalator displacement assay. *Org. Biomol. Chem.* 15, 5669–5673. doi: 10.1039/C7OB00710H
- Siddiqui-Jain, A., Grand, C. L., Bearss, D. J., and Hurley, L. H. (2002). Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U.S.A.* 99, 11593–11598. doi: 10.1073/pnas.182256799
- Sun, D., and Hurley, L. H. (2009). The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: Implications for drug targeting and control of gene expression. *J. Med. Chem.* 52, 2863–2874. doi: 10.1021/jm900055s
- Wright, E., Lamparska, K., Smith, S. S., and Waller, Z. A. E. (2017). Substitution of cytosine with guanylurea decreases the stability of i-Motif DNA. *Biochemistry* 56, 4879–4883. doi: 10.1021/acs.biochem.7b00628
- Wright, E. P., Day, H. A., Ibrahim, A. M., Kumar, J., Boswell, L. J. E., Huguin, C., et al. (2016a). Mitoxantrone and analogues bind and stabilize i-motif forming DNA sequences. *Sci. Rep.* 6:39456. doi: 10.1038/srep39456
- Wright, E. P., Huppert, J. L., and Waller, Z. A. E. (2016b). Identification of multiple genomic DNA sequences which form i-motif structures at neutral pH. *Nucleic Acids Res.* 45, 2951–2959. doi: 10.1093/nar/gkx090
- Xu, B., Devi, G., and Shao, F. (2015). Regulation of telomeric i-motif stability by 5-methylcytosine and 5-hydroxymethylcytosine modification. *Org. Biomol. Chem.* 13, 5646–5651. doi: 10.1039/C4OB02646B
- Xu, L., Hong, S., Sun, N., Wang, K., Zhou, L., Ji, L., et al. (2016). Berberine as a novel light-up i-motif fluorescence ligand and its application in designing molecular logic systems. *Chem. Commun.* 52, 179–182. doi: 10.1039/C5CC08242K
- Zeraati, M., Langley, D. B., Schofield, P., Moye, A. L., Rouet, R., Hughes, W. E., et al. (2018). I-motif DNA structures are formed in the nuclei of human cells. *Nat. Chem.* 10, 631–637. doi: 10.1038/s41557-018-0046-3
- Zhao, J., Bacolla, A., Wang, G., and Vasquez, K. M. (2010). Non-B DNA structure-induced genetic instability and evolution. *Cell. Mol. Life Sci.* 67, 43–62. doi: 10.1007/s00018-009-0131-2
- Zhou, J., Wei, C., Jia, G., Wang, X., Feng, Z., and Li, C. (2010). Formation of i-motif structure at neutral and slightly alkaline pH. *Mol. Biosyst.* 6, 580–586. doi: 10.1039/B919600E
- Zizza, P., Cingolani, C., Artuso, S., Salvati, E., Rizzo, A., D'Angelo, C., et al. (2016). Intragenic G-quadruplex structure formed in the human CD133 and its biological and translational relevance. *Nucleic Acids Res.* 44, 1579–1590. doi: 10.1093/nar/gkv1122

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Pagano, Iaccarino, Abdelhamid, Brancaccio, Garzarella, Di Porzio, Novellino, Waller, Pagano, Amato and Randazzo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.