



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

SEDE AMMINISTRATIVA: UNIVERSITA' DEGLI STUDI DI PADOVA
DIPARTIMENTO DI SCIENZE DEL FARMACO

SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE MOLECOLARI

INDIRIZZO FARMACEUTICO

XXVII CICLO

**CONFORMATIONAL SWITCH OF ONCOGENE
PROMOTORIAL SEQUENCES TOWARDS
NON-CANONICAL DNA SECONDARY STRUCTURES**

Direttore della Scuola : *Ch.mo* Prof. Antonino Polimeno

Coordinatore d'indirizzo: *Ch.mo* Prof. Stefano Moro

Supervisore : *Ch.ma* Prof.ssa Claudia Sissi

Dottoranda: Maria Laura Greco

31 GENNAIO 2015

A tutti coloro che hanno creduto in me

*“Iniziare un nuovo cammino spaventa.
Ma dopo ogni passo che percorriamo ci rendiamo conto
di come era pericoloso rimanere fermi”.*
Roberto Benigni.

TABLE OF CONTENTS

ABSTRACT	pag.III
RIASSUNTO	pag.V
<u>Chapter 1: STRUCTURAL CHARACTERIZATION OF EGFR AND BRAF PROMOTORIAL REGIONS</u>	
1.1 INTRODUCTION	pag.1
1.1.1 DNA polymorphism	pag.5
1.1.2 The G-quadruplex	pag.8
1.1.3 G-quadruplex in the cell	pag.12
1.1.3.1 G4 at telomeric level	pag.12
1.1.3.2 G4 structures as switching on/off of oncogene transcription	pag.16
1.1.4 The effect of cosolutes on G4 equilibria	pag.18
1.1.5 Targeting G4 structures	pag.21
1.1.6 I-motif structure	pag.26
1.2. AIM OF THE STUDY	pag.28
1.3. RESULTS AND DISCUSSIONS	pag.32
1.3.1 G4 characterization	pag.32
1.3.1.1 EGFR-272	pag.32
1.3.1.2 EGFR-37	pag.37
1.3.1.3 BRAF-176	pag.41
1.3.2 I-motif conformational study	pag.44
1.3.4 Effect of crowding conditions on i-motif structure	pag.48
1.3.5 Double strand stability at physiological conditions	pag.50
1.4. CONCLUSIONS	pag.59

Chapter 2: G-QUADRUPLEX FOLDING OF ONCOGENE PROMOTERS IN THE CELL pag.62

Chapter 3: BINDING STUDY OF ANTHRAQUINONES DERIVATIVES ON G4 TEMPLATES

3.1. INTRODUCTION	pag.68
3.2. AIM	pag.70
3.3 RESULTS AND DISCUSSIONS	pag.72
3.3.1 Anthraquinones derivatives on oncogenic sequences	pag.81
3.4. CONCLUSIONS	pag.84

Chapter 4: DISUBSTITUTED NAPHTHALENE DIIMIDES AS POTENTIAL SELECTIVE BINDERS FOR G4 TEMPLATES

4.1. INTRODUCTION	pag.85
4.2. AIM OF THE STUDY	pag.87
4.3. RESULTS AND DISCUSSIONS	pag.90
4.4 CONCLUSIONS	pag.107

Chapter 5: EXPERIMENTAL SECTION

5.1 Spectroscopic techniques	
5.1.1 Fluorescence melting assay	pag.109
5.1.2 Circular dichroism (CD) measurements	pag.110
5.1.3 Thermal differential spectrum (TDS)	pag.110
5.1.4 Surface Plasmon Resonance (SPR)	pag.111
5.1.5 Isothermal Titration Calorimetry (ITC)	pag.111
5.2 Electrophoretic techniques	
5.2.1 Electromobility Shift Assay	pag.112
5.2.2 Topoisomerase inhibition	pag.112
5.2.3 DNA polymerase stop assay	pag.113
5.2.4 Taq polymerase inhibition assay	pag.113

5.2.5 Telomerase Repeat Amplification Protocol (TRAP) assay	pag.113
5.2.6 Adducts formation	pag.114
5.2.7 DMS-treatment and LM-PCR (Ligation Mediated-Polymerase Chain Reaction)	pag.114
5.3 Cellular experiments	
5.3.1 Protein extract preparation	pag.118
5.3.2 MTT assay	pag.118
<u>References</u>	pag.119
<u>Appendix</u>	pag.131

LIST OF ABBREVIATIONS

AQ: anthraquinones

bp: base pairs

CD: circular dichroism

ctDNA: calf thymus DNA

ds: double strand

EDTA: ethylenediaminetetraacetic acid

EGFR: Epidermal Growth Factor Receptor

EMSA: electrophoretic mobility shift assay

FAM: 6-carboxyfluorescein

G4: G-quadruplex

HTS: Human Telomeric Sequence

iM: i-motif

ITC: Isothermal Titration Calorimetry

NDI: naphthalene diimides

PCR: polymerase chain reaction

RU: response units

SDS: sodium dodecyl sulfate

SPR: Surface Plasmon Resonance

ss: single strand

TAE: Tris- Acetate EDTA

Taq: thermos aquaticus

TBE: Tris-Borate EDTA

Tel: telomere

T_m: melting temperature

TRAP: Telomeric Repeat Amplification Protocol

TRIS: tetrahydroxymethylaminoethane

UV: ultraviolet

VIS: visible

ABSTRACT

The folding of DNA molecule into non-canonical secondary structures has been shown to be implicated in many important biological processes which regulate cell proliferation and proteins expression. In particular one of these peculiar secondary structures, called G-quadruplex (G4), has been shown to potentially impair cancer development. G4 occurs along DNA sequences rich of consecutive guanines which can fold through Hoostein pairs by forming stacked planes of guanines tetrads. This conformation prevalently forms along the termini of chromosomes (telomeres) but also along the promoter sites of several oncogenes directly involved in many cancers. The G4 formation leads to an hindrance on DNA molecule which hinder the telomere elongation and transcription process. The result is a switching off of these mechanisms which are directly involved in cancer progression. Several factors can influence the G4 equilibria for example, saline conditions, temperature, pH, the binding with specific proteins as well as the presence of dehydrating cosolutes. Additionally, the overall structural feature of the G4 is strictly dependent upon the DNA sequence. As a results, different G4 can be identified inside the cells.

In this project, we focused on the conformational study of the promotorial regions of EGFR and BRAF oncogenes since, on these sites the existence of G4 putative forming regions was found. In particular, the sequences at positions -272, -37 of EGFR and -176 of BRAF from the transcription start site were analyzed. Indeed, no previous literature data were reported about the structural equilibria in solution of these sequences. We found that our tested sequences are actually able to fold into G4 by setting the most proper experimental conditions and also close to the intracellular physiological environment (KCl 150 mM, pH 7.5).

However, oncogenes are double stranded sequences and the folding of the complementary cytosine rich strand into i-motif (iM) can be involved in the switching off of gene transcription. Although, so far, no physiological evidence has been observed for i-motif conformation, here, we aimed to investigate also the cytosine rich strand conformation, to assess if this folding in the case of our sequences is compatible with the physiological conditions and if it can synergically works with the G4 to destabilize the double strand. Our data showed that in physiological condition the preferential form is represented by the double strand . However, some selected ligands showed to shift the DNA B-form toward the

non canonical conformation. Indeed, here we implemented our work with the screening of two libraries of compounds in order to find a selective and efficient binder. We carried on the binding study of anthraquinones and naphthalene diimides derivatives, known to have the chemical features of efficient G4 binders. These ligands were first tested on different G4 templates, known to be validated models for G4 binding study, and their efficiency on G4 has been compared with the double strand. The most G4 selective derivatives were then investigated towards our oncogenic G4s. Although more work is required to identify a lead compound, we were able to demonstrate how the use of asymmetrical substitution pattern on an aromatic core can implement the selectivity among different G4s.

Finally, in order to map the occurrence of G4 conformation *in vivo*, we set up a novel technique which consists in an *in vivo* footprinting protocol. This work, performed at University of Mississippi, Oxford, MS (USA), under the supervision of Dr Tracy A. Brooks, should provide novel insight on the G4 formation in the cells according to their physiological and environmental conditions.

RIASSUNTO

Molti studi dimostrano che l'assunzione di strutture "non canoniche" da parte della molecola di DNA sia coinvolto in molti importanti processi biologici che regolano la proliferazione cellulare e l'espressione proteica. In particolare, è stata dimostrata l'implicazione di una di queste particolari strutture secondarie, chiamata G-quadruplex (G4), nel blocco della progressione del cancro. La struttura G4 è propria di sequenze di DNA ricche in guanine consecutive che assemblandosi tramite legami di Hoostein, formano piani di tetradi di guanine impilati tra loro. Questa particolare conformazione si forma prevalentemente lungo i tratti terminali dei cromosomi, i telomeri, ma anche lungo siti promotoriali di diversi oncogeni coinvolti in molti tipi di cancro. La formazione del G4 porta ad una sorta di ingombro sulla molecola di DNA che inibisce l'elongazione del telomero e i processi di trascrizione. Questo porta ad uno "spegnimento" di questi meccanismi che sono direttamente coinvolti nello sviluppo del cancro. Molti fattori possono influenzare gli equilibri delle conformazioni G4, per esempio, le condizioni saline, la temperatura, il pH, il legame con specifiche proteine, così come la presenza di cosoliti. Inoltre, la struttura globale del G4 è rigorosamente dipendente dalla sequenza oligonucleotidica. Pertanto, diverse strutture G4 possono essere identificate a livello cellulare.

In questo progetto, è stato condotto uno studio conformazionale di regioni promotoriali degli oncogeni EGFR e BRAF, dal momento che, su questi oncogeni è stata riscontrata la presenza di regioni "G-rich" (ricche in guanine) potenzialmente in grado di assumere una struttura G4. In particolare, sono state analizzate le sequenze a partire dalle posizioni -272, -37 di EGFR e -176 di BRAF dal "transcription start site" (sito di inizio della trascrizione). Finora, non sono presenti dati in letteratura riguardanti la caratterizzazione strutturale di queste sequenze in soluzione. Con questo studio, è stata dimostrata la capacità delle suddette sequenze di assumere una conformazione G4 nelle idonee condizioni sperimentali e soprattutto in un ambiente che mimi quello fisiologico (150mM KCl e pH 7.5). Poiché gli oncogeni sono sequenze a doppio filamento, anche la conformazione i-motif assunta dal filamento complementare ricco in citosine ("C-rich") può essere coinvolta nella regolazione del processo di trascrizione genica. Tuttavia, sinora non è stata riscontrata alcuna rilevanza fisiologica della conformazione i-motif. In questo lavoro, è stata caratterizzata anche la

conformazione assunta dal filamento “C-rich”, in particolare se essa possa esistere in condizioni fisiologiche e se fosse in grado di destabilizzare la doppia elica insieme al G4. I dati ottenuti dimostrano che in condizioni fisiologiche la forma prevalente è il doppio filamento. Tuttavia, è stato dimostrato come alcuni ligandi siano in grado di spostare l’equilibrio del DNA dalla sua forma di doppia elica-B, verso le conformazioni non canoniche. È stato infatti condotto uno studio su due librerie di composti con lo scopo di evidenziare un composto selettivo ed efficace. Ci siamo focalizzati su derivati antrachinonici e di naftalendiimidi noti come efficaci ligandi per il G4. Questi composti sono stati prima testati su diversi templati G4, noti per essere dei modelli validati per lo studio di binding sul G4. Quindi la loro efficienza sul G4 è stata poi comparata a quella sul doppio filamento. I derivati più selettivi verso il G4 sono stati poi testati su G4 oncogenici. Sebbene una continuazione dello studio fosse necessaria per identificare un composto “lead”, con questo lavoro è stato dimostrato come l’uso di una sostituzione asimmetrica sull’anello aromatico possa implementare la selettività tra più G4.

Infine, per identificare la formazione del G4 *in vivo*, è stata messa a punto una nuova tecnica che consiste in un protocollo di footprinting *in vivo*. Questo lavoro, svolto nell’Università del Mississippi, Oxford, MS (USA) sotto la supervisione della dr.ssa Tracy A. Brooks, dovrebbe fornire nuovi sviluppi per la formazione del G4 nelle cellule in accordo con le loro condizioni fisiologiche.

Chapter 1

STRUCTURAL CHARACTERIZATION OF EGFR AND BRAF PROMOTORIAL REGIONS

1.1 INTRODUCTION

Human and animal cancers are characterized by shared properties, which comprise invasion, metastasis, replicative immortality, angiogenesis, cell death resistance, sustained proliferative signaling and evasion of growth suppressors (fig. 1.1). These six hallmarks of cancer were first outlined by Hanahan and Weinberg in 2000 (Hanahan and Weinberg, 2000). Each of them is associated to peculiar gene mutations that bring to the expression of oncogenes with dominant gain of function or tumor suppressor genes with recessive loss of function. These gene mutations are responsible of the molecular, biochemical and cellular properties shared by the most, even all, types of cancers. Although various types of mutations, such as base substitutions, deletions, and insertions, induce proto-oncogenes overexpression, it has been shown that human cancers are mainly associated to point mutation of oncogenes(Kufe DW, Pollock RE, Weichselbaum RR, et al., 2003)

Of the 50.000-100.000 genes in our genome, about 100 have been identified as proto-oncogenes and about a dozen as tumor suppressor genes. Some examples of proto-oncogenes are *myc*, *raf*, *ras*, growth factors etc...which overexpression has been associated to different kind of cancer such as lymphomas, stomach, colon and lung carcinomas.

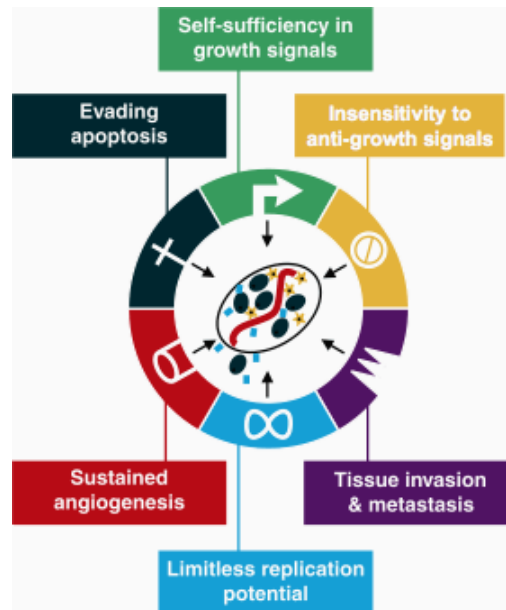


Fig.1.1: the six hallmarks of cancer

Conversely, tumor suppressors act by protecting cell from cancer progression by different mechanisms which are not completely known yet. Generally, they can suppress cell division, induce apoptosis and DNA damage repair and inhibit metastasis. Among them, the most important is p53, which misregulation is associated with about 50% of human cancers. It acts with more than one mechanism, for example by triggering DNA repair processes, inducing the transcription of other tumor suppressors, such as p21 and p16, and initiating cell apoptosis (Sun and Yang, 2010). Anticancer therapy uses different approaches depending on cancer progression and localization. They include surgery, radiotherapy, immunological therapy usually all combined with chemotherapy. The first anticancer compounds were introduced in 1948, when nitrogen mustards were used as weaponry during World War I. Their potential to cause bone marrow aplasia was exploited as a chemical approach for leukemia treatment. Besides mustards, during years, more other compounds were synthesized as anticancer agents such as folic acid derivatives, nitrogen bases synthesis inhibitors and vinca alkaloids (fig. 1.2).

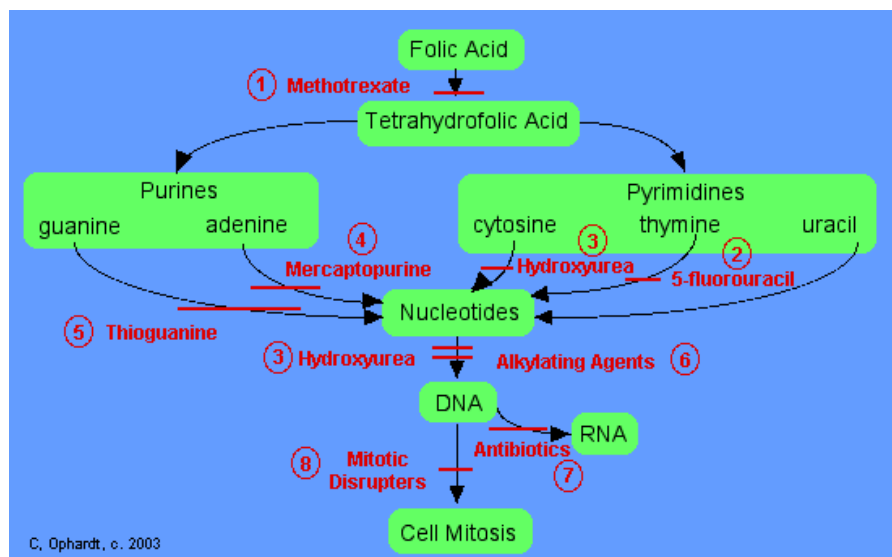


Fig.1.2: mechanism of action of anticancer drugs

Although chemotherapy approach is often essential for cancer eradication, the main negative aspect regards the high number of side effects. Nowadays, modern anticancer therapy looks at limiting this aspect with the utopia to develop a specific drug for a specific illness.

Since the discovery of the chemotherapeutics, the delivery of the drugs only to the damaged portion of the nucleic acids has been pursued with the aim to massively reduce side effects. These effects however, proved to partially fail when small molecules are used for their inability to “read” unique DNA sequences.

Cancer development is strictly related to specific mutations at DNA level which lead to the overexpression of defined growth factors or the lost of expression of tumor suppressing genes promoting cancer cells growth and immortality. For this reason, the targeting of DNA sequences involved in cancer could be a significant breakthrough in the field of cancer therapy.

In this contest, a novel strategy, involves gene therapy approach. Conceptually, the technique involves the identification of appropriate DNA sequences which can

interfere, once injected in the cell, with DNA translation processes or possibly even correct genetic mutations.

Many different approaches can be used for cancer gene therapy depending on the kind of the disorder. An example is the delivery of gene-specific antisense oligonucleotide or ribozyme to bind/cleave oncogene mRNA. A cancer gene therapy trial which is being conducted regards the delivery of antisense KRAS genes in the case of some forms of non-small-cell lung cancer (Strachan and Read, 1999).

The main issues regardless gene therapy approaches is linked to the delivery of nucleic acids and also the immunological response which can be caused by the introduction of foreign oligonucleotides. Nowadays, the sequence targeting approach is going to be replaced by the structure-specific recognition approach by the use of specific ligands for specific DNA structures.

1.1.1 DNA polymorphism

The predominant secondary structure assumed by DNA in solution is the right handed B-double helix. In this conformation, nitrogen bases are paired following Watson and Crick rules (Watson and Crick, 2003). The B-double helix is 23.7 Å wide and extends 34 Å per 10 bp of sequence. It makes one complete turn about its axis every 10.4-10.5 base pairs in solution. This frequency of twist (known as the helical pitch) depends largely on stacking forces that each base exerts on its neighbours in the chain. The bases are oriented perpendicular to the helix axis. The interaction energy between two bases in double-helical structure is therefore a combination of hydrogen-bonding between complementary bases, and hydrophobic interactions between the neighboring stacks of base-pairs.

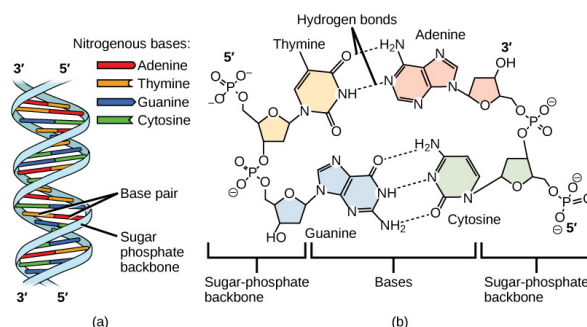


Fig.1.3: representation of Watson and Crick base pairings.

Besides the Watson and Crick pairings, DNA can assume other secondary structures, which commonly depend on solution conditions (salt, dehydration, binding with nuclear proteins), such as the A and Z form of the double helix.

The Z-DNA helix is an unstable structure of the double helix, it is left-handed and has a structure that repeats every 2 base pairs. The major and minor grooves, unlike A- and B-DNA, show little difference in width, it can also form a junction with B-DNA (called a "B-to-Z junction") in a structure which involves the extrusion of a base pair (de Rosa et al., 2010).

Furthermore, they exist some other secondary structures classified as “non-canonical” as they completely differ from the common double helix. Some examples are cruciform DNA, slipped hairpins, triplexes, G-quadruplexes (G-tetraplexes), and i-motifs (i-tetraplexes) (fig 1.4).


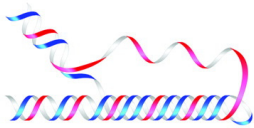
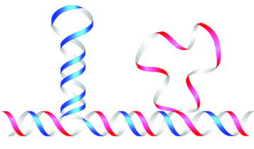
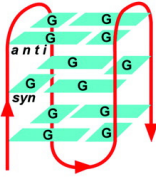

Name	Conformation	General Seq. Requirements	Sequence
Cruciform		Inverted Repeats	<pre> TCGGTACCGA AGCCATGGCT </pre>
Triplex		$(R \cdot Y)_n$ Mirror Repeats	<pre> AAGAGG GGAGAA TTCTCC CCTCTT </pre>
Slipped (Hairpin) Structure		Direct Repeats	<pre> TCGGTTCGGT AGCCAAGCCA </pre>
Tetraplex		Oligo $(G)_n$ Tracts	$AG_3(T_2AG_3)_3$ single strand
Left-handed Z - DNA		$(YR \cdot YR)_n$	<pre> CGCGTGC GTGTG GCGCACGCACAC </pre>

Fig. 1.4: examples of non-canonical DNA secondary structures (Bacolla and Wells, 2004).

These structures are involved in many important biological processes, including replication, regulation of gene expression, nucleosome structure and recombination. For example, cruciform DNA, are characterized by the presence of inverted complementary repeats thus inducing the folding of the one strand sequence, and they have been implicated in the evolution and development of diseases including cancer, Werner's syndrome and others (Brazda et al., 2011). Furthermore, a DNA

triplex is formed when pyrimidine or purine bases occupy the major groove of the DNA double helix through Hoogsteen pairs with purines of the Watson-Crick basepairs (Frank-Kamenetskii and Mirkin, 1995). Their formation prevents the processing of the DNA by several nuclear proteins. Hairpins is an intramolecular folding of the single stranded DNA and it occurs when two regions of the same strand, usually complementary in nucleotide sequence pair to form a double helix that ends in an unpaired loop.

1.1.2 The G-quadruplex

The G-quadruplex (G4) is a “non-canonical” secondary structure assumed by gene sequences rich in consecutive guanines (“G-rich sequences”). In this case, four guanines assemble each other through Hoogsteen hydrogen bonds, involving the N1, N7, O6, and N2 of guanine molecule, forming the G-tetrad plane. The stacking of more G-tetrads brings to the G4 structure (fig. 1.5, 1.6).

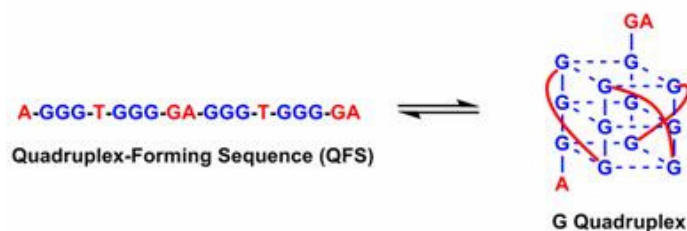


Fig. 1.5: guanines assembly in a G-quadruplex conformation([age855.html](#)).()

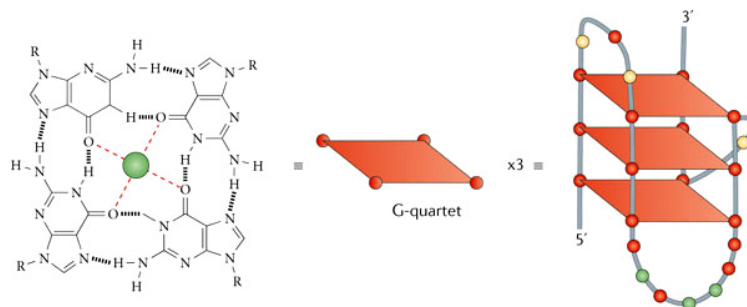


Fig. 1.6: stacking of guanines tetrads to form a G4 conformation(Balasubramanian et al., 2011)

Several factors can influence this conformation for example, saline conditions, the number of guanines, molecular crowding and dehydrating agents as well as the binding with nuclear proteins. The presence of salts is very important for quadruplex stability. As the O6 guanine carbonylic group determines a negative charged cavity in the center of the tetrad plane, the selection of a suitable cation based on size and charge dramatically determines the overall stability of the final folded quadruplex. The contribution of several ions has been studied so far, such as K^+ , Na^+ , NH_4^+ , Li^+ etc...,

The most relevant turned out to be Na^+ and K^+ which are also largely prevalent in the physiological environment. Since Na^+ has a smaller ionic radius compared to K^+ , it intercalates in the tetrad plane conversely, K^+ interposes between two tetrads (fig. 1.7).

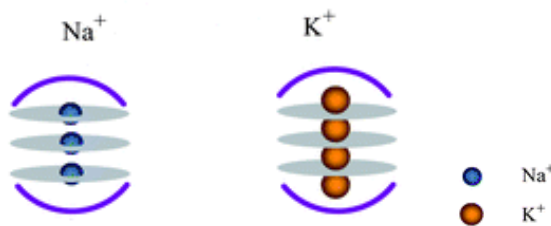


Fig. 1.7: stabilizing effect of K^+ and Na^+ cations. Na^+ intercalates in the tetrad plane, whereas K^+ among two stacked tetrads.

The preference of quadruplex central cavity for potassium over sodium ion is the result of two opposite effects: from one side the free energy of Na^+ binding to a G4 is more favorable than that of K^+ , but from the other side this effect is neutralized by a greater cost for Na^+ dehydration. The net result is a more negative free energy change for the potassium form.

Besides stability, these factors can influence also the topology of the G-quadruplex structure. The switching between various conformations can be induced by ions, temperature and, above all, by the bases sequence (Neidle and Balasubramanian, 2006). In order to classify, the different G4 several properties can be considered. Indeed, topology is determined by the relationship between backbone strand orientation (+,-), glycosidic torsion angles (*syn-anti*-) and groove widths (w,m,n). The glycosidic is the one bond linking the guanine base (N9) to the sugar in the G-tetrad; the torsion angle χ (C4-N9-C1'-O4') is limited in the stacked tetrad to be either *syn* (0 to 90°) or *anti* (-120 to 180°). In turn, depending on strand orientation the four G-containing strand can be all parallel, antiparallel or mixed type (fig. 1.9)(Neidle and

Balasubramanian, 2006; Patel et al., 2007; Webba da Silva, 2007; Webba da Silva et al., 2009)

Strand orientation and glycosidic torsion angle contribute to define groove widths: all-parallel backbone strand alignments generate equal groove widths, while anti-parallel arrangements of the phosphate backbones generate both wide and narrow grooves. The combinations of *syn* and *anti* base orientations around the tetrads changes significantly the access to hydrogen bond donors and acceptors within the grooves, altering hydration networks, accessibility and a variety of surfaces for proteins, DNA and ligands interactions(Collie and Parkinson, 2011).

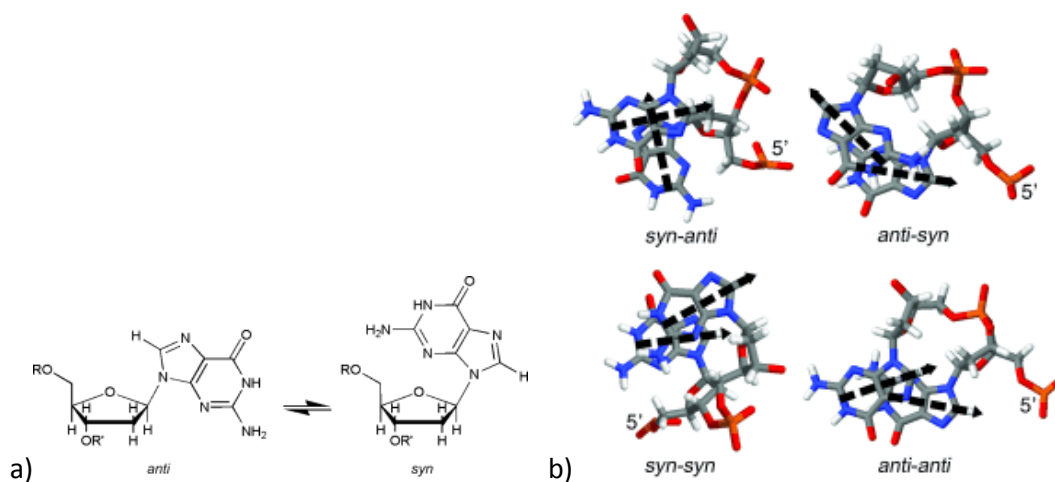


Fig. 1.8: chemical structures of *anti* and *syn* conformers of 2'-deoxyguanosine a)(Webba da Silva et al., 2009) (Webba da Silva, 2007)

Another level of G4 polymorphism is related to the number of strands it involves. For example, G4 can be monomeric, dimeric or tetrameric.

Combining strand orientation, glycosidic torsion angle, groove width and number of strands, different G4 structures occur; as a result, G4 can be compared to a library of different tridimensional structures where each single topology can be a valuable molecular target to study important biological processes and to develop possible selective therapeutic ligands.

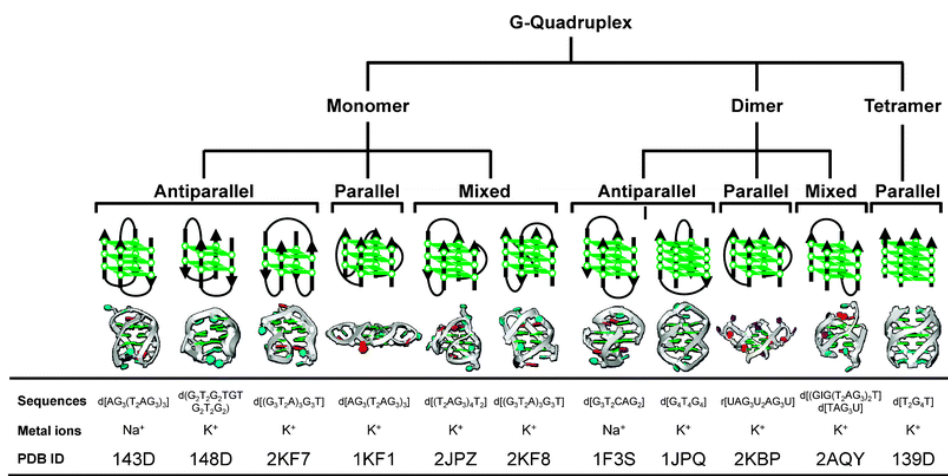


Fig. 1.9: G4 polymorphism(Yaku et al., 2012).

1.4 G-quadruplex in the cell

The first discovery of guanines assembly to form a planar tetrad, is dated back to 1962, when Gellert M. et al. proposed a model of arrangement of the bases in a guanylic acid gel (GELLERT et al., 1962). The recognition of G4 structures and their implication in many biological processes such as aging and cancer, led many researchers to point their attention on this topic. Only several years later, nucleic acid G4, was first proposed to occur along the termini of eukaryotic chromosomes, telomeres, due to the presence of consecutive guanines on these sequences $-(TTAGGG)_n-$. However, G-rich sequences were identified not only along telomeres but also along different genomic DNA sites. For this reason, G4 structures can occur also along promoter sites and the 5'UTR coding regions of several oncogenes, transient single stranded DNA occurring during replication processes as well as on 5'UTR of mRNA outside the nucleus (fig. 1.10) (Lipps and Rhodes, 2009).

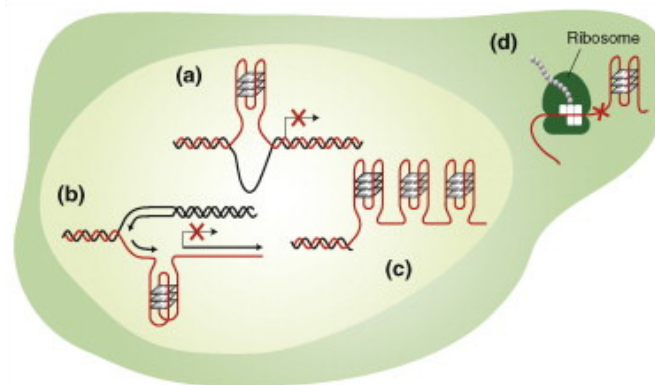


Fig. 1.10: G-quadruplex formation at different level into the cell (Lipps and Rhodes, 2009).

1.4.1 G4 at telomeric level

Telomeres consist of tandem repeats of $(TTAGGG/CCCTAA)_n$ where the G-rich strand protrudes at the 3' terminal as single stranded overhang of about 100-300 bases. Telomere length is reduced at each cell replication cycle due to the end-replication effect. Since DNA polymerase is not capable to completely copy DNA at the very ends

of chromosomes, approximately 50 nucleotides are lost during each cell cycle, which results in gradual telomere length shortening and loss of genetic information. When the number of tandem repeats falls below a critical value, the cell stops to divide and enters a state of senescence (Cech, 2004; Zvereva et al., 2010). The enzyme responsible for the physiological maintenance of the length of telomeres is telomerase. This is a ribonucleoprotein which acts as a reverse transcriptase by adding guanine-rich repetitive sequences and reducing telomeres shortening (Zvereva et al., 2010). Nevertheless, telomerase is overexpressed in the 80-90% of cancer cells, leading to an uncontrolled cellular proliferation and consequent immortalization. The shift of the conformational balance of telomeres towards G-quadruplex, has been proposed as a novel strategy to prevent telomerase activity because the enzyme recognizes only linear oligonucleotides (Fig.1.11). It is for this reason that the development of selective drugs, which recognize and stabilize the G-quadruplex folded form of telomeres could be a great strategy to control cancer diffusion (Neidle and Balasubramanian, 2006)(Sun et al., 1997).

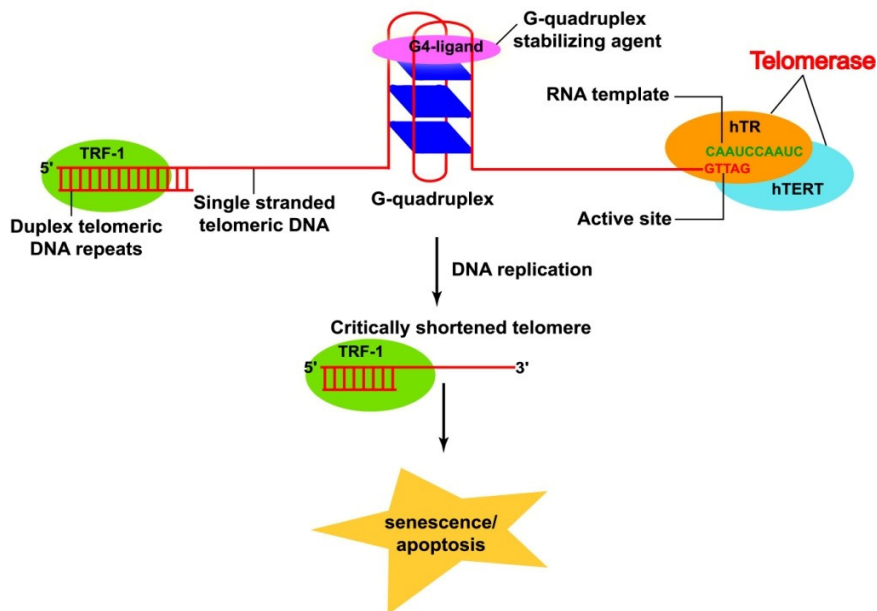


Fig. 1.11: telomerase inhibition by G4 structure at telomeric level
(<http://www.chem.iitb.ac.in/~pradeep/G-quadruplex.html>).

Nevertheless, also telomere is characterized by a high polymorphism which mainly depends on sequence length and salt conditions. NMR and X-ray evidences assess the potential of the Human Telomeric Sequence (HTS) ($d[AG_3(T_2AG_3)_3T]$) to mainly fold in a population of prevalently hybrid conformations in the presence of K^+ . Conversely, in Na^+ it assumes a defined antiparallel folding. Furthermore, Tel24 [$d(T_2AG_3)_4$] and mutTel26 $d[(A_3(G_3T_2A)_3G_3A_2)]$ adopt a hybrid-1 folding in K^+ containing solutions, whereas wtTel26 $d[(T_2AG_3)_4T_2]$ in the same conditions, folds into a hybrid-2 type of arrangement (fig. 1.12).

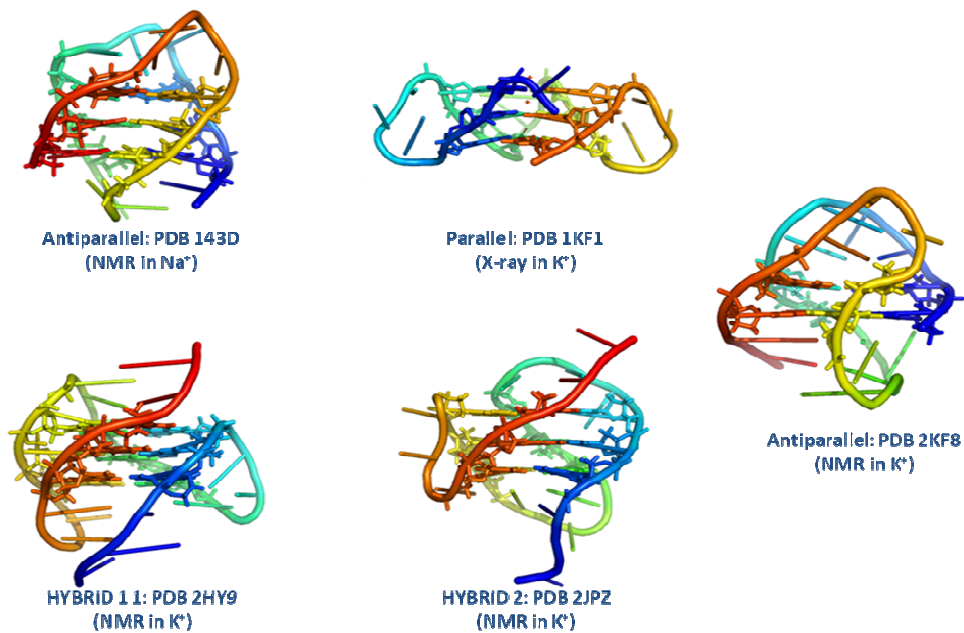


Fig. 1.12: X-ray and NMR representation of telomeric G4 structures assumed at different saline conditions.

Here is reported a resuming table with the topologies assumed by short telomeric sequences of different length in K^+ (fig. 1.13).

wtTel27	TTAGGGTTAGGGTTAGGGTTAGGGTTA	(hybrid-2)
wtTel26	TTAGGGTTAGGGTTAGGGTTAGGGTT	(hybrid-2)
wtTel25b	TAGGGTTAGGGTTAGGGTTAGGGTT	(hybrid-2)
wtTel24b	AGGGTTAGGGTTAGGGTTAGGGTT	(hybrid-2)
wtTel25a	TTAGGGTTAGGGTTAGGGTTAGGGT	(hybrid-2/1)
wtTel24a	TAGGGTTAGGGTTAGGGTTAGGGT	(hybrid-2/1)
wtTel23a	AGGGTTAGGGTTAGGGTTAGGGT	(hybrid-2)
wtTel24	TTAGGGTTAGGGTTAGGGTTAGGG	(hybrid-1)
wtTel23	TAGGGTTAGGGTTAGGGTTAGGG	(hybrid-1)

Fig. 1.13: G4 structures assumed by different length telomeres(Dai et al., 2007).

To date the existence of several G4 topologies under physiological conditions has been shown and the predominance of a unique G4 structure of the telomeric sequence *in vivo* has not been identified yet. Several works aim to outline this aspect by trying to isolate and identify G4 from genomic DNA. An important breakthrough was reached in 2013, when S.Balasubramanian research group visualized the G4 structures at the telomeres of ciliate macronuclei using a fluorescent labelled antibody, selective for G4 structures. In fig. 1.14 it can be observed the G4 formation within the telomeres. Interestingly, this approach evidenced also the occurrence of G4 structure at non-telomeric regions (Biffi et al., 2013).

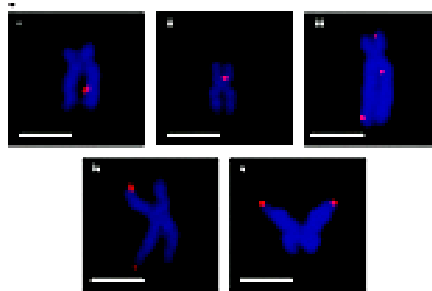


Fig. 1.14: immunofluorescence for G4 structure on metaphase chromosomes isolated from HeLa cervical cancer cells. Discrete G4 foci (red) were observed both within the non-telomeric regions and at the telomeres. Adapted from:(Biffi et al., 2013)

1.4 2. G4 structures as switching on/off of oncogenes transcription

Oncogenes are physiologically expressed in humans, where they are involved in cell proliferation and differentiation. Several factors can alter the physiological expression of these oncogenes leading to their overexpression or to the expression of mutant forms which can be responsible for cancer development (Huppert and Balasubramanian, 2005; Huppert and Balasubramanian, 2007; Huppert et al., 2008; Hurley, 2009; Zhao et al., 2010). Bioinformatic analysis identified G-rich sequences upstream the promotorial region of several oncogene but not on tumor suppressor gene. In this regard, a correlation between the G4 folded state of the promoter sequence and the repression of gene transcription exists, which confirms a role of G-quadruplex structures in controlling gene expression in physiological conditions (Brooks and Hurley, 2010). This is due to the inability of proteins involved in the translation processes to bind to G-quadruplex and to carry on transcription process (Fig. 1.15).

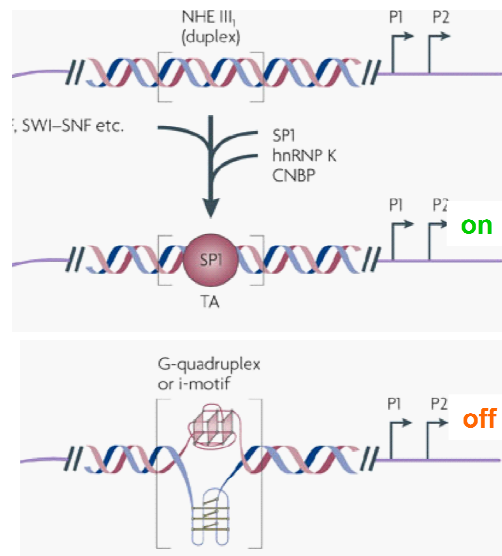


Fig. 1.15: Mechanism of the inhibition of gene expression in eukaryotic cells by G-quadruplex folding (Daekyu Sun and Laurence H. Hurley, 2009).

Despite the large number of putative G4 folding sequences identified along the human genome, only few of them have been structurally characterized in solution (NMR) or by X-ray crystallography (fig. 1.16). Some examples are *c-myc*, *c-kit*, *KRAS* and *VEGF*.

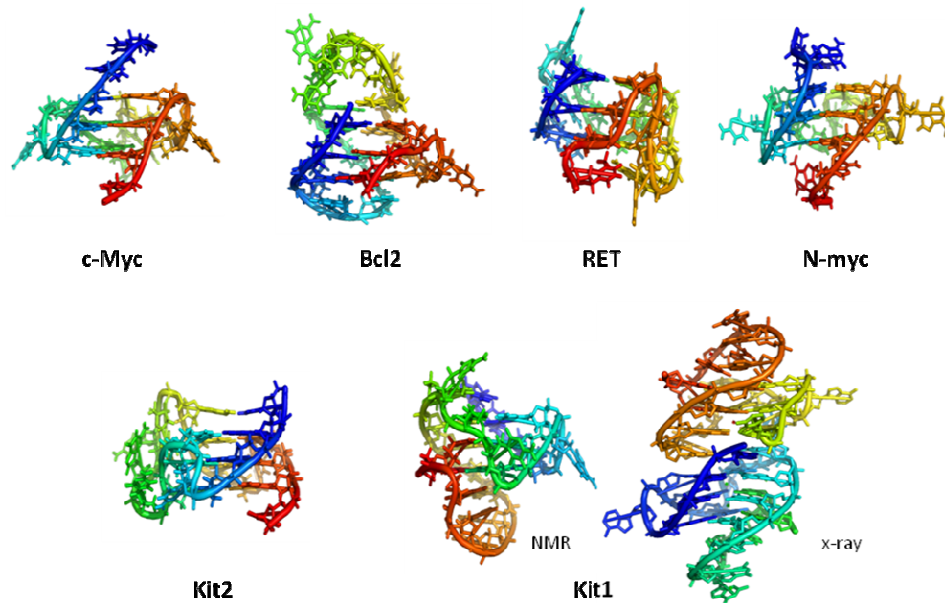


Fig. 1.16: NMR and X-ray representation of the G4 of several oncogenes

Again a variegate landscape of tridimensional structures emerged.

1.1.4 The effect of cosolutes on G4 equilibria

The accurate structural characterization of the G4 structures under physiological conditions is a very important starting point for the design of selective ligands. Data on G4 structures obtained so far, refer to NMR and crystallographic studies which can implement each other and although sometimes are even discordant. For example, in 2002 Parkinson et al. published the crystal structure of the telomeric sequence as a mainly parallel folded (Parkinson et al., 2002), but three years later, Chaires et al., through NMR studies, showed that it was not the main conformation assumed in solution (Li et al., 2005). Thus they outlined the potential existence of an antiparallel structure, too. Later on, NMR studies underlined the presence of hybrid structures according to the sequence length.

Nowadays, cosolvents as molecular crowding and dehydrating agents have been extensively used to reproduce the physiological conditions (Lannan et al., 2012; Petraccone et al., 2012a; Petraccone et al., 2012b). Indeed, molecular “crowding” and dehydrating agents such as polyethylen glycol (PEG), Ficoll, sucrose, acetonitrile and ethanol, simulate the intracellular environment which is crowded with biomolecules that occupy a significant fraction (up to 40%) of the cellular volume, with a total concentration in the range 300–400 mg/ml.

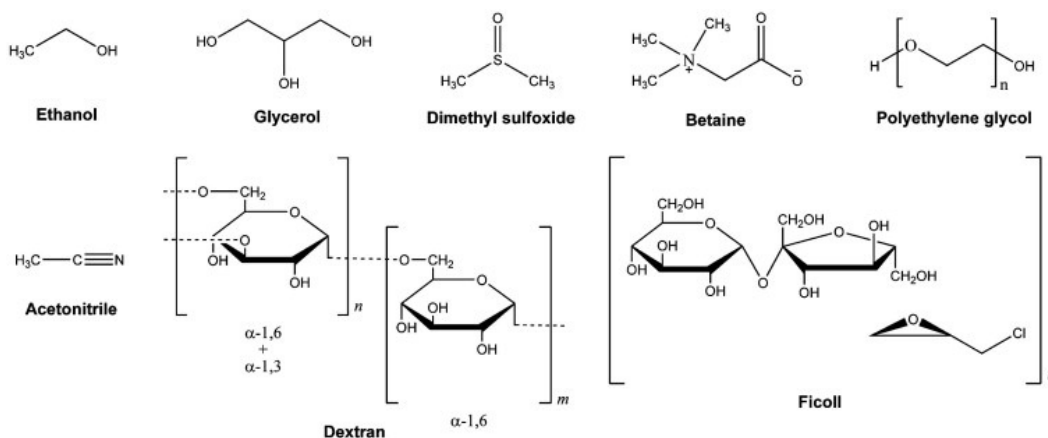


Fig. 1.17: examples of cosolutes commonly employed to mimic intracellular “crowded” environment.

Cosolutes significantly affect the structural and thermodynamic equilibria of G4 as they increase the excluded volume and they alter water activity and molecules hydration, thus affecting any equilibrium involving changes in bound water molecules upon complex formation. Depending on the cosolute nature, they can act as crowding and/or decrease water activity. For example, small cosolutes (ethanol, acetonitrile, DMSO, betaine) at high molar fractions, greatly reduce the water activity and act mainly as dehydrating agents, whereas larger molecules such as water soluble proteins or polysaccharides (BSA, hemoglobin, lysozyme, ficoll, dextrans, PEG) generate a significant excluded volume (crowding) also at low molar concentration and have a small effect on water activity. As regards G4 structural studies, the most studied “crowding” solute is the highly water-soluble synthetic polymer polyethylene glycol (PEG), usually with low molecular weight (PEG 200 and PEG 400). However, recent evidences outlined that PEG interacts with macromolecules rather than being an inert agent thus contributing significantly to lower the water activity by mainly acting as a dehydrating rather than steric crowding agent (Petraccone et al., 2012b).

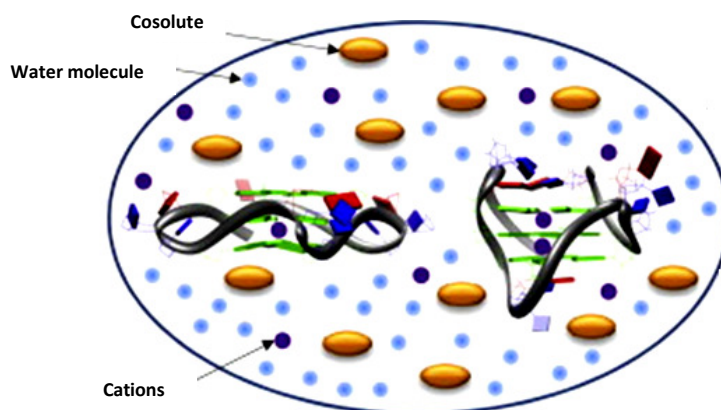


Fig. 1.18: representation of dehydration and crowding conditions generated by cosolutes affecting G4 folding.

Several studies outlined that molecular crowded conditions simulated by 40% PEG₂₀₀ (w/v) switch the double strand vs G4 equilibrium towards the G4 conformation, in particular towards a parallel G4 structure (fig. 1.19).

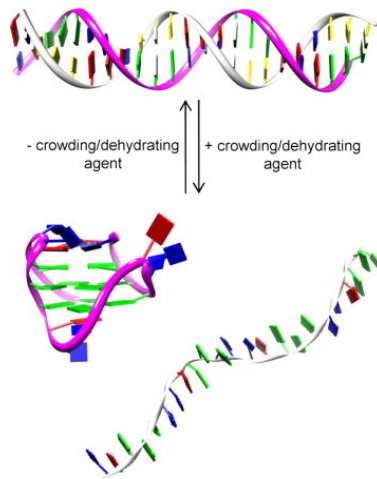


Fig. 1.19: shift of the double strand vs quadruplex equilibrium induced by a crowding/dehydrating agent.

1.6 Targeting G4 structures

The many experimental evidences on G4 existence in cells and on its multiple role *in vivo* indicate G4 as a novel target potentially leading to novel strategies with a wide therapeutic potential (Balasubramanian et al., 2011; Brooks and Hurley, 2009; Brooks and Hurley, 2010; Brown et al., 2011). The recognition of the G4 tridimensional fold by a small molecule might replace the nucleic acid-based approaches by overcoming the poor availability of this therapy (Alcaro et al., 2013; Murat et al., 2011). However so far, a G4 binding drug has not been developed yet, but modern pharmaceutical chemistry is focusing its attention on this topic. The ability to bind DNA is the first assumption to design putative G4 binders. The goal would be the development of a chemical/biological agent, which could discriminate between different folding of DNA in particular between G4 and the dsDNA that is overexpressed inside the cell. A model lead compound for G4 recognition can be considered the telomerase inhibitor telomestatin, isolated from *Streptomyces anulatus*. The solution structure of the complex quadruplex-telomestatin, shows the interaction mode of the compound with the G4. In particular, the telomestatin molecule stabilizes the folding by stacking on the G-tetrad (Chung et al., 2014).

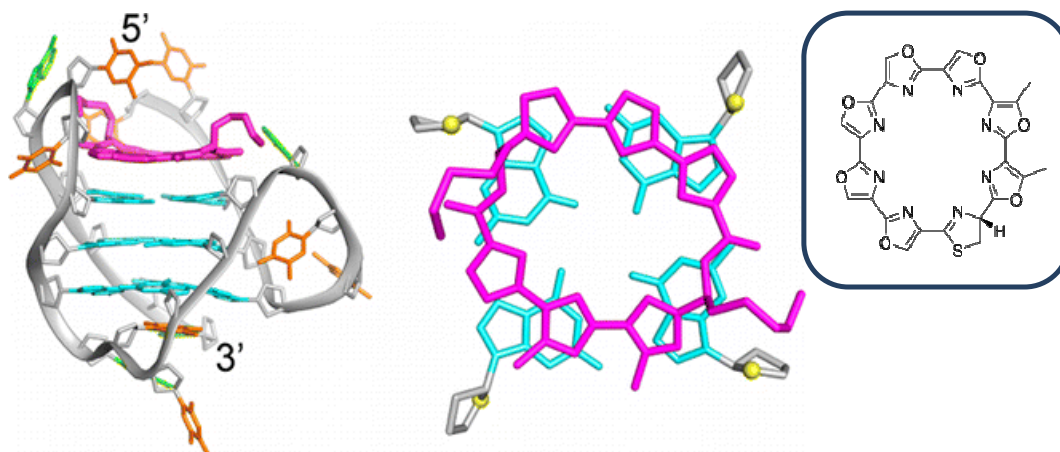


Fig. 1.20: G4 stabilization by telomestatin molecule (Chung et al., 2014).

In the last decades, several G4 stabilizer compounds have been screened, among them, porphyrines, anthraquinones and naphthalene diimides derivatives provided promising results (Collie et al., 2012; Di Antonio et al., 2007; Gatto B., Zagotto G., Sissi C., Cera C., Uriarte E., Palu` G., Capranico G., and Palumbo M.; Grand et al., 2002). As an example, the porphyrine TMPyP4 is the most studied G4 binding agent until now and its ability to discriminate between double stranded DNA and G4 has been confirmed many times (Grand et al., 2002). Though, solubility issues, as well as problems related to the permeation into the cell nucleus and the lack of significant binding specificity *in vivo* did not allow it to overcome pre-clinical trials.

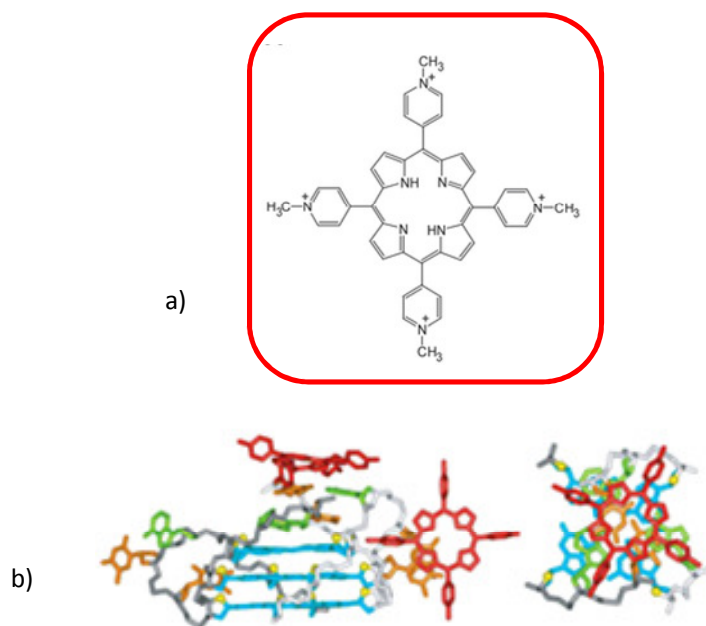


Fig. 1.21: molecular structure of the porphyrin TMPyP4 (a) and X-ray structure of the compound (red) bound to the d(TAG3T2AG3) bimolecular G-quadruplex (Patel et al., 2007).

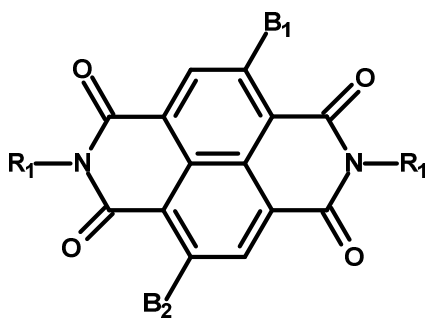


Fig. 1.22: chemical structure of a tetrasubstituted naphthalene diimide derivative.

As regards anthraquinones and naphthalene diimides scaffolds, they have been considered as suitable starting point for G4 binding ligands thanks to their chemical accessibility and large planar surfaces (Collie G.W., Promontorio R., Hampel S.M., Micco M., Neidle S., and Parkinson G.N.; Milelli et al., 2012; Nadai et al., 2011). These classes of derivatives are able to efficiently bind to double strand DNA thanks to their planarity, which allows an intercalation between base pairs of DNA in the B conformation. However, their chemical and biological properties are greatly affected by the different substituents on the planar ring system. Side chains are usually introduced with one or two positive charges, in order to establish an electrostatic interaction with the phosphate backbone of the polynucleotide. The volume and position of side chain substituents balance the preferential recognition of G-quadruplex versus double strand structure. Despite their distinct chemical features, all these derivatives share a general binding mode on G4. In particular, they are hypothesized to interact with the G-quadruplex through π -stacking and electrostatic interactions. This is first of all, due to their aromatic surface and planarity (Sissi and Palumbo, 2014) (Zagotto et al., 2008a; Zagotto et al., 2008b; Zagotto et al., 2011)

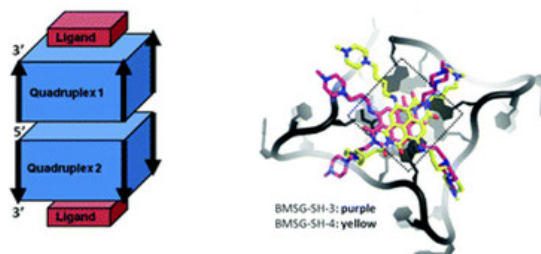


Fig. 1.23: representation of the binding of two tetrasubstituted naphthalene diimide (BMSG-SH-3 and BMSG-SH-4) derivatives on a G4 tetrad.

However, other binding modes have been proposed for G-quadruplex ligands, including binding into the grooves(Jain and Bhattacharya, 2010), electrostatic interactions with the loops and with negatively charged phosphate groups of the backbone(Sissi and Palumbo, 2014). In particular, G-quadruplex groove targeting is a very attractive strategy that can lead to an increased selectivity of recognition between duplex and quadruplex DNA. Some example of compounds that fit these regions are reported in fig. 1.24.

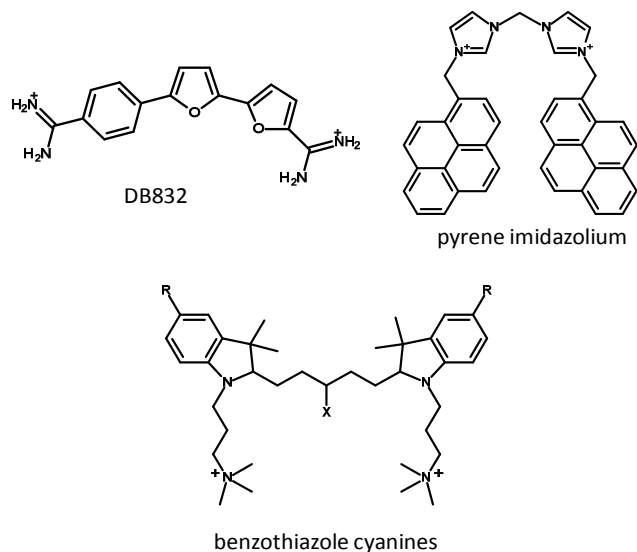


Fig. 1.24: examples of G4 groove binders(Sissi and Palumbo, 2014).

Quadruplex grooves are very different from double helix grooves for shape and for dimensions. Moreover, due to the high polymorphism of the structure, quadruplex

grooves can vary significantly according to the type of sequence, the nature of cations and the syn/anti conformation of guanine residues resulting in an ensemble of possible groove geometries. Selective targeting of the different grooves will allow the recognition of quadruplex DNA with a high degree of selectivity(Sissi and Palumbo, 2014).

Thus, this open the chance to design derivatives that can discriminate among G4 of different topologies and it is valuable for the development of new potent anticancer drugs.

1.7 I-motif structure

Distinctly from the telomeres end, G-rich regions in oncogene promoters are paired to their complementary C-rich strand. Under particular conditions also these sequences can fold, , into a peculiar non-canonical secondary structure called i-motif(Day HA, Pavlou P, Waller ZA, 2014). This is a four-stranded DNA secondary structure comprised of two parallel-stranded DNA duplexes held together in an antiparallel orientation by intercalated, cytosine–cytosine⁺ base pairs (fig. 1.25). This conformation requires hemiprotonated cytosines thus it preferentially occurs at low pH (4-5) due to the pKa of N3 of cytosines (~4.5).

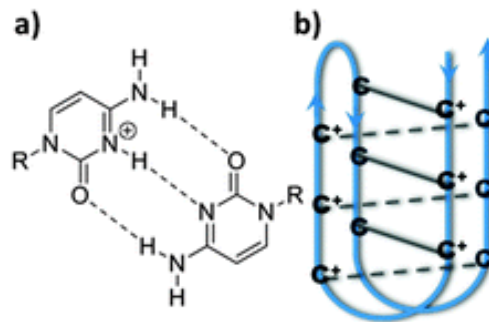


Fig. 1.25: (a) A hemiprotonated cytosine–cytosine base pair which intercalate to form the i-motif structure. (b) A schematic representation of i-motif structure(Day HA, Pavlou P, Waller ZA, 2014).

No physiological evidence of iMs has been shown yet, but these structures are now widely employed as pH switches in nanotechnology devices. However, as previously outlined, the folding into i-motif can reinforce the switching off of transcription processes mediated by non-canonical G4 secondary structures, thus leading to efficient cancer arrest. For this reason, several studies are now facing the topic of the occurrence of iMs at pH values, closer to the physiological one (pH 7.5). For example, the presence of a crowding agent such as PEG₂₀₀ at 40% w/w induces the iMs formation at slight higher pH values(Bhavsar-Jog et al., 2014). Furthermore, the presence of silver cation (Ag⁺) has been showed to induce i-motif folding at room

temperature and physiological pH and that folding is reversible with the addition of cysteine(Day et al.).

Actually, the study of the best conditions at which this structure exist, could be a starting point for the design of selective ligands for gene silencing and the possibility for this structure to assume an important role in anticancer therapy as G4 does.

A considerable breakthrough in targeting iM with small molecules regards BCL2 oncogene. Indeed, it was shown that two molecules, IMC-48 and IMC-76, act with antagonistic effects on the formation of the complex of i-motif with the activating transcription factor (hnRNP LL), thus controlling BCL2 expression. IMC-48 stabilizes the complex by activating transcription, whereas IMC-76 stabilizes the hairpin structure of the C-rich strand by hindering the process. This was the first evidence in which i-motif was exploited as controlling element for oncogene expression through the use of small molecules(Kang et al., 2014).

1.2 AIM OF THE STUDY

An *in silico* prediction tool (<http://www.quadruplex.org>) disclosed the existence of several G4 putative sequences among the upstream of the promotorial start sites of other several oncogenes. Among these, along BRAF and EGFR oncogenes promoter site . BRAF encodes for a serine-threonine protein kinase. Activating mutations in the BRAF oncogene (BRAFFV600E) are seen in about 70% of primary melanomas, 10% of colorectal cancers and some 30-70% of papillary thyroid carcinoma. Specific BRAF kinase inhibitors are undergoing rapid clinical development and promising data on efficacy have been demonstrated. However, clinical responses to the highly selective small molecule inhibitor of the BRAF (V600E) oncoprotein, PLX4032, differ widely, ranging from a response rate of approximately 80% in melanoma to only 5% in BRAF mutant colorectal cancer. It has been demonstrated that the unresponsiveness of colon cancer to BRAF (V600E) inhibition relies on a rapid feedback activation of EGFR. EGFR is the epidermal growth factor receptor. Upon ligand binding it undergoes to dimerization which activates its tyrosine kinases activity leading to a complex signal transduction cascade that modulates cell proliferation, survival, adhesion, migration and differentiation. Several evidences correlate its overexpression with different types of cancer(Goffin and Zbuk, 2013).

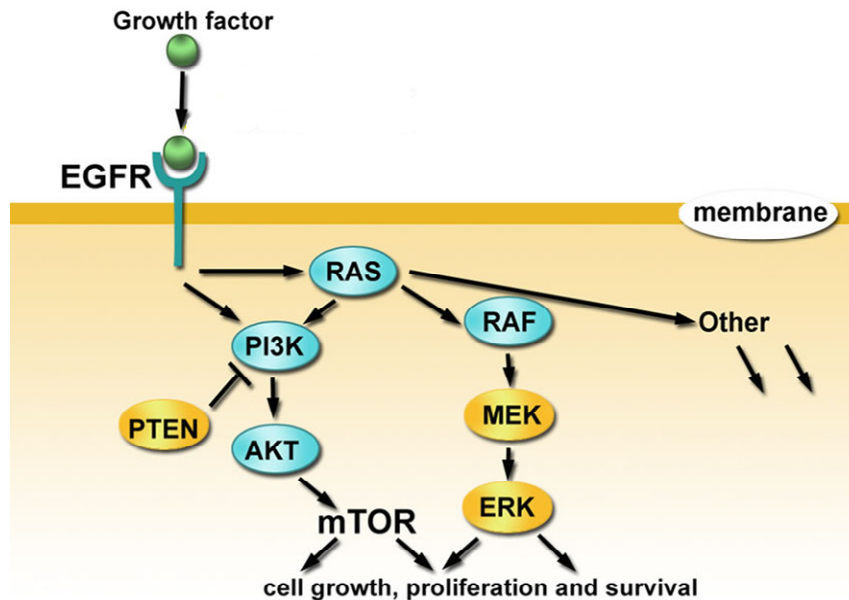


Fig. 1.26: effect of the feedback activation of EGFR (Epidermal Growth Factor) by mutation of RAF protein(Sun and Yang, 2010).

Nowadays, therapeutic resources to counteract EGFR overexpression concern, for example, humanized monoclonal antibodies against the extracellular domain of growth factor receptors or ATP mimetics which compete with ATP for binding to the receptor kinase pocket, thus disabling signal transduction (Arteaga, 2001).

The structural characterization of the promoter regions of EGFR and BRAF oncogenes have been carried on in this project. A prediction software allowed to identify two G-rich sequences at position -37 and -272 from the transcription start site of EGFR of 26 and 30 bases respectively. No data regardless the structural characterization of these sequences are reported so far. Besides EGFR we then explored also BRAF oncogene and we identified a 30 bases (rich of 27 guanines) G4 putative forming sequence at position -176 from transcription start site of BRAF promoter. Another G-rich sequence along BRAF promoter have already been identified and characterized. Indeed, the crystal structure of the 5'UTR coding region of BRAF oncogene has been characterized as an intertwined dimeric G4 (fig. 1.27)(Wei et al., 2013).

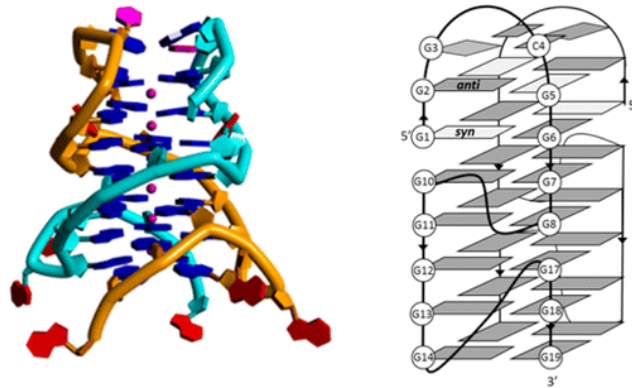
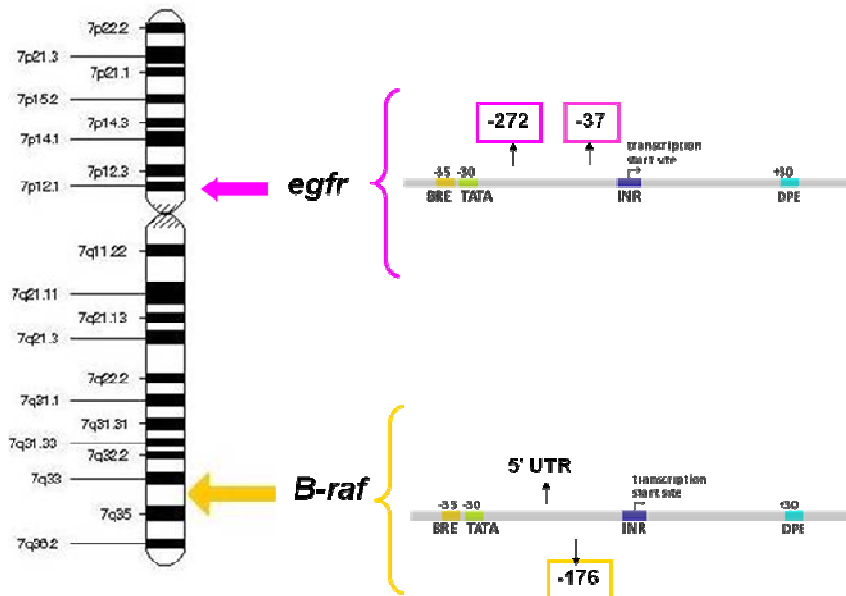


Fig. 1.27: crystal structure of 5'UTR coding region of BRAF oncogene.

Our main goal is to first structurally characterize the above mentioned sequences belonging to EGFR and BRAF oncogene.



EGFR-272: 5'-GGG GAC CGG GTC CAG AGG GGC AGT GCT GGG-3'

EGFR-37: 5'-GGG GAG GCA GGG CGG GAG GAG GAG GG-3'

BRAF-176: 5'-GGG GGT GCG GGG GGG AGC GGG GGA AGG GGG-3'

Fig. 1.28: localization and representation of EGFR and BRAF G4 forming regions, object of the study.

As previously mentioned, the reported sequences were not previously characterized, for this reason beside the structural investigation, we also explored the conditions which best allow the folding and which best fit with the physiological ones. In this contest, since oncogenes are double stranded sequences it was interesting to investigate also the folding of the complementary cytosine rich strand into i-motif as it could also be implicated in switching off oncogene transcription. Furthermore, the G4/i-motif folding vs the double strand has been studied in order to identify the prevalent conformation when at physiological conditions.

1.3 RESULTS AND DISCUSSIONS

1.3.1 G-quadruplex characterization

1.3.1.1 EGFR-272

In order to assess the potential of EGFR-272 to fold into a defined secondary structure, the effect of KCl as a G-quadruplex stabilizer was investigated.

In this context, the Electrophoretic Mobility Shift Assay (EMSA)(Sun and Hurley, 2010) is a commonly used technique to observe the formation of species characterized by different folding states. DNA is a negative charged molecule, so it always migrates towards the positive electrode. However, DNA molecules present the same mass/charge ratio irrespectively of their length. Thus, when they are subjected to an electric field, the migration on a polyacrylamide native gel is largely influenced by their hydrodynamic volume. In particular, species with higher hydrodynamic volume will migrate slower in the gel, conversely the ones with lower hydrodynamic volume will have a higher electrophoretic mobility. Considering that G4 is a more compacted form with respect the linear form, it will run faster in the polyacrylamide gel, and this is the feature which allows the recognition of the G4 formation.

The EMSA performed for EGFR-272 titrated by increasing KCl concentrations, showed the presence of two forms, which differ for their hydrodynamic volume, in particular, one of them has a higher electrophoretic mobility, thus indicating it as an intramolecular folded form (fig. 1.29).

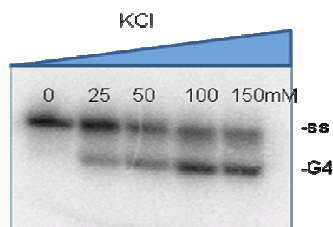


Fig.1.29: 20% acrylamide/bisacrylamide native gel (0.5% TBE, 10mM KCl, pH 7.5) of single stranded EGFR-272 folded in the presence of increasing KCl concentrations.

In order to identify the nature of the species resolved by the gel shift experiment, spectroscopic analysis were performed which consist in thermal difference spectra (TDS) and circular dichroism (CD). Nucleic acid circular dichroism provides important information on DNA conformational properties. DNA structures differ one from each other for their strands polarities, sugar conformation and nucleobases orientation around the glycosidic bond. As a result each DNA structure present a different CD spectrum. For example double stranded DNA has a dichroic signal generally characterized by one positive and one negative band centered at about 275 and 245nm, respectively. Conversely, a positive band at 260nm and a negative one at 240nm are characteristic of parallel G-quadruplex, whereas the spectra of anti-parallel G-quadruplexes have a negative band at 260nm and positive band at 290nm. Thus, CD technique allows to monitor the structural equilibria which occur on the selected G-rich sequences(Randazzo et al., 2013).

The titration of EGFR-272 sequence with increasing KCl concentrations showed that the spectrum in the absence of KCl is different from the one of the oligonucleotide titrated by the ion. In particular, the metal ion induced a positive pick at 260nm and a negative one at 240nm. This suggests the ability of this sequence to assume a parallel G4 conformation. This is the explication that the parallel G4 folding is promoted the presence of the potassium ion. The plotting of the molar ellipticity obtained at 265nm for increasing KCl concentration, showed that saturation is reached at about 200mM KCl (fig. 1.30). By fitting data with one site saturation binding equation ($y=B_{max} * x / (K_d + x)$) a dissociation constant of $28,46 \pm 4,19$ mM towards KCl was estimated. The final DNA conformation showed a melting temperature of 58 °C when saturated with 200mM KCl moreover, it can be recovered after a melting-annealing step, thus confirming the reversibility of the process.

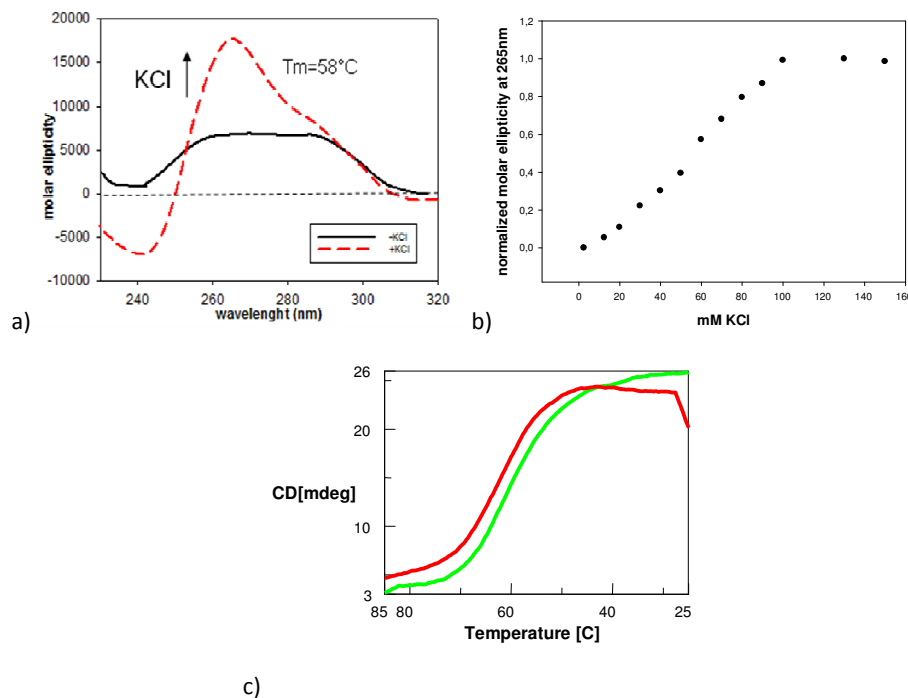


Fig. 1.30: dichroic spectra of EGFR-272 in the absence of KCl and in the presence of 200mM KCl (saturation) (a); variation of molar ellipticity in function of increasing KCl concentrations (b); melting (green) and annealing (red) curves of EGFR-272 after KCl titration (c).

Also the thermal difference spectrum (TDS)(Mergny et al., 2005) is unique for each type of nucleic acid structure as it reflects the base stacking interactions that occur uniquely within each type of conformation. It consists in the difference of the UV spectrum of the denatured oligonucleotide at 95 °C and the one of the folded oligonucleotide at 25 °C. The absorbance spectra of nucleic acid bases are deceptively simple: they actually correspond to a number of different $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions. At 295nm, hypochromicity (negative TDS pick) at 295nm, due to the changes in light absorption occurring upon denaturation, is attributed to the interaction between the dipoles induced in the chromophore by the light. This might be the result of the significant contribution of $n \rightarrow \pi^*$ transition moments at this wavelength when stacking interactions are lost due to denaturation. However, Although the dissociation of all G-quadruplexes analyzed so far show an inverted transition at 295 nm, this property is not unambiguously indicative of G-quadruplexes. Indeed, Z-DNA, i-motif, Hoogsteen

duplexes, and pyrimidine triplex formation also leads to this effect. It is for this reason that the global shape rather than the wavelength of the peak is specific for each structure (Mergny et al., 2005; Mergny and Lacroix, 2009)

Table 1I summarizes the different structures which involve a negative differential absorbance at 295 nm and fig 1.31 the characteristic TDS shape of G4 assumed by the telomeric sequence in KCl.

Structure	Major positive peak	Negative peaks
Z-DNA	241 ± 2 nm	295 ± 1 nm (-0.37)
Triplexes		
Triplex TC	247 ± 1 nm	295.5 ± 1 nm (-0.30)
Quadruplexes		
i-motif	239 ± 1 nm	294.5 nm (-0.60)
G-quartet	243 ± 2 nm	295 ± 1 nm (-0.73)
	273 ± 1 nm	

Table 1I: attribution of characteristic picks of TDS to peculiar DNA folding

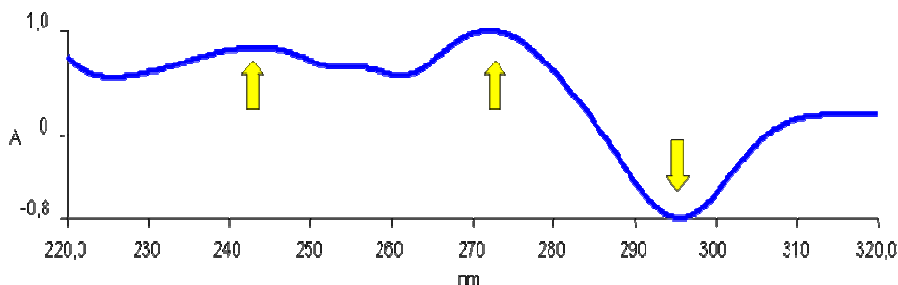


Fig. 1.31: TDS of the G-quadruplex structure of the telomeric sequence in KCl, arrows outline the characteristic hyperchromisms and hypochromism (positive and negative picks) of the G4 folding.

The TDS shape of more than 200 oligonucleotidic sequences have been compared each other so that the TDS shape of a G-quadruplex and i-motif conformation have been outlined (Mergny et al., 2005; Mergny and Lacroix, 2009). As reported in literature (Mergny et al., 2005) a G-quadruplex TDS spectra shows two positive picks at 243 (±2)nm and 273 (±1) nm, and a negative one at 295(±1) nm at -0.73 AU , but its shape is highly sequence dependent.

TDS performed on EGFR-272 confirmed CD data. Its shape is different from the TDS of the telomeric G4, but considering that TDS is sequence dependent, the negative pick at about 295-300nm and the two positive ones at about 245 and 270nm (fig. 1.32) assess the assumption of the G-quadruplex conformation when oligonucleotide is folded at physiological KCl concentrations (150mM).

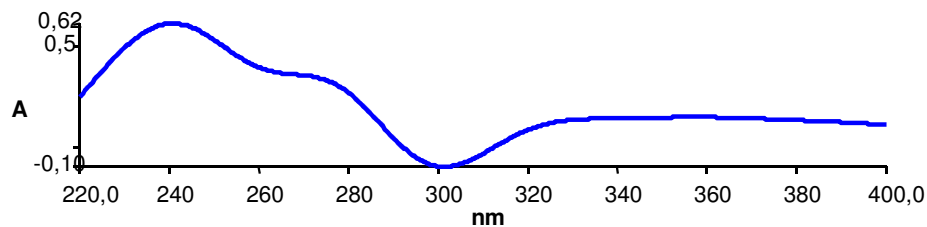


Fig. 1.32: TDS of the G-rich EGFR-272 strand recorded at 150mM KCl concentrations.

Thus, from data obtained so far, we can conclude that the G-rich tract of EGFR-272 assumes a stable intramolecular folded form, attributed to a parallel G4 conformation. Its stability at 150mM KCl, suggests its occurrence at physiological conditions.

1.3.1.2 EGFR-37

The sequence EGFR-37 showed a peculiar folding behavior depending upon its concentration in solution. Indeed, when the oligonucleotide is at mM concentration, the dichroic spectrum has a high positive signal at 260nm and a negative one at 240nm. In these conditions, KCl further stabilizes the dichroic signal without altering the overall shape of the spectrum (fig. 1.33). The resulting species is thermally stable with a of melting temperature >90 °C (fig. 1.34).

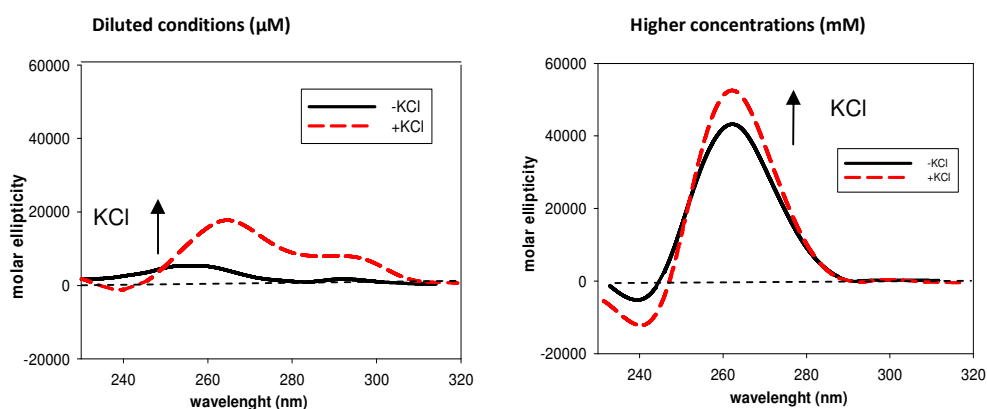


Fig. 1.33: CD spectra of EGFR-37 recorded at μM and mM oligonucleotide concentrations in the presence and in the absence of KCl

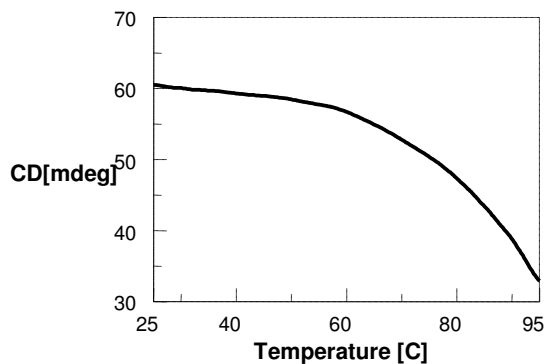


Fig. 1.34: melting curve of EGFR-37 at mM concentration after KCl titration (50mM).

However, mM DNA strand concentrations are not compatible with physiological conditions as the oligonucleotide in cells is at very lower concentrations. For this reason, also the dichroic spectrum at μM strand concentration was acquired and oligonucleotide was titrated with KCl. In this case, the oligonucleotide in the absence of KCl didn't show a defined spectrum, suggesting that the sequence is not significantly folded. Conversely, the spectrum obtained after KCl titration suggested the conversion into a more structured form and it showed two positive picks; one at 295 and one at 260nm, only with a modest negative one at 240nm. This may suggest the presence of different G4 conformations, in particular of parallel and antiparallel form (fig. 1.33). Indeed the variation of the 260 vs 290 nm contribution was differently regulated by KCl concentration. In particular, the conformation at 260nm was induced only at metal ion concentrations higher than the physiological ones (fig. 1.35). The affinity constant towards KCl was also estimated at the previous mentioned different experimental conditions and data are summarized in table 11I.

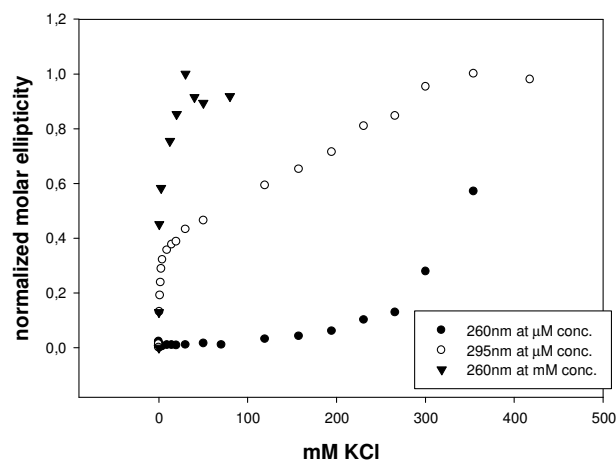


Fig. 1.35: normalized molar ellipticity in function of KCl increasing concentrations. Data refers to the oligonucleotide at mM and μM concentrations at 260 and 295nm.

	Kd (mM)
260nm at mM conc.	0,7042 ± 0,2279
295nm at μM concentration	19,41 ± 4,32
260nm at μM concentration	>300

Table 1II: Kd value obtained from oligonucleotide KCl titration at mM and μ M concentrations.

Thus, CD data preliminary suggest that the selected sequence is able to fold into G4 but that its topology is a function of strand concentration. These first hypothesis was supported by gel shift and thermal-UV experiments. The gel shift outlined the formation of multimeric species at high concentrations oligonucleotide (fig. 1.36). Consistently, the thermal difference spectrum could not be attributed to a unique G-quadruplex conformation (1.37b). Conversely, as it emerges from the gel shift, samples prepared at μ M concentration, in KCl migrates only as a unique intramolecular folded form (fig. 1.36). Taking into account CD and TDS results, it can be identified as a G4 folding, in particular into a mixed antiparallel parallel G-quadruplex structure (1.37a).

Focusing on the experimental conditions which comprise the diluted oligonucleotide and physiological KCl concentrations (150mM), the high melting temperature ($T_m \sim 74^\circ\text{C}$) of the G4 conformation at 295nm further suggests EGFR-37 G4 as a possible species in the intracellular environment.

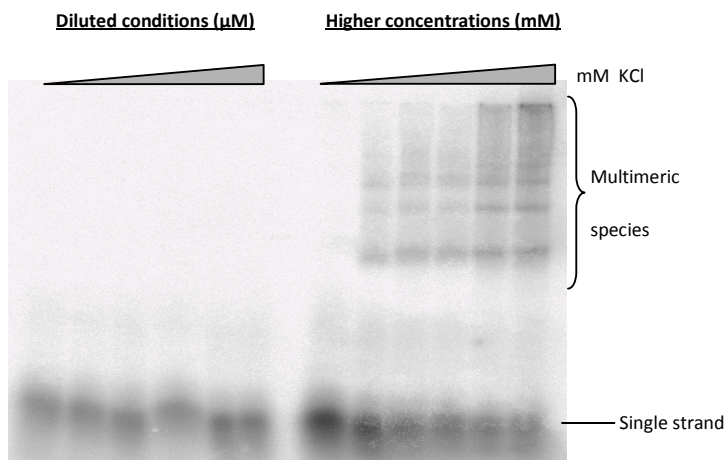


Fig. 1.36: gel shift assay of EGFR-37 performed in diluted conditions and at higher concentrations oligonucleotide at increasing KCl concentrations (0-5-10-50-100-200 mM).

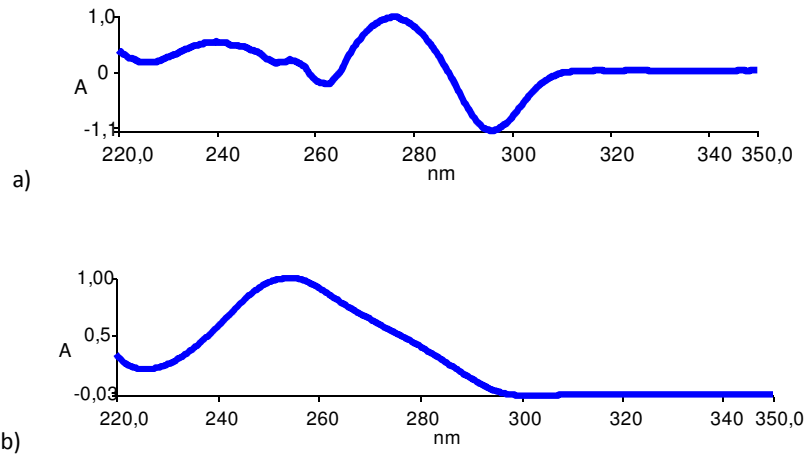


Fig. 1.37: TDS of EGFR-37 in KCl at diluted (a) and higher (b) concentrations.

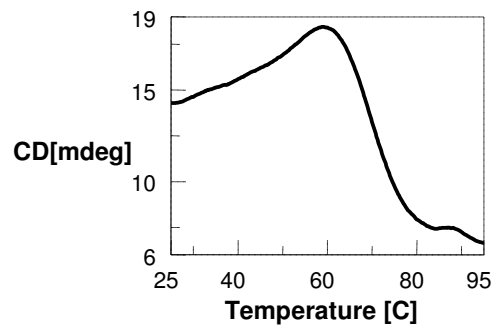


Fig. 1.38: Melting curve of the diluted oligonucleotide after KCl titration.

1.3.1.3 BRAF-176

The 30 bases promotorial sequence of BRAF oncogene is characterized by 24 guanines which can easily lead to the potential formation of multiple G4 structures. PAGE analysis confirmed that the presence of KCl causes the samples to migrate faster than a linear oligonucleotide of comparable length. To monitor the associated structural equilibria the conformational changes associated to the folding were analyzed by CD spectroscopy in the presence/absence of K^+ . Titration curves showed that the metal ion altered the dichroic spectrum and that the process reached saturation at low K^+ concentration (about 10mM) with an apparent $K_d=0.64 \pm 0.03$ mM. The extent of such a variation was modest but led to a final form identified by a strong positive band at 295 nm and a modest negative contribution at 248 nm. The melting temperature of the folded form was clearly directly related to the presence of the metal ion reaching values higher than 90 °C at physiological concentrations of K^+ (> 100 mM). A good correlation between the thermal stabilization driven by the metal ion and the variation of the CD signal it promotes, was found. These evidences allowed us to propose also that this sequence is actually able to fold into a G-quadruplex structure, likely arranged into a main antiparallel folding deriving from overlapping of guanosine of alternating syn/anti glycosidic bond. This model was in line with TDS results.

Nevertheless, the potential folding of BRAF-176 into G-quadruplexes of different topologies was not fully excluded. Indeed, working at KCl concentrations higher than 50 mM two thermal transitions were observed. Additionally, a fast annealing step allowed to trap a CD signal characterized by a strong positive band centered at 265 nm. Since along the time, it slowly converts to the above described form, we can indicate it as a kinetically but not thermodynamically favored form (fig. 1.39).

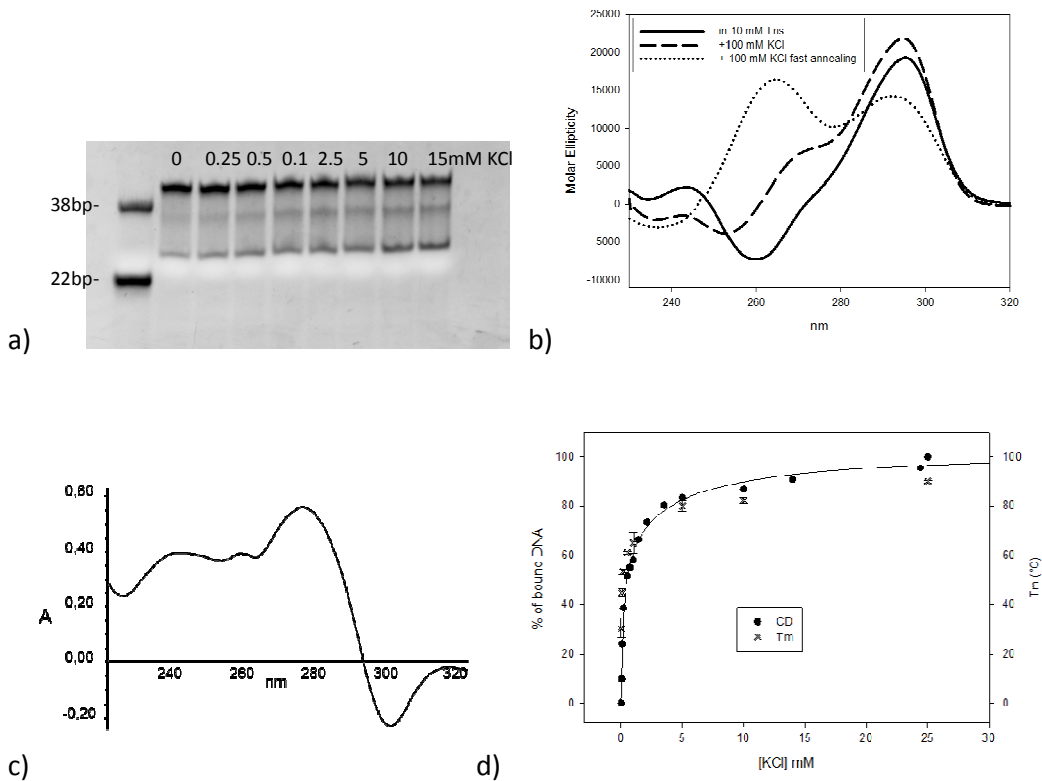


Fig. 1.39: EMSA of BRAF-176 folded in the presence of increasing KCl concentrations (0-150mM) (a); CD spectra of 4 μ M BraF-176 recorded in the absence (solid lines) or in the presence of 100mM KCl in 10 mM Tris, 1 mM EDTA, pH 7.0 after a slow (dashed lines) or a fast annealing step (dotted lines) (b); the thermal difference spectrum corresponding to the slow annealed sample is reported in c. Melting temperature and % of bound DNA at increasing KCl concentrations (0-30mM) is reported in d.

To further explore G4 formation, we included in our analysis a 40% of polyethylene glycol (200 w/v), too. Indeed, it is reported in literature that this polymer acts as a crowding agent, and thus it mimics the high nucleic acid and proteins concentration typical of the cell nucleus. Moreover, it is known that PEG₂₀₀ causes dehydration and condensed environment. These conditions allow a highest stabilization of the tetrads in a unique parallel conformation.

The dichroic spectrum of BRAF-176 annealed in 100mM KCl was not influenced by the addition of 40% PEG₂₀₀ showing a high contribution of the antiparallel structure and a very low contribution of the parallel conformation. However, when the sequence was melted and

annealed in the presence of the crowding agent the pick at 295nm (which is related to an antiparallel structure) consistently decreased in favor of the parallel one (pick at 260nm), suggesting that PEG₂₀₀ can provide the rearrangement of the G4 structure assumed by BRAF-176 towards a parallel arrangement. This data agrees with literature which shows that also the telomeric sequence (TTAGG)_n which is hybrid type in KCl, shifts towards a parallel G4 conformation (Miyoshi and Sugimoto, 2008) in 40% PEG (fig. 1.40).

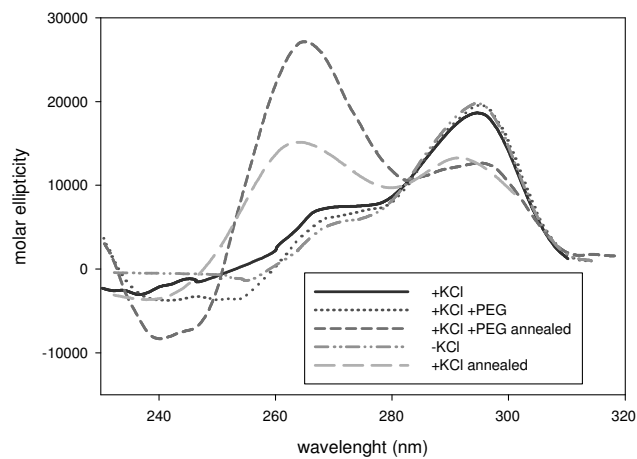


Fig. 1.40: effect of 40% of PEG₂₀₀ on the CD spectra of the G-rich BRAF-176 strand in the presence of 100mM KCl.

1.3.2 i-Motif conformational study

As previously outlined, oncogenes are double stranded sequences, so it was interesting to characterize also the folding of the C-rich complementary strand. Indeed, at particular conditions it can assume the non-canonical secondary structure called i-motif that can also be implicated in the switching off transcription mechanism. However, i-motif conformation is not favored at physiological pH because it prevalently forms at pH 5.0 where cytosines are hemiprotonated. Also i-motif has a peculiar CD spectrum: in particular, it was demonstrated for short C-rich oligonucleotide sequences that the CD spectrum of an i-motif exhibits a positive band around 285 nm and a negative band around 265 nm (Zikich et al., 2011).

Since the maximum stabilization of the i-motif is achieved under mild acidic conditions (pH 4–5), where both concentration of the protonated and the deprotonated forms of cytosines is nearly equal, besides KCl contribution, we also observed the influence of pH variations on the dichroic spectrum of the C-rich strands folded in 100mM KCl. Fig. 1.41 outlines the dependency from pH of i-motif formation.

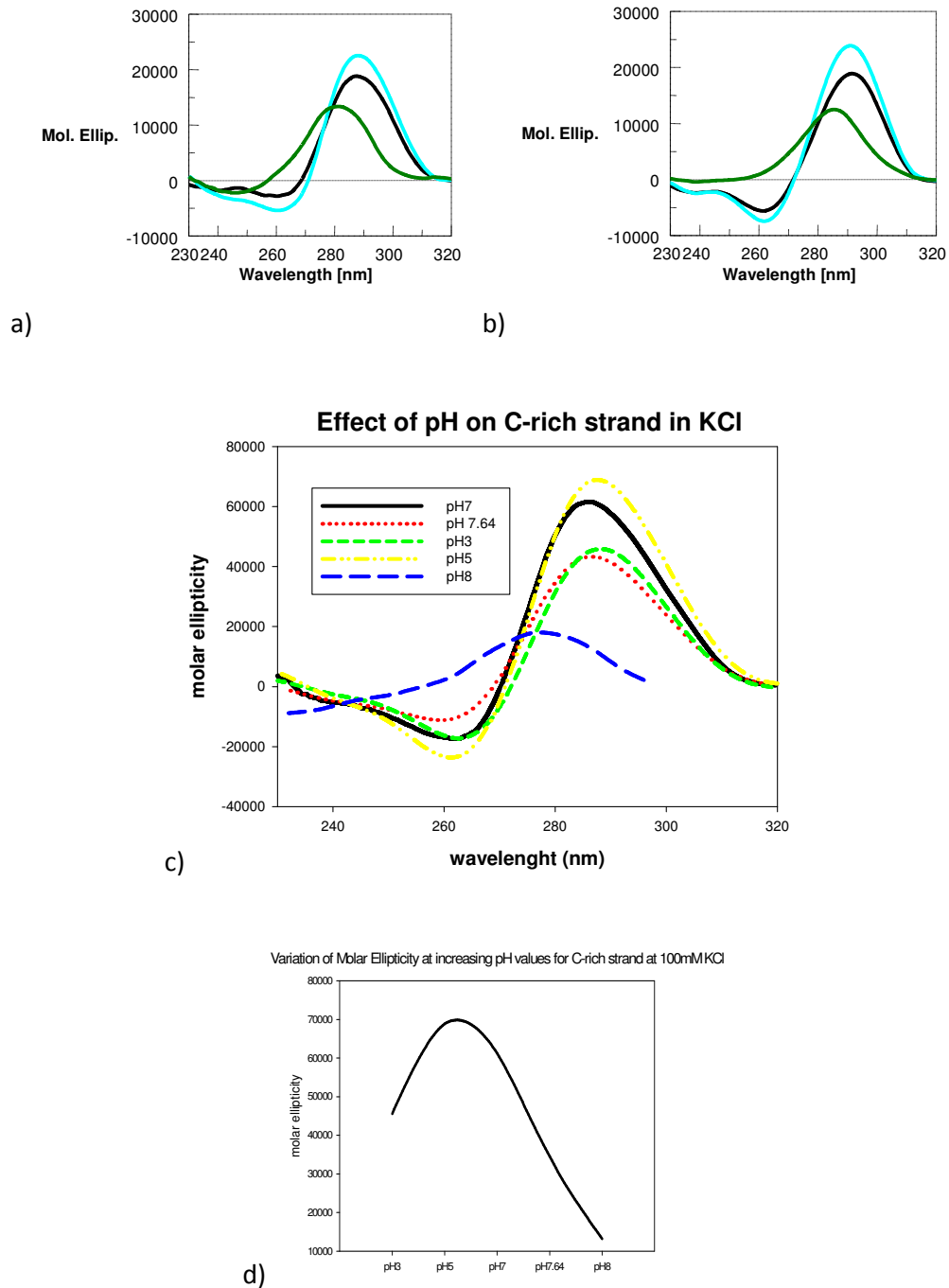


Fig. 1.41: CD spectra of EGFR-272 (a), EGFR-37 (b) C-rich strands at different pH values: 7,0 (green); 5,0 (blue); 3,0 (black) in the presence of 100mM KCl, influence of pH on the CD spectrum of BRAF-176 C-rich strand (c) and variation of molar ellipticity in function of pH (d).

As expected, the optimal folding occurred at pH 5.0, thus suggesting a remarkable dependence of the folded fraction with the protonation state. Consistently, the thermal difference spectra (TDS) recorded at pH 5,0 showed the characteristic signature of the i-motif conformation with a maximum centered at $239 \pm 1\text{nm}$ and a minimum at 295nm (fig. 1.42). These data were reproducible for all the three sequences analyzed.

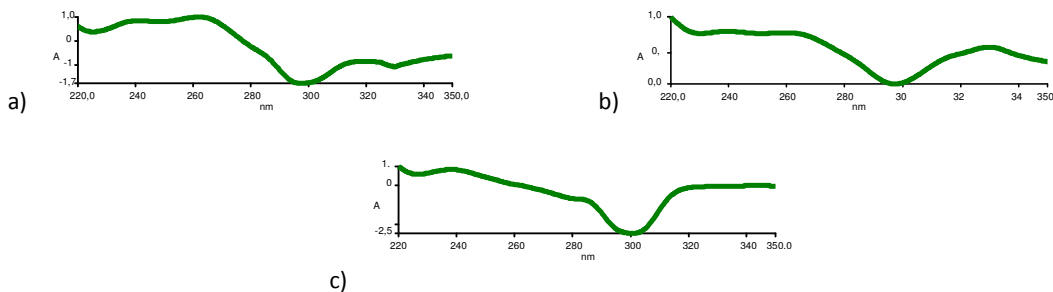


Fig. 1.42: TDS of EGFR-272 (a), EGFR-37 (b), BRAF-176 (c) C-rich strand in 10mM NaCacodilate/ 100mM KCl pH5.

Conversely, small pH variations lead to large variations of the extend of the structured form. In particular, fig. 1.41 shows that the dichroic spectra progressively decreases lowering pH up to 3. This is due to a full protonation of cytosines, which bring to higher repulsions by hindering the i-motif structuring. As expected, a more dramatic effect is observed by increasing pH up to 8.0 where all secondary structure appears to be lost. The interesting emerging data is the persistence of the i-motif at pH7.0, a condition more close to the physiological one.

As far as it concerns the role of KCl, previous studies, say that it does not have any effect in incrementing i-motif folding. Spectra acquired in the presence/absence of KCl showed that in our conditions EGFR-37 i-motif is negatively perturbed by the presence of 100mM KCl (saturating conditions) at pH7. Indeed, the molar ellipticity at 290nm is decreased indicating a reduced folded fraction. This is not the case for EGFR-272 and BRAF-176 whose i-motif structures are not significantly influenced by KCl (fig. 1.43).

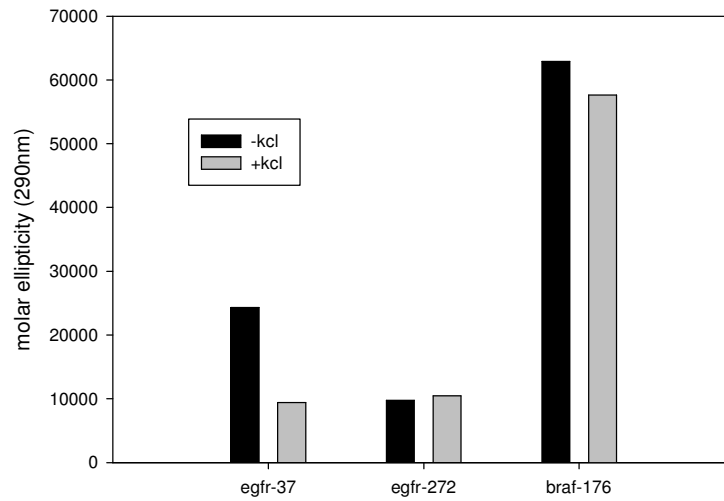


Fig. 1.43: molar ellipticity of the C-rich sequences at pH 7.0 in the presence or in the absence of 100mM KCl

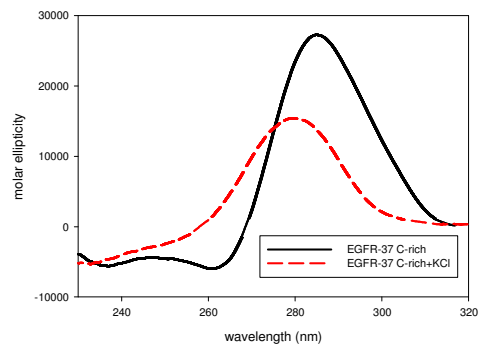


Fig. 1.44: CD spectra of c-rich EGFR-37 recorded in the presence and in the absence of KCl (100mM).

1.3.3 Effect of crowding conditions on i-motif structure

So far, the stabilization induced by crowding agents on G-quadruplex structures has been confirmed by several works but, Bhavsare-Jog et al.(Bhavsar-Jog et al., 2014), reported a study of the influence of pH, crowding agents such as polyethylene glycole (PEG), dehydrating agents and epigenetic mutations also on i-motif. Although it is known that this conformation preferentially occurs at not physiological pH values, they observed the potential of PEG to increase stabilization at slightly higher pH (about 6.5). For this reason, we considered interesting to evaluate the effect of 40% PEG₂₀₀ in inducing the i-motif at increasing pH. We focused on the sequence BRAF-176 as we observed a higher intensity of the dichroic spectrum which could suggest a easier propensity to fold into i-motif. We observed that at low pH values (5.0 and 3.0), PEG impairs the i-motif of the BRAF-176 C-rich strand. This can be probably due to inappropriate PEG/pH conditions, which lead to an increase of the electrostatic repulsion between the cytosine pairs due to the decrease of the dielectric constant for the water/PEG, mixed solvent. Actually, this is not the case at pH 7.0 where cytosines are not charged and this avoid the predominance of repulsion forces by favoring the iMs folding. Indeed, the iM spectrum, at physiological pH value and in the presence of 40% PEG₂₀₀, has the same signature of the iM at pH 5.0 in the absence of PEG₂₀₀ (fig.1.46) thus suggesting that crowding conditions, which reproduce the physiological environment, allow the occurrence of iM at slightly higher pH.

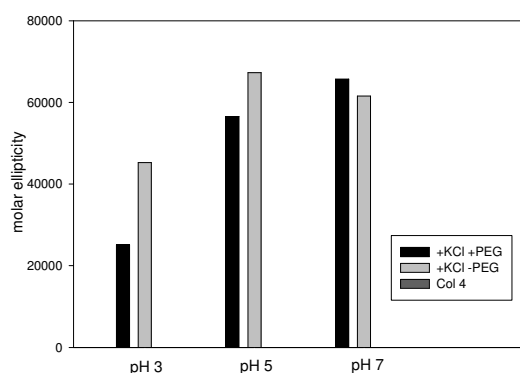


Fig. 1.45: dependence of molar ellipticity on pH and on the presence/absence of 40% PEG₂₀₀.

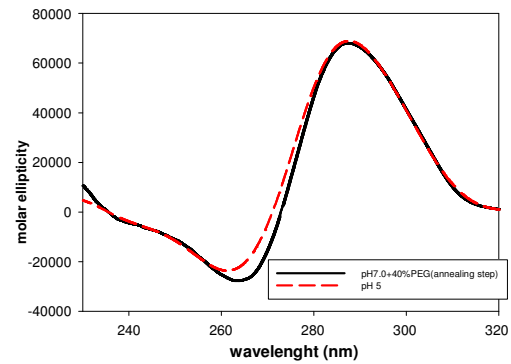


Fig. 1.46: CD of C-rich EGFR-37 recorded in 100mM KCl at pH 5.0 and at pH 7.0 in 40% PEG₂₀₀. In this latter case, samples were annealed in 40% PEG₂₀₀.

1.3.4 Double strand stability at physiological conditions

The rearrangement of both the G-rich and C-rich strands in peculiar non-canonical secondary structures has been characterized at different experimental conditions and data obtained by different techniques are consistent each other in order to identify specific foldings. So far, the isolated single stranded sequences were studied therefore, to better highlight the potential role of the herein described G4 or iM conformations in cells we investigated their formation within double stranded fragments at KCl intracellular concentrations (150mM).

Generally, the G-C pairing highly contributes to the formation of a stable double stranded sequence as it involves three hydrogen bonds with a ΔG° of -24,42 kcal/mol with respect the A-T pairing which is supported by just two hydrogen bonds (ΔG° -17,20 kcal/mol). The high content in GC pairs, in our tested sequences, allows to predict a high stability for the double stranded forms. Besides sequence itself, also ions which are physiologically present in cells can influence DNA structural equilibria. Indeed, positive ions, neutralize the negative charges of the phosphates in the DNA backbone and induce a consequent stabilization of the paired strands.

According to the metal ion nature, this effect can be more pronounced on selected DNA structural arrangements. For this reason, here we investigated the effect of KCl on double strand, if it is able to sufficiently promote G4 formation over the double strand or conversely if the stabilization of the double helix predominates.

The CD spectrum of the double helix EGFR-272 showed a positive signal at 265nm with a shoulder at 285nm and a negative pick at about 240nm. By monitoring the variations of the 265nm CD signal a melting temperature of about 71 °C in the absence of KCl was found (fig. 1.47). This is relevant since the stability of G-quadruplex EGFR-272 is lower ($T_m=58$ °C) in the same experimental conditions. Upon addition of KCl, the double strand dichroic spectrum was not perturbed by the presence of the ion.

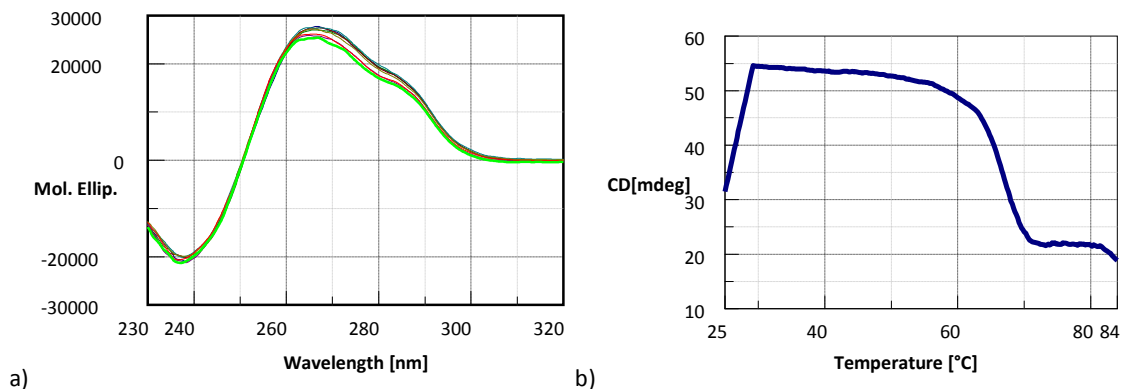


Fig. 1.47: a) KCl titration on double strand EGFR-272 (0-200mM); b) melting profile of double stranded EGFR-272 in the absence of KCl.

Comparable results were obtained on EGFR-37 and BRAF-176 double strands. CD spectra are not significantly influenced by KCl titration (fig. 1.48). Furthermore, they both show a melting temperature of 68 °C in the absence of the potassium ion which is incremented by the addition of the ion (tab. 1III). This is due to the neutralization of the phosphate negative charges of the nucleotides by incrementing double helix stability.

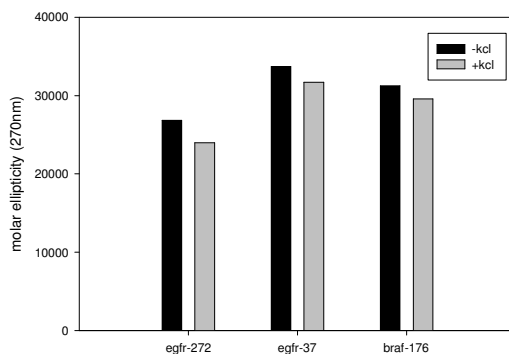


Fig. 1.48: molar ellipticity of double stranded sequences in the presence and in the absence of KCl.

Through a predictive software, the theoretical thermal stabilities of the double stranded oligonucleotides at physiological KCl concentrations were evaluated and data were compared with the experimental ones obtained recording the variation of the dichroic signal of the double strand when temperature is progressively increased. Data are resumed in table 1III from which it emerges that experimental data well agree with the predicted ones.

Double strand	EXPERIMENTAL		TEORETHICAL	
	no KCl	150mM KCl	no KCl	150mM KCl
EGFR-272 pH7	71 °C	86 °C	52.4 °C	81.6 °C
EGFR-37 pH7	68 °C	83.4 °C	52.1 °C	80.4 °C
BRAF-176 pH7	68 °C	82.4 °C	58.1 °C	86.5 °C

Table 1III: comparison of the theoretical and experimental melting temperatures of the double stranded sequences obtained in the presence (150mM) and in the absence of KCl.

In addition to the effect of KCl, the structural study on double stranded sequences has been carried on by testing other conditions. In particular, we focused on pH since iM is favored at pH 5.0.

Resuming experiments of data obtained on this aspect are the gels reported in fig.2 which refer to the EGFR-272 sequence. It emerged that by increasing the strand ratio of the G-rich versus its complementary the formation of the double strand occurs. As evidenced by the appearance of a band at low migration mobility. At physiological pH (7.5) this process is efficient. The presence of K⁺ allows to identify the excess of G-rich strand in its G4 form. However, it is completely absent in samples containing stoichiometry amounts of C-rich, thus confirming that the double strand is the preferred one. Interestingly, working at acidic conditions (pH 4.5) the folding of the cytosine rich strand into a dimeric i-motif appears. Indeed, two bands are visible in the lane containing the cytosine rich strand, in the gel of fig. 1.49. The lower band refers to the single stranded C-rich strand conversely, due to its higher electrophoretic mobility, the upper band refers to the dimeric form. This results is in line with literature data which report i-motif as a dimeric or tetrameric structure, and further support the ability of the sequence to fold into i-motif. The formation of the dimeric form is supported by KCl since in its absence it is not more detectable. However, also the dimeric iM is not sufficiently stable to impair the ds formation.

However, the dimeric i-motif species do not occur when the experiment is performed at physiological pH. In this case the G-quadruplex structure is more evident.

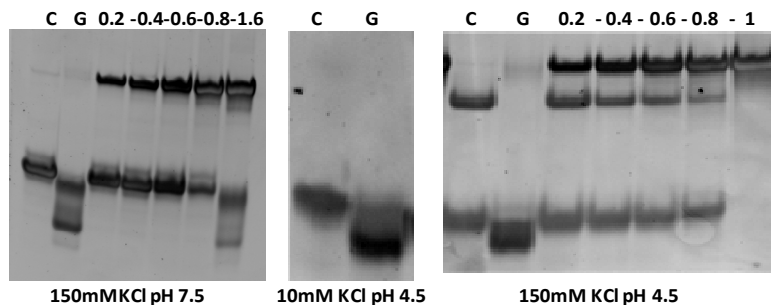


Fig. 1.49: EMSA performed at physiological pH (7.5) and at pH 4.5. C and G correspond respectively to the cytosine and guanine rich strands and 0.2, 0.4, 0.6, 0.8, 1 and 1.6 refer to the fraction of G-rich strand added to the complementary one in the reaction.

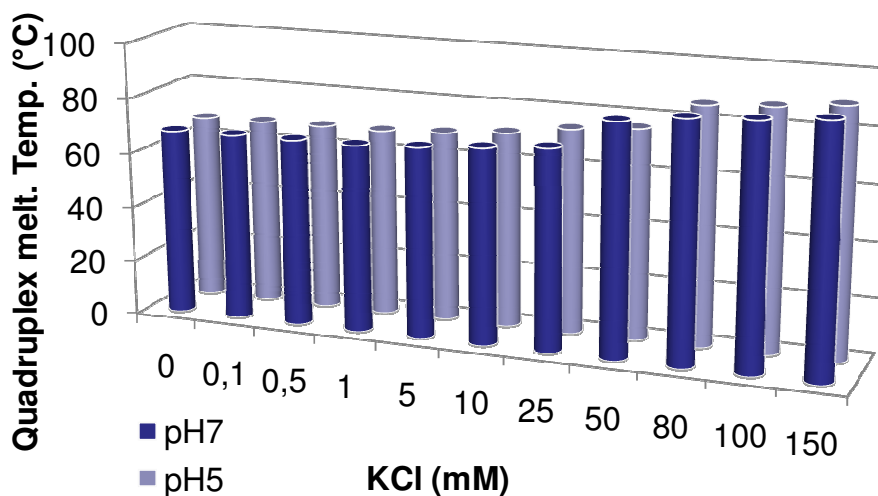


Fig. 1.50: melting temperatures of the G4 structures when the C-rich and G-rich strands are paired at different KCl concentrations and at pH 7.0 and 5.0.

The modest effect of pH variations on the double stranded EGFR-272 stability was also observed by melting assay (fig. 1.50). Data obtained perfectly agree with electrophoretic ones. Indeed, the double strand melting temperature was not significantly reduced by switching pH from 7.0 to 5.0 thus making difficult to actually relate it to iM formation or to a direct destabilization of the ds by acidic conditions. The same melting profiles were obtained also for the sequences EGFR-37 and BRAF-176.

The reduced occurrence of the G-quadruplex structure over the double stranded at pH 7.0 and 150mM KCl was confirmed by gel shift and fluorescence melting experiments. By comparing the behavior of the G-rich vs double strand EGFR-272 by a native acrylamide gel, it appears that, starting from the paired form, the formation of a second band with a higher electrophoretic mobility is observed at higher KCl concentrations and at a lower intensity (fig. 1.51).

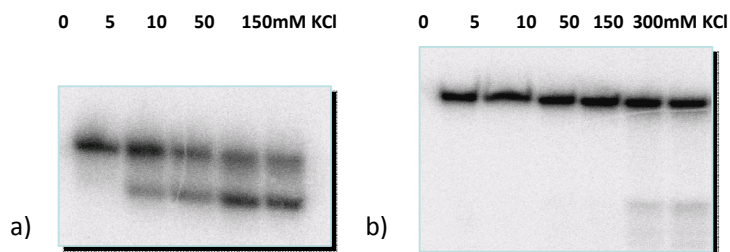


Fig. 1.51: EMSA performed on G-rich single stranded EGFR-272 (a) and on the double stranded EGFR-272.

This model was verified also on dsBRAF-176. Despite the higher affinity of BRAF for KCl, EMSA showed that addition of KCl was not sufficient to cause the strand separation required in order to form G4. However, not the same occurred when the two complementary strands were annealed in the presence of KCl: in these conditions a significant competition between the ds and G4 structures was evident as highlighted by PAGE.

For all the tested sequences, the G4 formation over the double strand was further highlighted by melting curves of the double stranded sequences in which only the labelled G-rich strand was monitored. Fluorescence changes are strictly related to oligonucleotide conformational changes. In particular, when the oligonucleotide, labeled with fluorophore and quencher at each end, is in G4 conformation, the probes are close and fluorescence is lower as its intensity depends on Förster distance. Conversely, when annealed in double strand, fluorophore and quencher are more distant and fluorescence high.

Regardless the sequence BRAF-176, starting from its double strand form, in the absence of KCl, we observed only the transition from the double strand to the single stranded form (1.52). Conversely, when KCl was included in the reaction mixture, this step was associated

to a remarkable reduction of the fluorescence signal which describes the folding of the labeled G-rich strand into a G-quadruplex structure.

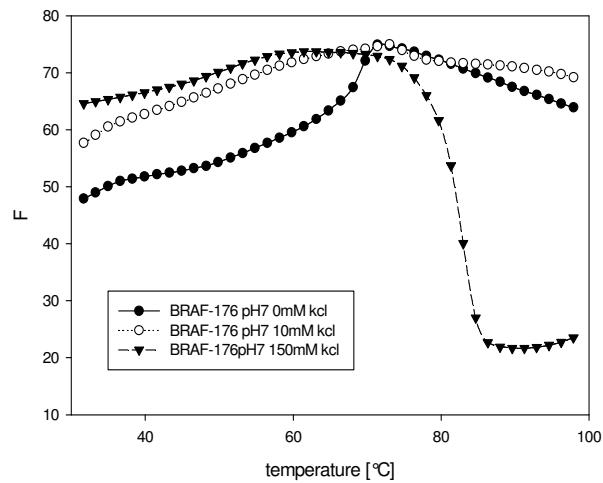


Fig. 1.52: melting curves of the double stranded BRAF-176 at different KCl concentrations.

In fig. 1.53 the example of the melting profiles of the three double stranded sequences are reported in the presence of 150mM KCl.

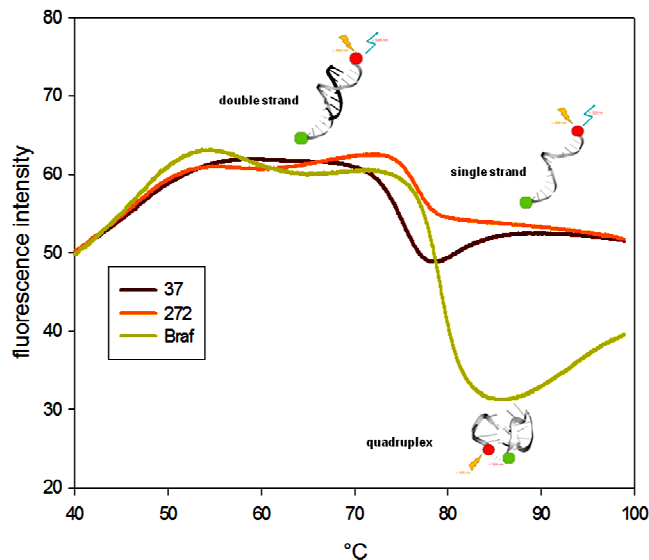


Fig. 1.53: melting profile of the double stranded three sequences at 150mM KCl

Despite a clear competition between double strand/G-quadruplex is evident for all the three sequences, BRAF-176 shows a higher propensity in structuring into G-4 (fig. 1.53).

This behaviour is probably the result of the high guanines number (24 guanines on 30 bases sequence) which characterize the sequence BRAF-176. Another explication is the BRAF-176 Kd value ($0.64 \pm 0.05 \text{mM}$) towards KCl which is significantly lower with respect the ones of the other sequences, thus suggesting a higher propensity to fold into G4.

By merging these data we can conclude that G4 formation can be clearly detected only upon heating the dsDNA at a temperature higher or close to its T_m and by including a KCl concentration sufficient to stabilize the tetrahelix to a larger extent. In these conditions, the preferential stabilizing effect of the metal ion on G4 versus ds, shifts the equilibrium toward the tetrahelix thus allowing to detect the G4 form although always at temperatures higher of the physiological one (fig. 1.54). Nevertheless, we can take into account that transient lowering of the stability of the double helix can be promoted by different other physiologically relevant events like a structural stress of the ds form, helicases, etc.

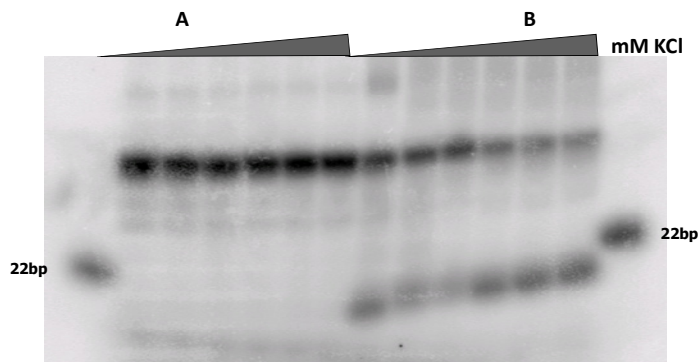


Fig. 1.54: EMSA of dsBRAF-176 when pre-folded in the absence of KCl than incubated with increasing concentrations of the ion (A) and when oligonucleotide was annealed in KCl (B).

In the case of BRAF-176 sequence, we extended the competition study, by investigating other physiological factors which can further modulate the G4/double strand equilibria, among these, we take into account the potential role of the crowding conditions. Starting from these evidences we monitored the structural equilibria of the double stranded also in the presence of PEG₂₀₀. Distinctly from what observed by adding just the potassium ion, the addition of the polymer perturbed the structural features of the double helix which results into an increment of the 280 nm signal. Interestingly this increments reaches a maximum after an annealing step. It is significant to underline that this signature perfectly matches the one obtained by combining the contribution related to the G-and C-rich strand. This allows to assume that PEG can actually play a favorable role in promoting the structural conversion of the double helix towards the G-quadruplex and iM structures (fig. 1.5).

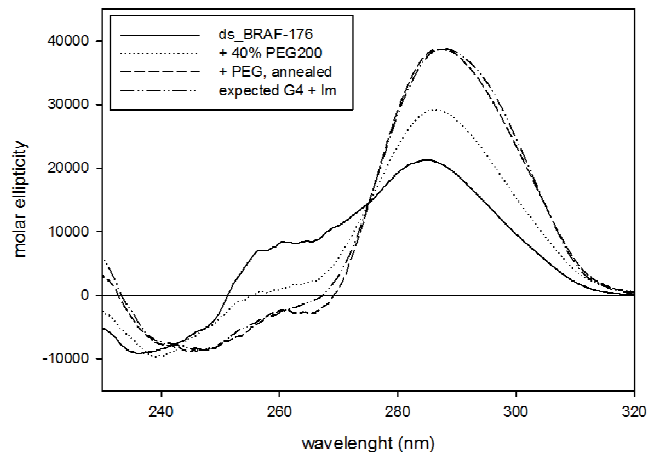


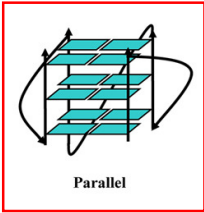
Fig. 1.55: comparison of dsBRAF-176 dicroic spectra in the presence of KCl and with or without PEG₂₀₀.

1.4 CONCLUSIONS


In conclusion, we can say that all the three tested sequences are actually able to fold into a G4 structure. However, the features of the folded forms are very different.

Here is reported a summarizing table (table 1IV) with the Kd values towards KCl obtained for the three sequences, the Tm at 50mM and the main conformation assumed in KCl.


	EGFR-272	EGFR-37	BRAF-176
Kd	28,46±4,19 mM	19,41±4,32 mM	0,64± 0,05 mM
Tm (50mM KCl)	55°C	74°C	90 °C
G-quadruplex conformation	Parallel	Antiparallel/ Parallel	Antiparallel



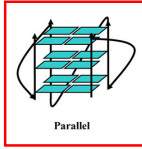
Parallel



Antiparallel



Antiparallel



Parallel

Table 1IV: summary of Kd values towards KCl calculated for the three sequences, Tm values obtained at 50mM KCl and prediction of the conformation assumed by the G-rich strand at physiological conditions (150mM KCl, pH7.5).

Since our aim was to assess the potential of these sequences as drug target the result herein collected are very promising. Indeed, they underline that for all tested sequences the G4 can exist in conditions comparable to those found in the nucleus.

Moreover, all of them fold according to different topologies thus representing a different target for a potential drug.

Conversely, the structural properties of the three complementary C-rich strands are more conserved. In any instances, the iM folding is strictly dependent on pH. KCl does not affect this conformation except for the sequence EGFR-37 which is negatively influenced by the presence of the ion. However, crowding conditions, which mimic the intracellular environment are able to increase the stabilization of the iM folding even at pH close to the physiological one (pH 7.0).

As a result we observed a direct correlation between the binding affinity of the three G-rich sequenced for KCl and their attitude to assume a G4 folding starting from the double strand form. Nevertheless, in all the experimental conditions we applied we were never able to observe a complete conversion of the B-form to the non-canonical structures at 37 °C. In the next sections we will take into account also this point in order to evaluate if the binding of proper ligands can implement the structural conversion.

Chapter 2

G-QUADRUPLEX FOLDING OF ONCOGENE PROMOTERS IN CELLS

So far, we always worked in vitro with synthetic short sequences and choosing the best conditions to promote their structural rearrangements. We considered interesting, for our purpose, to study also the potential G-quadruplex formation into a cellular environment (Pfeifer et al., 1990). For this reason, we set up an “in vivo” footprinting by inducing G-quadruplex directly in cells. Footprinting assay is a useful technique that can be suitably applied in DNA structural studies. Different chemical agents or enzymes are used for footprinting analysis depending on the needed information. For examples, S1 nuclease enzyme preferentially cleaves single-stranded regions over locally unwound regions, normal duplex regions, or secondary structures, such as cruciform DNA. KMnO₄ is a chemical probe that preferentially reacts with the C5–C6 double bond of unpaired thymidines. Br₂ is used to probe for secondary structures that can be formed by the C-rich strand.

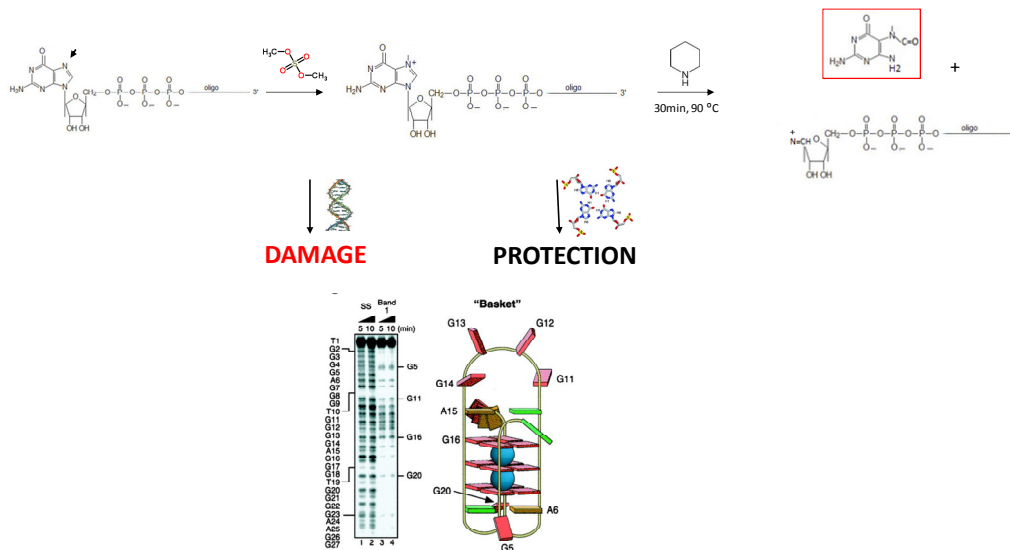


Fig. 2.1: DMS-footprinting reaction on G4 structure. In the gel, the guanines pattern of the unfolded (ss) and G4 (band1) folded oligonucleotide is shown (Siddiqui-Jain et al., 2002).

In fact Br₂ reacts with the C5–C6 double bond of the pyrimidine base within DNA, but its reaction with the impaired cytosine is at least 10-fold higher than that in duplex DNA. We

focussed on DMS footprinting that is particularly useful for fine-mapping the presence of G-quadruplex structures. In fact, DMS (dimethyl sulfate) is a chemical agent which selectively methylates the guanines and, to a lower efficiency, adenines. In particular, the methylation occurs on the N7 of the guanines and it generates a bond tension which produces a cleavage site in basic conditions. Since the N7 of the guanines are involved in the hydrogen bonds in the G-tetrads, if the G-quadruplex occurs, the N7 are protected by DMS attack. Consequently, no cleavage is detected at this sites on a sequencing gel [9] (fig.2.1). This protocol is widely applied on short sequences, thus we decided to assess its suitability in cells.

To set up the protocol, we choose as G-quadruplex template the purine-rich sequence in the anti-sense strand of the nuclease hypersensitive element III₁ (NHE III₁) upstream of the P1 promoter of c-myc oncogene (fig.2.2). Indeed, the G-quadruplex structure of this oncogene has been well characterized in physiological conditions and its implication as silencer element in cancer cells has been firmly confirmed (Brooks and Hurley, 2009; Brown et al., 2011; Grand et al., 2005; Siddiqui-Jain et al., 2002).

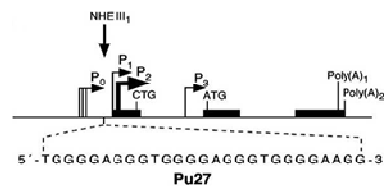


Fig.2.2: NHE III₁ region on c-myc oncogene.

Pu27 is the purine rich sequence upstream the P1 promoter which fold in G4.

By using a synthetic fragment of the c-myc promotorial region we verify by gel shift assay (fig.2.3), the G- quadruplex formation promoted by KCl physiological conditions. The occurrence of a band corresponding to the G4 form is evident starting from 50mM KCl.

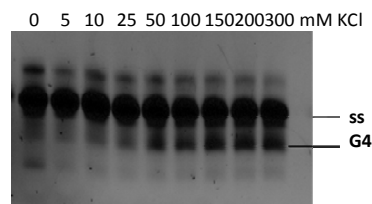


Fig.2.3: gel shift assay performed on Pu27 c-myc after incubation at increasing KCl concentrations for 48hours. 15% polyacrilamide (19:1) native gel with 12.5mM KCl/NaCl in 0.5X TBE.

We then moved to DMS-footprinting on Del4 plasmid, in which the c-myc G-quadruplex forming region was previously inserted between the sequence p3X2 and p3r used as selecting primers at the last step of the protocol (fig.2.4).

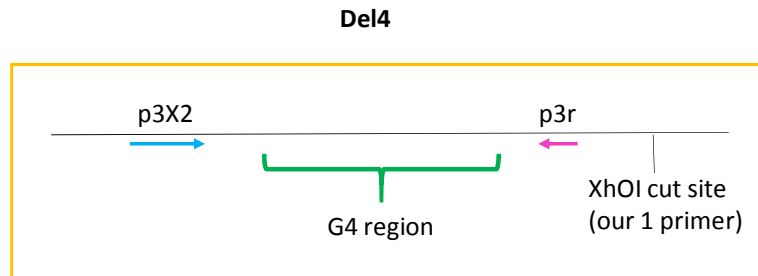


Fig.2.4: Schematic representation of Del4 plasmid with its G4 (G-quadruplex) region and primers design.

A scheme of the footprinting protocol is reported in fig.2.5. We first induced G-quadruplex on the plasmid with KCl (100mM). Then, we treated it with DMS and piperidine to generate the cleavage sites at the N7 of guanines not involved in G-quadruplex. Then, we transcribed the G-rich strand with 1 cycle of PCR (Polymerase Chain Reaction), step 2, using as primer1 the complementary sequence of XhoI cleavage site. The so obtained fragments are blunt ends at the 5' terminal of the G-rich strand. Moreover here the end contain unknown terminal sequences, due to the DMS-piperidine cleavage. So, in order to amplify them, we added a double strand blunt end short sequence containing p3x2 site at the cleavage site by a T4-lygase catalyzed ligation reaction (step3). At this point, a 25 cycles PCR (step4) was performed with two primers (forward and reverse) each pairing at known ends. Finally, we labelled the complementary strand of the G- rich sequence by an elongation step, using two sets of FAM-labelled primers: p3X2 in the forward or p3r in the reverse direction of the G-quadruplex region.

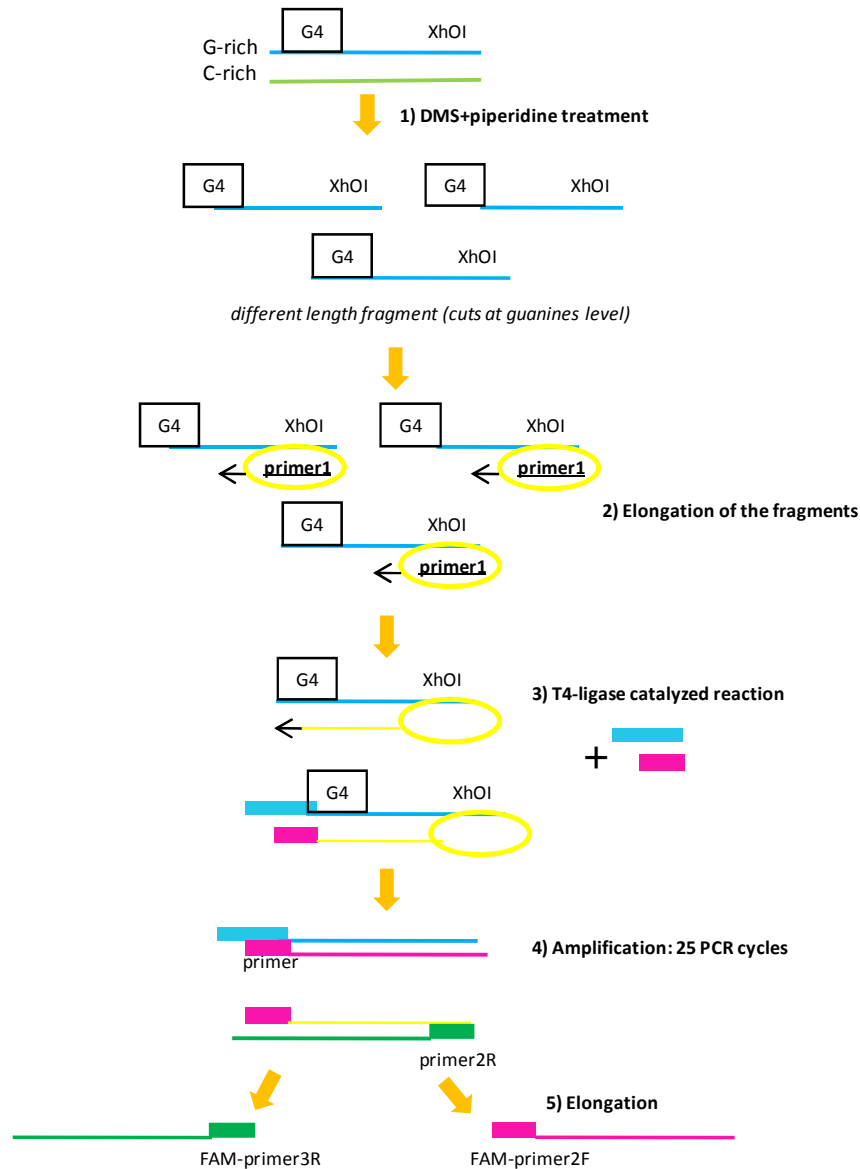


Fig.2.5: Schematic representation of in cells DMS-footprinting procedure

Due to the complexity of the protocol we actually were not able to obtain a defined DMS pattern from treated cells but we focused to verify the suitability of each step. In particular, we first verify step 1, to verify that after DMS reaction, piperidine was able to induce the cleavage of the methylated DNA sites. This is shown in the gel in fig.10 in which a band of <500 base pairs is related to the formation of Del4 cleavage products after DMS+piperidine treatment (fig. 2.6). This suggested us that the use of restriction enzymes

to select the G-quadruplex region is not required and that piperidine and specific primers would act as selecting elements for the G-quadruplex sequences.

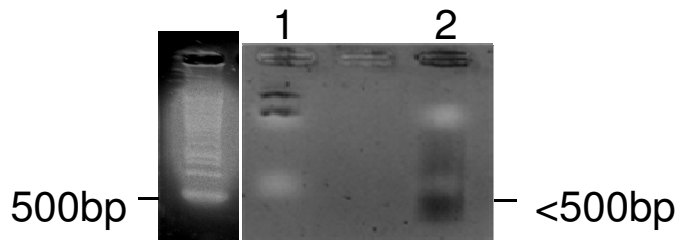


Fig.2.6: 1) Del4 DMS treated without piperidine treatment; 2)Del4 DMS treated and then incubated with 10% piperidine at 90°C for 30min.

So, we checked the suitability for our assay of selected primers. In lane 5 in the gel of fig.2.7 is visible the band (<500bp) referring to PCR products obtained combining primer2F and primer3r (see also fig.2.5).

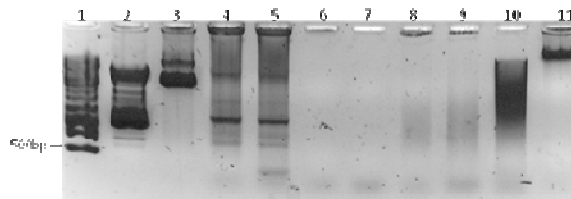


Fig.2.7: 1)molecular weight marker (1kbp); 2) del4 cut with PvuII enzyme; 3) del4 not cut 4) del4 cut with PvuII and amplified with the primers couple p1 and p3x2; 5) del4 cut with PvuII and amplified with the primers couple p2F and pr; 6)H₂O control with primers p1 and p3x2 and p2F and pr (7); 8) genomic DNA cut with PvuII and amplified with p1 and p3x2 and px2 and pr (9); 10 and 11) respectively: genomic DNA cut with PvuII and uncut.

Once assessed the suitability of this technique and the right functioning of the primers and PCR cycles, we shifted *in vivo*, by transfecting HEK-293 (Human Embryonic Kidney) cells with Del4 plasmid. We also tried on genomic DNA, by using lymphoma cells (Raji), which overexpress c-myc oncogene. We treated cells with two different compounds: Gcq-05 (NSC338258), which is a potent G-quadruplex stabilizer and SN-38 which is a topoisomerase I inhibitor, known also to induce G-quadruplex structure (Brooks and Hurley, 2009; Brown et al., 2011)(fig. 2.9).

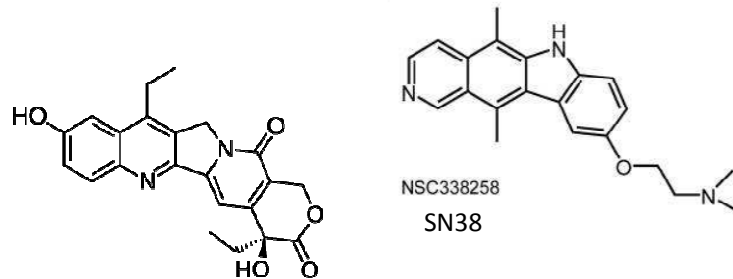


Fig. 2.9: chemical structures of GCQ-05 and SN38 compounds.

As in literature is reported that an efficient DMS reaction occurs when DNA fragments are of an average of <500 nucleotides long (Pfeifer et al., 1990) we also checked the efficient of DMS/piperidine reaction in vivo through an agarose gel (fig.2.10). However, we obtained DMS/piperidine treated DNA with a length higher than 1000bp.

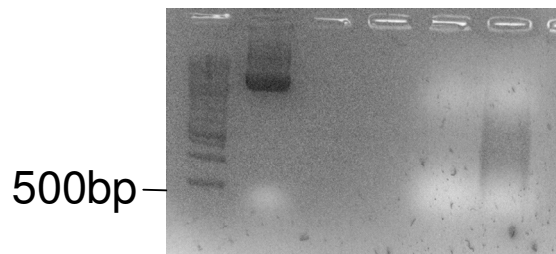


Fig. 2.10: to in vivo DMS/piperidine reaction. The average length of the nucleotides doesn't allow to continue the protocol with a high efficiency.

Nevertheless, we completed the protocol up to the last step and we performed sequencing gel on the obtained products. The bands on the sequencing gel are not related to myc sequence. We hypothesize first we should implement the plasmid transfection and extraction yield but also that, during PCR, aspecific amplifications can occur. Thus, we started the design of different sets of primers. Also, we thought that electrophoresis technique is not so sensitive for the detection of the products. Our future perspective on this set up protocol is to check the remaining steps and also to try products detection by capillary electrophoresis as a more sensitive alternative to sequencing gel.

Chapter 3:

BINDING STUDY OF ANTHRAQUINONES DERIVATIVES ON G4 TEMPLATES

3.1 INTRODUCTION

Anthraquinone derivative (**AQs**) represent an interesting scaffold to develop selective and multifunctional G4 ligands, with many potential applications. This derives from their well characterized DNA-binding properties(Cairns et al., 2002; Ndlebe and Schuster, 2006; Sun et al., 1997), fairly low redox potential and their ability to act as photosensitizers by one-electron oxidation(Ndlebe and Schuster, 2006). Structurally, **AQs** are strictly related to the anthracycline antibiotics like doxorubicin and daunomycin(Monneret, 2001). It has been shown that doxorubicin and daunomycin can interact with telomeric DNA *via* G4 stabilization, mediated by the anthraquinone scaffold and demonstrated by the crystal structure of a complex between the telomeric G4 DNA with daunomycin (fig. 3.1)(Neidle and Parkinson, 2008).

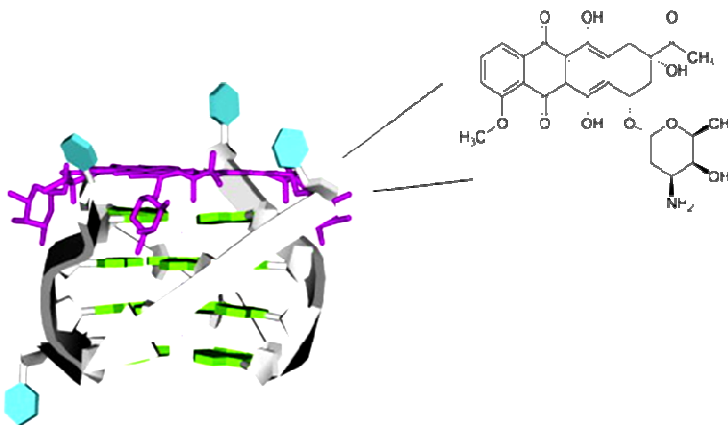


Fig.3.1: daunomycin-d(TG₄T) complex. Adapted from: (Neidle and Parkinson, 2008)

With the aim to optimize G4 recognition the synthesis of several 1,4- 1,5- 1,8- 2,6- and 2,7-difunctionalized amidoanthracene-9,10-diones, have been performed and the resulting compounds have been tested as G4 ligands(Clark et al., 2003; Perry et al., 1998; Perry et al.,

1999; Zagotto et al., 2008a; Zagotto et al., 2008b). The five different region-isomers showed different abilities to recognize G4 telomeric structures according to the nature and the position of the substituent side chains. Consistently, conjugation of the anthraquinone core with aminosugars or amino acids (Zagotto et al., 2008a; Zagotto et al., 2008b; Zagotto et al., 2011) was applied to modulate their G4 binding properties. One such example is a neomycin-anthraquinone conjugate that exhibits a nanomolar affinity for telomeric G4 DNA, which is 1000 fold higher when compared to its constituent units (Zagotto et al., 2008a; Zagotto et al., 2011). This higher affinity is ascribed to a dual binding mode of the conjugate which can interact with the grooves (neomycin) and with the guanines of the G4 (anthraquinone) *via* π -stacking interactions. For **AQ**-amino acid conjugates, the combination of a basic amino acid (Lys) with a more hydrophobic residue (Phe) has provided a better G4 selectivity *versus* the duplex DNA (Zagotto et al., 2011).

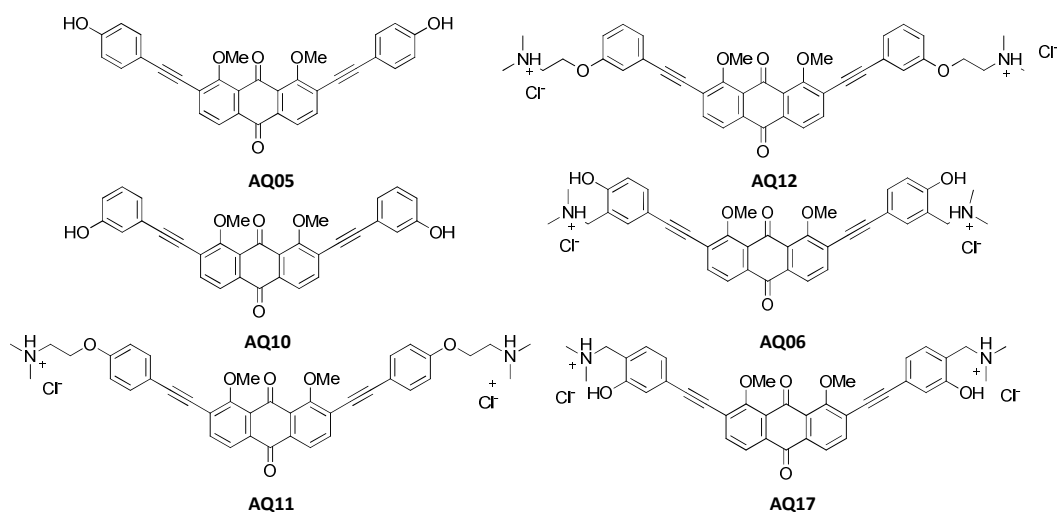
Unlike the large majority of G4 ligands, **AQs** exhibit interesting redox properties, as they easily generate radical anions and di-anions by bio-compatible reduction (van Dijk et al., 2006). Recently, we demonstrated that formation of stable radical anion can be exploited to generate alkylating agents such as Quinone Methides (QMs, generated from *o*- or *p*-benzyl substituted phenols) (Di Antonio et al., 2007; Modica et al., 2001). We anticipate that **AQs** could be similarly exploited for *in situ* generation of QMs at G4 sites, thus enabling G4 covalent targeting. Alkylation has been proposed as an alternative approach to physically lock the DNA G4 in its folded conformation, enabling the investigation of the biological implications associated with G4s stabilization (Doria et al., 2012; Doria et al., 2013). Effective electronic conjugation between the **AQ** core and the aryl moiety (embedding the QM precursor) will ensure generation of the alkylating QM under reductive conditions. Furthermore, it has been shown, recently, that V-shaped compounds have an enhanced affinity towards G4-DNA suggesting disubstituted compounds with a central symmetric and planar core. Starting from this data, we postulated that increasing the aromatic surface and the structural constraints by introducing aryl side chains at positions 1 and 7 of the anthraquinone molecule, would have been beneficial on both the **AQs** G4 binding properties and on G4 vs duplex selectivity.

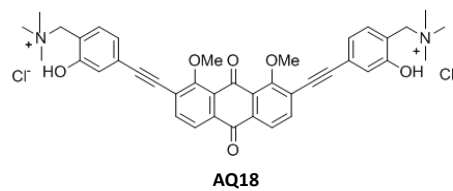
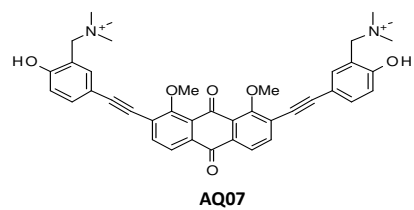
3.2 AIM OF THE STUDY

We started our work using the human telomeric DNA. Indeed, it is widely used as a model for the screening of novel G4 binders. Moreover, by properly turning its length and experimental conditions it is possible to trap it into different G4 conformations. In this study, we focused on a small library of aryl ethynylantraquinones which were synthesized as potential G4 binders. The synthesis involved the symmetric functionalization of the anthraquinone core at 2 and 7 positions with chemically diverse aryl moieties, such as: negatively charged phenolates arising from AQ05 and AQ10, positively charged secondary amines (AQ11, AQ12) and zwitterionic Mannich bases (AQ06, AQ18).

All the synthesized ligands have been tested for their G4 binding properties in comparison to double stranded DNA. Furthermore, in order to explore the potential recognition of G4 of different topologies, selectivity of the compounds was investigated by testing them on oncogene EGFR and BRAF selected sequences.

Due to the differential binding properties observed on the selected DNA substrates, the selected aryl ethynylantraquinones represent an interesting platform for the development of a new generation of multifunctional G4 interacting ligands.





3.3 RESULTS AND DISCUSSIONS

To assess the potential of the AQs to stabilize peculiar G4 topologies we screened them by fluorescence melting using DNA telomeric sequences properly labeled at the 5'-end with a quencher (dabcyl) and at the 3'-end with a fluorophore (fluorescein). An increase of the oligonucleotide melting temperature upon addition of the tested compound relies on the ability of the ligands to stabilize the DNA G4 folded structure. Since the human telomeric G4 is characterized by a large conformational flexibility, the analysis was performed under different conditions and with different sequences known to promote distinct folding: HTS (d[AG3(T2AG3)3T]) which in the presence of K⁺ folds mainly in a population of prevalently hybrid conformations, whereas in Na⁺ it assumes a defined antiparallel folding, and Tel24 [d(T2AG3)4] which adopts a hybrid-1 folding in K⁺ containing solutions. The same analysis was additionally performed using a double stranded random sequence (dsDNA) to check for duplex vs. quadruplex selectivity. To summarize our results we report the variation of the oligonucleotide melting temperature as a function of ligand concentration (Fig. 3.2)(Darby et al., 2002). Among the tested ligands only AQ05 and AQ10 did not induce any modification of the melting profile of the tested DNA sequences. This sustained the fundamental requirement of protonable groups in the side chain to grant effective nucleic acid recognition. Although the ΔT_m values remain quite low at 1 μ M ligand concentration, all the other compounds stabilized the G4 forms. In particular, a sigmoidal correlation emerged between the observed ΔT_m and ligand concentrations, which suggested the presence of cooperative binding events. In particular, AQ06 and AQ07 were the most active compounds towards telomere. Furthermore, at concentrations higher than 5 μ M AQ11 and AQ12 behaved as good ligands for Tel22 in K⁺. If we compare the behavior of each anthraquinone derivative for the different tested G4 targets, we did not observe prominent selectivity for any of them. The only exception was AQ12, which was less active on Tel24. Thus, the presence of oxygen in the meta position on the aromatic ring of the side chains seems to negatively perturb the DNA recognition process. Interestingly, the regio-isomers AQ06 and AQ17 showed very different binding profiles and this can suggest a peculiar binding interaction for the para isomer AQ06 with the G4 structures. Finally, all tested compounds

almost did not affect the thermal stability of the double stranded DNA (Fig. 3.2), indicating their use as potential G4 selective binders.

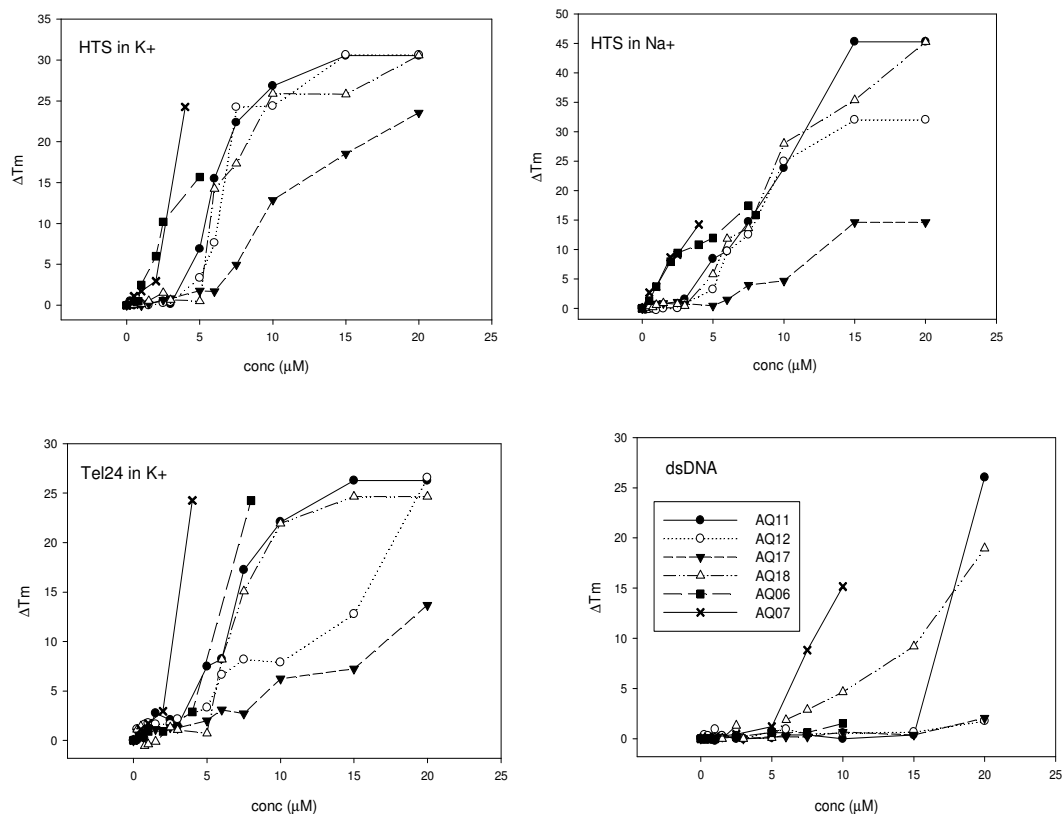


Fig. 3.2: Variation of the melting temperature of HTS, Tel24 and double stranded DNA (dsDNA) by increasing concentrations of the tested ligands in K⁺ or Na⁺ containing buffer.

CD titrations were performed to investigate the ability of the novel compounds to induce structural modifications to the tested G4. Thus, the study was performed using the same oligonucleotide sequences used for thermal stabilization experiments. Moreover, we extended our analysis to wtTel26: d[(T2AG3)4T2] which, in the presence of K⁺, folds into a hybrid-2 type of arrangement. The recorded dichroic spectra of the oligonucleotides in the presence of potassium are all characterized by two positive bands, one centered at 290 nm and the other at 265–268 nm which reflect the principal 3 + 1 hybrid arrangement assumed. Conversely, in the presence of sodium, the dichroic spectrum of Tel22 shows a negative band at 260 nm and a positive band at 290 nm, typical of the antiparallel conformation

signature identified by NMR spectroscopy. Fig. 3.2 Variation of the melting temperature of HTS, Tel24 and double stranded DNA (dsDNA) by increasing concentrations of the tested ligands in K⁺ or Na⁺ containing buffer. Variations of the intensity of the dichroic features of all tested G4 folded DNA sequences were detected upon addition of the ligands (Fig. 3.3). This confirmed the occurrence of a DNA–ligand interaction, which does not affect the G4 topology to a large extent. Interestingly, the most relevant CD variations occurred generally with AQ06. In this instance, induced dichroic bands (ICD) in the ligand absorption range were also observed (Fig. 3.3). Such contribution should derive from the insertion of the ligand chromophore into the chiral environment provided by the nucleic acid. Since this effect is a function of the mutual orientation of the AQ chromophore and DNA, we can assume that its presence/lack among the tested derivatives is linked to a significant repositioning of the ligand in the complex as a consequence of the side chain nature and position. Thus, distinct binding modes for AQ06 vs. AQ11, AQ12 and AQ17 can be further inferred. When the DNA substrate was arranged into a double helix, the most prominent effect was reduction of the 275 and 245 nm DNA dichroic bands. This should exclude the occurrence of an efficient intercalation binding mode for these ligands, since this process usually causes an increment of these optical contributions. This result is in agreement with the above reported lack of thermal stabilization induced by the tested ligands in this nucleic acid conformation.

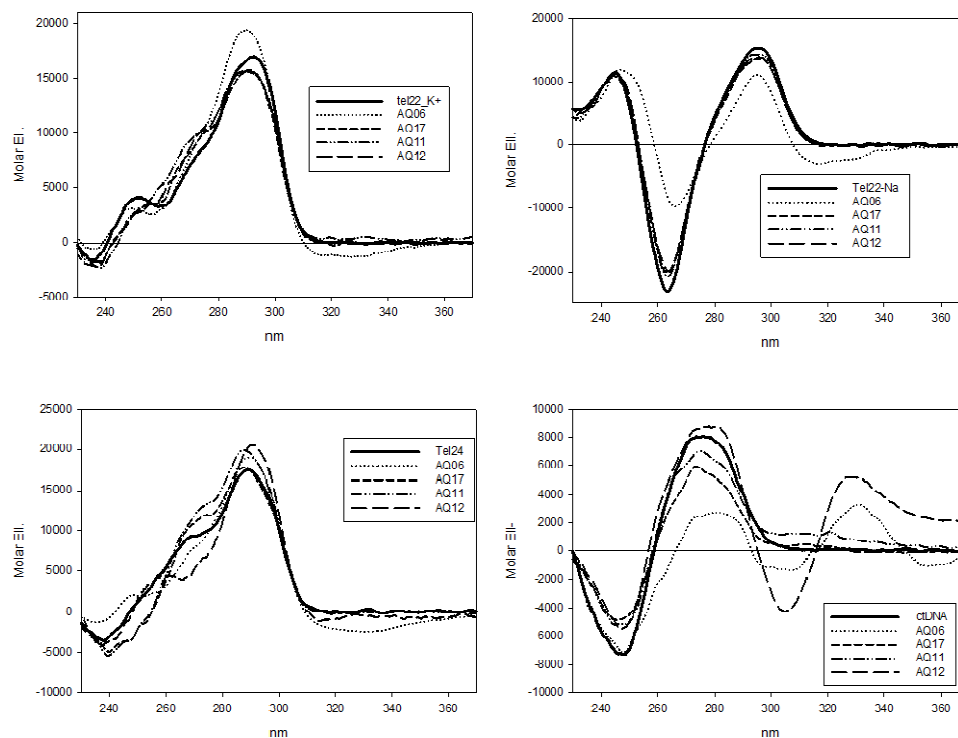


Fig. 3.3: circular dichroism spectra of DNA templates (4 μ M strand concentrations) alone (solid lines) or upon addition of 4 equivalents of tested AQs recorded in 10 mM TRIS, 50 mM KCl or NaCl, pH 7.0, 25 $^{\circ}$ C.

The above described results were collected to evaluate the recognition of a G4 structure by the tested ligands. Additionally, we analyzed whether they can promote G4 folding generating species sufficiently stable to interfere with enzymes devoted to the processing of the nucleic acid. Thus, the DNA polymerase stop assay was performed using a template containing a four repeat human telomeric sequence (HT4-temp)(Han et al., 1999). In a typical experiment, if the compounds under investigation promote G4 formation by the template, the formation of truncated products due to the collision of polymerase with the folded G-rich tract (see the cartoon on the right side of gel reported in fig. 3.4) appears. The results summarized in Fig. 3.3 showed that increasing concentrations of the tested ligands lead to a slight decrease of the intensity of the band relative to the fully processed oligonucleotide, which is more evident for AQ06. At the same time, the tested compounds blocked, by some other means, the primer extension by DNA polymerase starting from 5 to 10 μ M, the same concentration range in which they stabilized the G-quadruplex as

determined by the melting assay. The observed stop occurs at a well-defined site corresponding to the template G-rich stretch. Interestingly, AQ11 and AQ17 tend to arrest the enzyme at position-1 with reference to the G-rich tract at 5 μM concentration. This behavior is not shared by AQ06, thus further sustaining the different binding mode of this derivative. Conversely, the same reaction performed on a DNA sequence not G-rich (HT4sc-temp) failed to evidence any interference in the enzymatic activity by tested ligands up to 40 μM . This result correlates with the increased ability of the tested compound to recognize G4 over other nucleic acid arrangements.

Referring to spectroscopic analysis on anthraquinones compounds, we focused on AQ06, AQ11 and AQ17 to evaluate if they are able to stop the progression of polymerase due to induction of G-quadruplex structure on the template sequence (Han et al., 1999; Sun and Hurley, 2010). Stop assay performed with these derivatives (fig.3.4) highlights that all of them stop polymerase activity starting from 5-10 μM concentration. Interestingly, AQ11 and AQ17 behave in a similar way and in particular they tend to arrest the enzyme also at position -1 with reference to the G-rich tract. This behavior was not shared by AQ06 which start arresting the enzyme at 10 μM and also it can sustain differences in the binding mode of these derivatives.

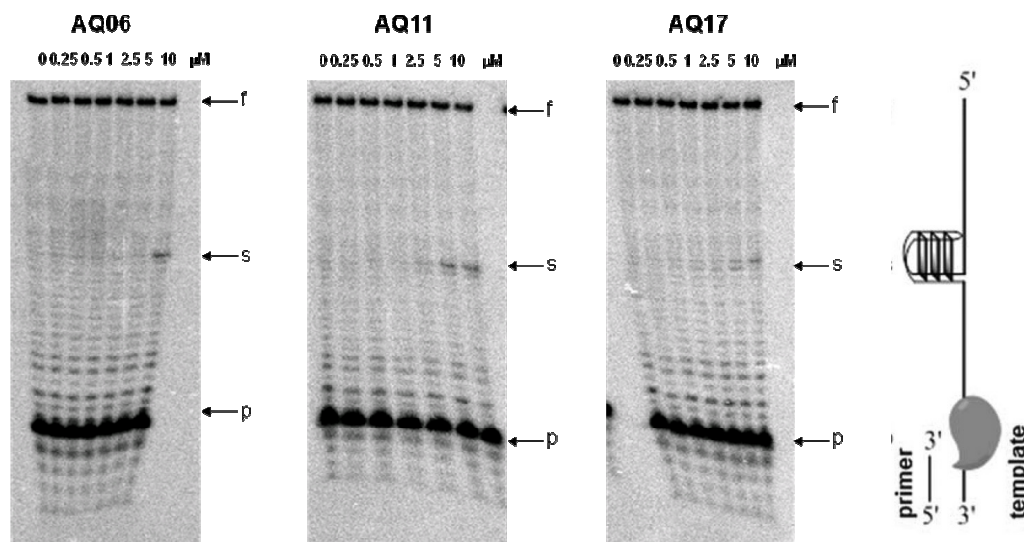


Fig. 3.4: *Taq*-polymerase stop assay performed with increasing concentrations of AQ06, AQ11 and AQ17 on the HT4-temp template in the presence of 50 mM K⁺.

As previously mentioned, the potential of anthraquinone derivatives to alkylate DNA has been exploited as a strategy to trap the G4 folded form. Thus, we investigated this effect in the presence of the most selective of our compounds. Since different conditions can positively contribute to the activation of the quinonemethide, we incubated the telomeric G4 with our derivatives at different temperature (50 °C) as well as in the presence of reducing agent dithiothreitol (DTT) or in basic conditions (pH 9.5). The results are summarized in fig. 3.5. At the higher AQ concentrations (50 μM AQ18), the formation of a higher electrophoretic mobility band can be observed. Due to the denaturation conditions applied during the product resolution, it can be attributed to the formation of a covalent complex ligand-DNA. We expected that upon changing the experimental conditions a modulation in the extent of the formation of this product could be evidenced. This was not the case. This suggests that the DNA-AQ18 adducts occur but with low efficiency. The same result was obtained also for the other screened compounds.

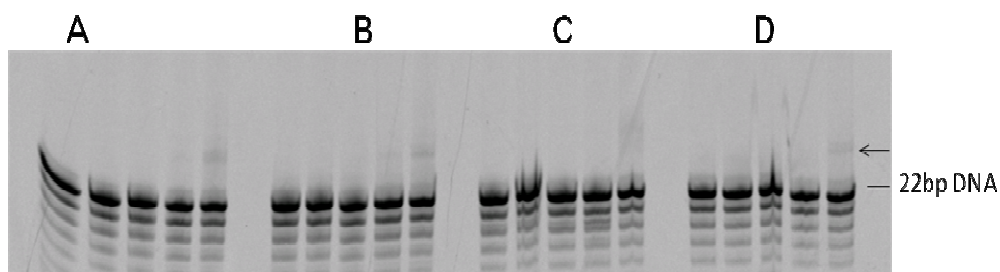


Fig.3.5: adducts experiment performed on 5'FAM-labelled 22 bases scramble sequence at 0-0.1-1-10-50 μM concentration of AQ18 at three different conditions: A refers to incubation for 1h at room temperature, B refers to the incubation at 50°C and C at pH 9.5 for 24hours and D to the incubation for 1h in 2mM DTT.

Considering the G4 stabilization of these class of compounds, and the consequent stop of polymerase activity, it was interesting to assess if they were able to interfere with telomere elongation, too. We took AQ11 and AQ06 as representative of good binders with different G4 binding modes. The telomeric repeat amplification protocol (TRAP) outlined that compounds AQ11 and AQ06 actually interfere with telomerase activity starting from 20 μM concentration (fig. 3.6). This suggests that tested compounds significantly stabilize telomeric G4 by inducing the arrest of telomerase enzyme.

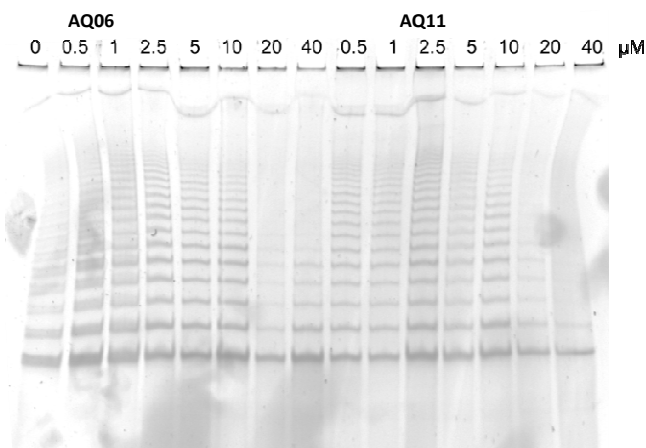


Fig. 3.6: TRAP assay performed with AQ11 and AQ06 compounds using protein extract obtained by HeLa cell line.

Finally, MTT assay was performed with tested anthraquinones on HeLa (cervical cancer) cells in order to assess if they can impair cell growth on this cell line. As reference compounds we

Compound	IC ₅₀
AQ05	-
AQ10	-
AQ11	2.55 ± 0.40 μM
AQ12	3.66 ± 1.25 μM
AQ06	2.10 ± 0.34 μM
AQ18	3.00 ± 0.51 μM
AQ07	>100 μM
AQ17	7.78 ± 2.69 μM
mitoxantrone	0.54 ± 0.18 μM

Table 3I: summary of the IC₅₀ values of AQs compounds obtained from MTT assay on HeLa cells

It is worth to underline that the above reported enzymatic assays showed impairment of telomerase and polymerase activity at higher ligands concentrations. Thus, we cannot relate the observed cytotoxic effect only to an interference with telomerase.

Nevertheless, our evidenced indicate these compounds as promising G-4 binders.

4.3.1 ANTHRAQUINONES DERIVATIVES ON ONCOGENIC SEQUENCES

Due to the significative G4 stabilization properties of tested AQ, we decided to test them also on our oncogenic sequences. In order to screen these potential binders, first we needed to identify proper experimental conditions which would allow to properly identify the melting of the free G-quadruplex. Indeed, the temperature at which the unfolding occurs must be not too high in order to allow its determination even when DNA is bound to the ligand. Thus we monitor the behavior of our sequences in the presence of increasing KCl concentrations since this metal ion modulates the thermal stability of the corresponding G-quadruplexes. As reported in fig. 3.9 the three sequences require different KCl concentration to melt at a temperature between 50-60 °C. This is in line with their different G-quadruplex arrangements and with their different affinity for the metal ion. Once assessed the proper KCl concentration for each oncogenic sequence, the stabilizing effect of the ligands was investigated.

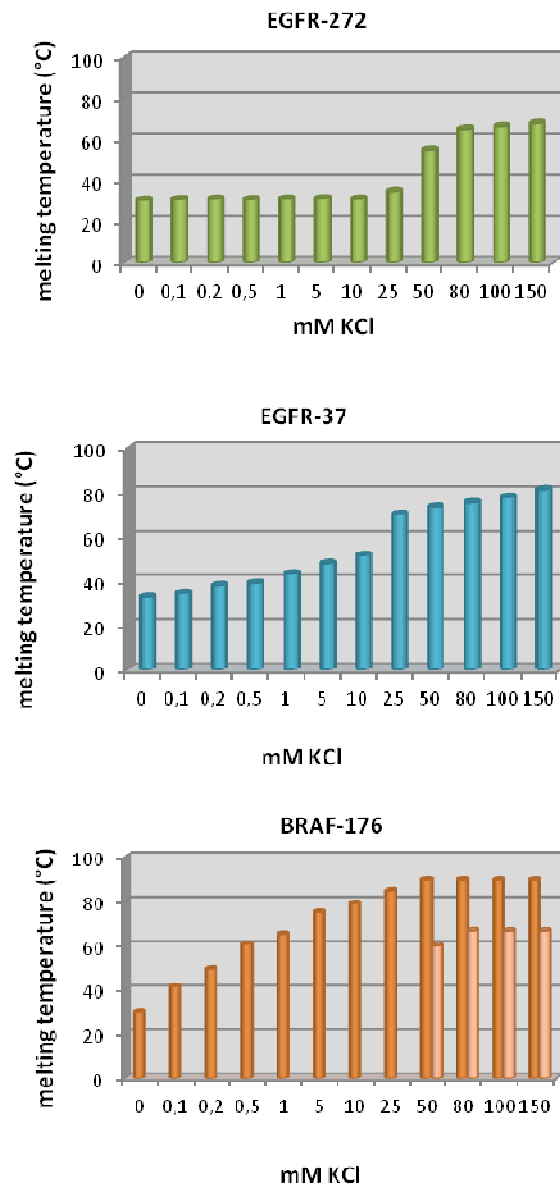


Fig. 3.9: melting temperature obtained for the three G-rich sequences at increasing KCl concentrations at 0.5 μ M concentration of oligonucleotide.

In particular, the fluorescence melting analysis were performed at 50 mM, 0.1mM and 1mM KCl for EGFR-272, EGFR-37 and BRAF-176 respectively. Considering their higher efficiency on telomeric G4, AQ11, 18, 06 and 07 were chosen as candidates of this screening.

Fig. 3.5 summarizes our screening results, and it outlines the preference of the selected derivatives to stabilize the G-quadruplex conformation with respect the double strand and the telomeric sequence.

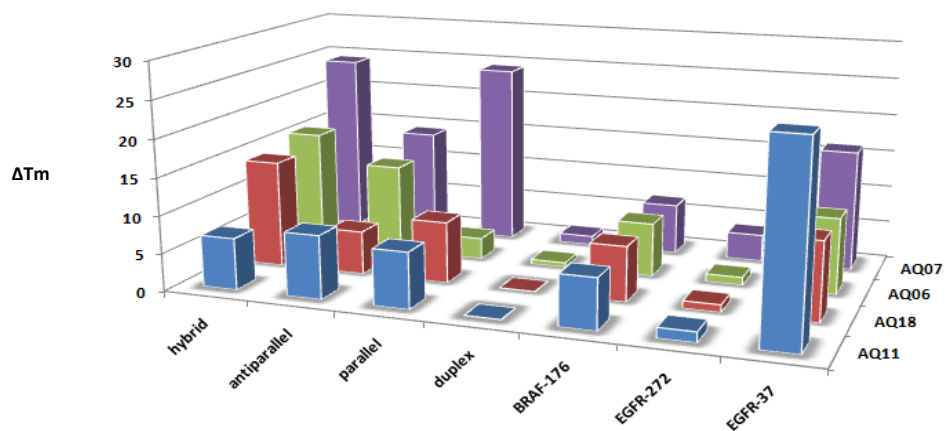


Fig. 3.5: melting temperature differences (ΔT_m) of some AQs derivatives on different telomeric and oncogenic G4 templates.

All the selected compounds showed the same stabilizing potential on BRAF-176 although no compound emerged as selective on this template. Conversely, EGFR-272 was not significantly stabilized by any of them. Not the same occurred on EGFR-37. In this case, AQ11 emerged as the most selective for the parallel/antiparallel mixed type G-quadruplex structure of EGFR-37, even if AQ07 also showed a high thermal stabilization, too.

3. CONCLUSIONS

According to all of the performed assays, compound AQ06 was the most effective ligand. The change of the relative position of the substituents on the phenolic aromatic ring from *para* to *ortho* (AQ06 and AQ17 respectively) provided compounds with quite different G4 binding and stabilization properties. Additionally, despite an overall conserved binding mode we showed that the hydroxyl groups on the lateral chains are crucial to cause a substantial shift in the localization of the AQ core over the G-tetrads.

Thus, our investigation highlights that structural positioning of Mannich bases is a key parameter for the design of efficient G4 binding. Although our best candidate is not as efficient as some of the most potent G4 binders among the AQ derivatives tested so far (Perry et al., 1998; Perry et al., 1999; Zagotto et al., 2008a), it apparently lacks significant intercalation into the double helix. This finding highlights the potential correlation of the observed cytotoxicity with G4 recognition. In this connection we have to take into account that for these compounds telomeric DNA might not represent the real preferential G4 target in the cells. Indeed, our preliminary results evidenced one of the screened compounds, AQ11, as a potential selective binder for the oncogenic sequence EGFR-37.

Although further structural refinements are required to increase to optimize the recognition of this scaffold towards G4-DNA, these preliminary results are encouraging. We are currently working on the development of a larger library of aryl (QM precursor) ethynylanthraquinone as G4 bi-modal ligands, acting as selective reversible binders endowed with a subsequent DNA alkylation mechanism upon reductive activation.

This would provide more potent G4 stabilizers with unprecedented potential applications as therapeutic agents as well as in diagnostic.

Chapter 4

DISUBSTITUTED NAPHTHALENE DIIMIDES AS POTENTIAL SELECTIVE BINDERS FOR G4 TEMPLATES

4.1 INTRODUCTION

As previously mentioned, the binding of G4 structures is supported by several chemical structures generally characterized by notably planar aromatic chromophores and positively charged lateral chains (Haider et al., 2011). It is known that intercalation between the two G-tetrads inside the quadruplex is thought to be difficult because of the high stability and rigidity of the G4 structure. Thus, stacking of the drug on the outer planes of the G-tetrads seems to be energetically more favored. Naphthalene diimides and related derivatives are confirmed G-quadruplex stabilizing compounds, particularly due to their chemical nature. Some substitution of the chemical scaffold have been shown to be essential for affinity and selectivity. In particular, the length of the lateral chain affects selectivity and the N-methyl piperazine at the termini of the lateral chain affects affinity towards G4. Also the number of side chain can be relevant for selectivity. Recently developed tetrasubstituted naphthalene diimide-based compounds (Collie et al., 2012) showed to significantly thermal stabilize human telomeric G-quadruplex DNA and inhibit the growth of several cancer cells. Furthermore, several studies show that positive charged polyamines, can interfere on gene expression through different mechanisms, for example, by interacting with DNA, they can modulate its flexibility or interconvert it into alternative structures, thus hindering the binding with proteins involved in transcriptions processes. The association of polyamines with DNA molecule has been confirmed by several biophysical techniques which include NMR spectroscopy, X-ray crystallography, Raman spectroscopy and molecular dynamics. As it concerns the G4 recognition, the interaction of the physiological polyamine spermine on the G4 of the thrombin-binding aptamer (TBA) has been shown and, most of all, the localization

of the polyamine on the narrow groove of the [d(G4T4G4)]₂ has been related to long-lives complex. This is clearly due to the high negative electrostatic field on the G4 groove with the positive charged spermine molecule(Keniry and Owen, 2013). Furthermore, as polyamines have been found to be in higher concentrations of in some tumor cells with respect the normal cells, their role in cancer proliferation has been investigated. In particular, it has been shown that these molecules at lower concentrations (less than 1mM) promote G4, by inducing cancer cells apoptosis, whereas at higher concentrations (more than 1mM) denature the G4 structure(Sun et al., 2011). This data on polyamines mechanism is important to outline the possibility to enhance G4 recognition by targeting the grooves with ligands containing a polyamine lateral chain.

4.2 AIM OF THE STUDY

Considering these important evidences, we aimed to design a strategic ligand with at least two specific features: affinity, and selectivity. In particular, affinity can be provided by compound disubstitution, and the introduction of a protonable polyamine lateral chain brings to an increment in selectivity by groove recognition (fig. 4.1).

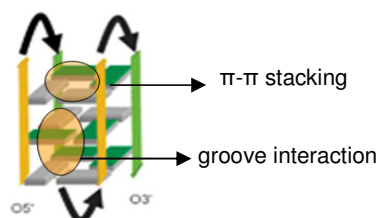


Fig.4.1: binding model of a disubstituted naphthalene diimide derivative containing a polyamine lateral chain, with an antiparallel G4 structure.

A previous work performed in our lab identified some NDI derivatives as good G4 binders. They were characterized by symmetrical substituents on the lateral chain. Although most of them lack a relevant selectivity for this conformation showing a good affinity also towards the double strand, two compounds emerged for their higher preferential for the G4 template (fig. 4.2).

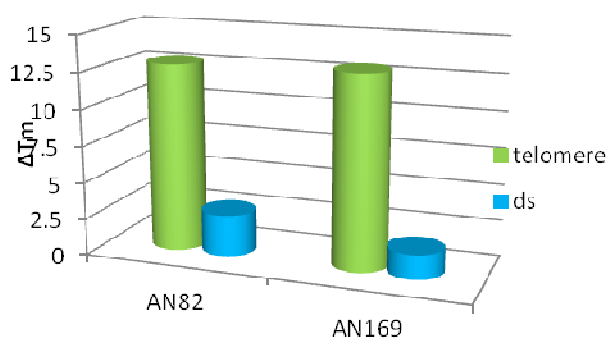
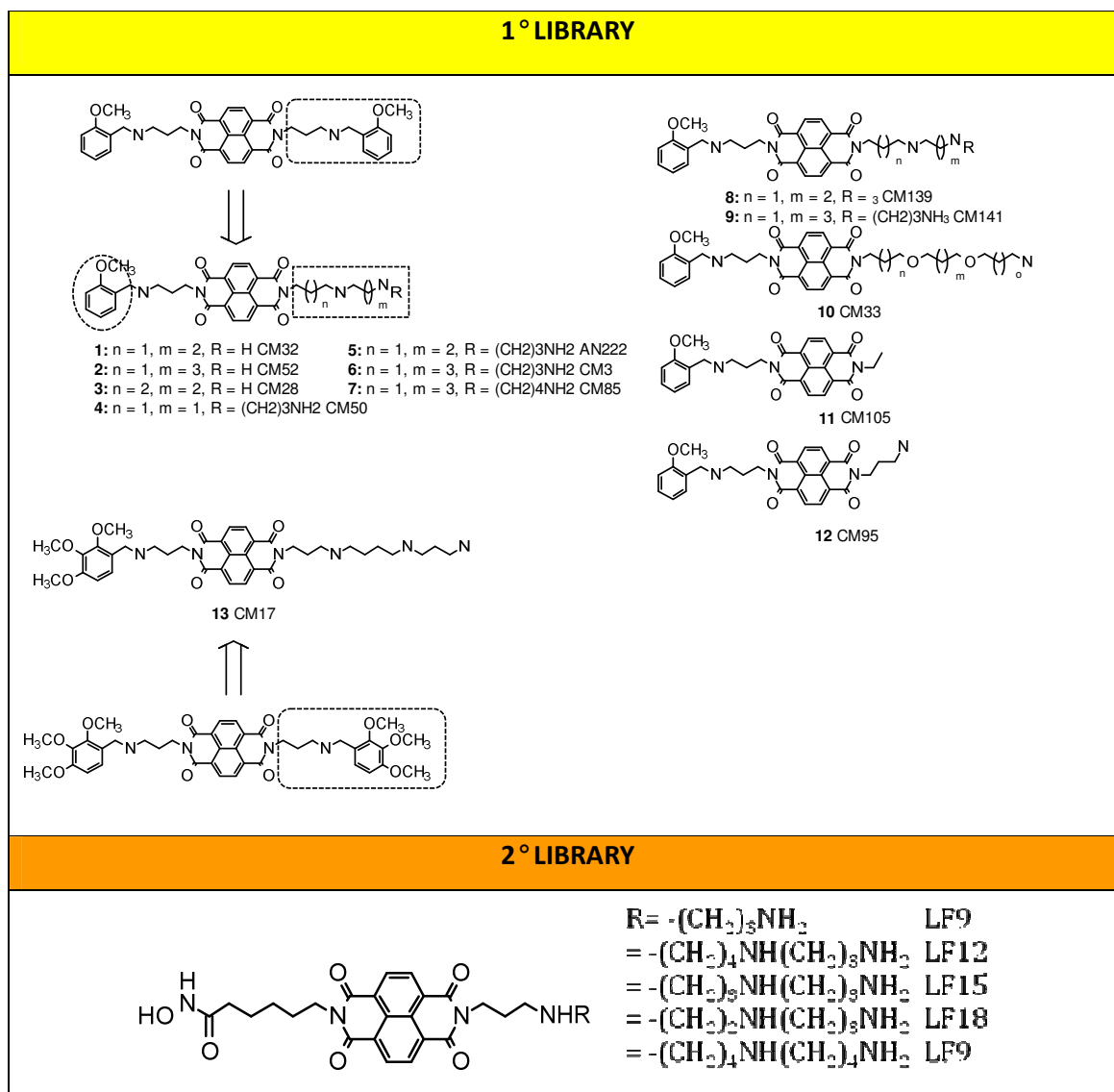


Fig. 4.2: melting temperature differences on telomeric G4 and double strand of AN82 and AN169.

Two libraries of disubstituted naphthalene diimides derivatives were thus synthesized by the pharmaceutical department of University of Bologna. In particular, the first library comprises derivatives containing an aromatic ring on the lateral chain. Starting from the symmetrical

AN82 and AN169, several derivatives were synthesized and the asymmetrical lateral chain containing a polyamine substituent was also considered. Conversely, the second library comprises naphthalene diimides without the aromatic ring on the lateral chain.



We carried on the binding study of these derivatives on several G4 templates assumed by the telomeric sequence and we compared their affinity towards double strand. Starting from a symmetrical chemical structure, the effect of the introduction of a asymmetrical substituents was explored and the selectivity towards G4 templates was studied.

4.3 RESULTS AND DISCUSSIONS

The stabilizing effect on the telomeric G4 and the random double stranded DNA (dsDNA) was first considered. Generally, all derivatives were active on both templates. The fluorescence melting assay allowed to select the derivatives with the higher G4/dsDNA ΔT_m ratio. In particular, as shown in fig. 4.3, the LF series compounds showed no significant selectivity towards the telomere, and they were excluded from additional screening. A similar behavior was shown by the asymmetrical CM17, AN222, CM141, CM95 whereas CM105 was not able to significantly stabilize the two secondary DNA structures. Conversely, the asymmetricals CM52, CM28, CM32, CM3, CM33 apparently have the telomeric sequence as preferential template (fig. 4.3), with CM52 being the most active.

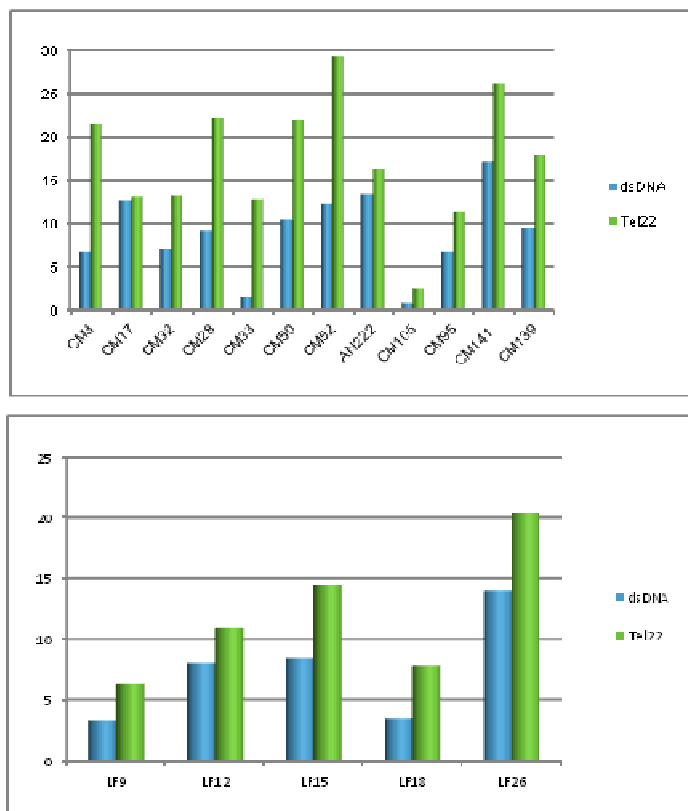


Fig. 4.3: ΔT_m (y axis) on double stranded DNA and telomeric G4.

By comparing these data with those related to the symmetrical AN82 and AN169, the selectivity is conserved but the efficiency is improved. Thus, we selected CM52 and CM28 as the most promising compounds and CM32 was considered for its modest activity (selectivity) for further analysis in comparison to the symmetrical derivatives.

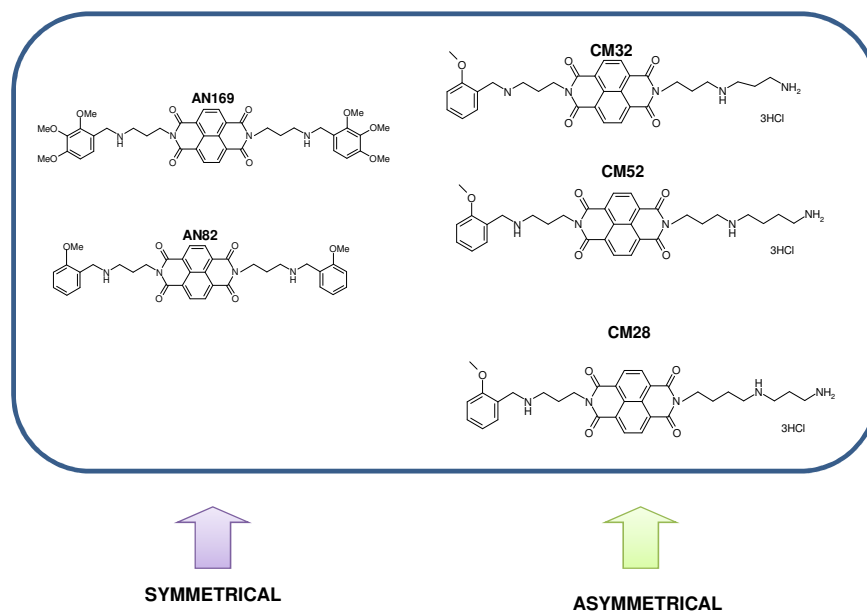


Fig.4.4: symmetrical and asymmetrical selected derivatives.

Considering the potential of most compounds to bind also the double stranded DNA, even with different affinities, we first point our attention on this template. By agarose gel electrophoresis (fig. 4.5) we actually confirmed that our compounds bind, also the plasmidic DNA (pBR322) even with different affinities. These data perfectly agree with the previous spectroscopic analysis. Indeed, it clearly emerged that CM52 lower the pBR322 electrophoretic mobility at higher concentrations with reference to CM28 and CM32. Compound CM32 was considered as positive control, as from the fluorescence melting experiment it emerged to preferentially bind the double strand form. Also AN82 and AN169 are able to bind the plasmid even with less efficiency in comparison to the asymmetrical derivatives.

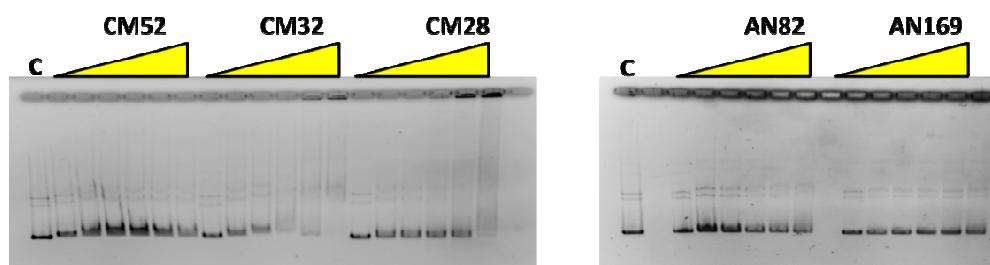


Fig. 4.5: plasmid binding of compounds CM52, CM28, CM32, AN82 and AN169 at the following concentrations: 1-5-10-25-50-100 μM

Another aspect we focused on, was the potential of the selected compounds to interfere with the activity of nuclear proteins such as Taq-polymerase, topoisomerase II α and telomerase, all involved in cell replication processes. In particular, we observed a clear difference among the symmetrical and asymmetrical substituents to hinder taq-polymerase activity (fig. 4.6).

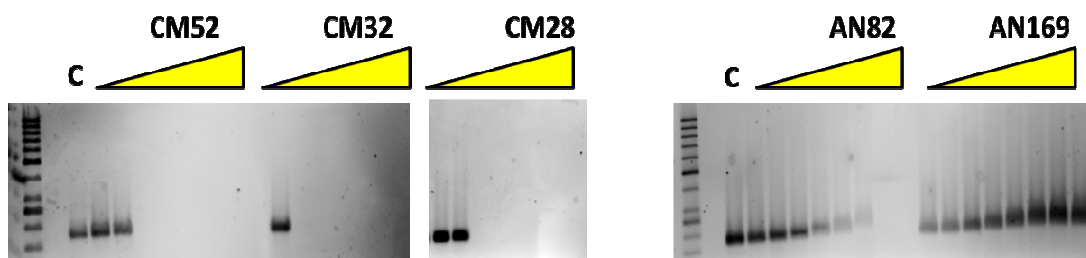


Fig. 4.6: Taq polymerase inhibition assay performed for compounds CM52, CM28, CM32, AN82 and AN169 at the following concentrations: 0.25-0.5-1.0-2.5-5.0-10-20-40 μM .

Conversely, none of them is able to interfere with telomerase activity, as shown by TRAP assay in fig. 4.7. The exception was AN169 which showed inhibition, even low, starting from 8 μM concentration. It is worth to underline that TRAP assay cannot be performed at ligand concentrations that affect *taq* polymerase. Thus, only for AN169 we were able to test NDI concentrations $\geq 1\mu\text{M}$ (table 4I). Nevertheless, its lower *taq* polymerase inhibition and its consequently higher interference with telomerase activity is probably due to AN169 lower affinity towards the double strand DNA.

COMPOUND	Taq INHIBITION
CM52	>0.5 μ M
CM32	>0.25 μ M
CM28	>0.5 μ M
AN82	>10 μ M
AN169	>40 μ M

Table 4I: taq inhibition concentrations for the selected compounds.

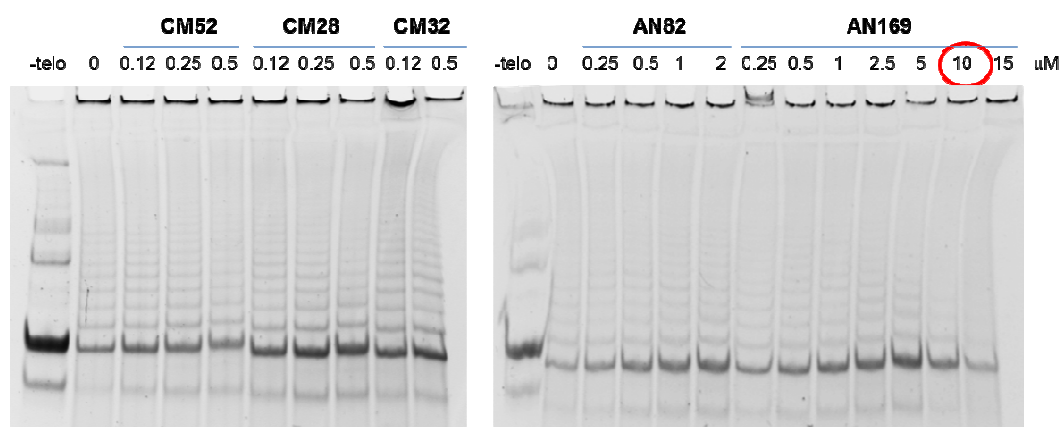


Fig. 4.7: TRAP assay performed for compounds CM52, CM28, CM32, AN82 and AN169.

As previously mentioned, also the interference with topoisomerase II α mechanism of action was studied. First we explored the potential of NDIs to inhibit the enzyme activity. For this assay, we used protein which fully relax the supercoiled plasmid. Moreover, due to the previously discussed DNA binders properties of our ligands before loading, samples were added of 200mM NaCl and 0.1% sodium dodecyl sulphate (SDS). The high ionic strength allows to break potential DNA-Topoisomerase adduct which would result in a reduction of the nucleic acid electrophoretic mobility and the SDS remove the DNA-ligand interaction. As a result, variations on the electrophoretic run rely only on the activity of the enzyme before loading. It emerged again that all compounds were actually able to hinder topoisomerase relaxation effect on pBR322 which the asymmetricals being the most active compounds.

Indeed in the gel (fig. 4.8), the band corresponding to the relaxed plasmid (lower band) is visible starting from low concentrations.

Then, we moved to verify the ability of the tested compounds to stabilize the DNA covalent complex where the enzyme is covalently linked to the DNA in a cleaved form. In order to do this, the experiments were repeated in the presence of higher enzyme amount. In this case before loading the enzyme was digested by proteinase K in order to trap cleavage complexes which should appear as a linear plasmid (band L in the etoposide line in the gel of fig. 4.9).

From the gel (example in fig. 4.9) it emerges that no compound was able to induce the poisoning of topoisomerase II α . The slowing down of the band at higher compound concentrations was attributed to the tight binding of the compound to DNA rather than to the enzyme inhibition.

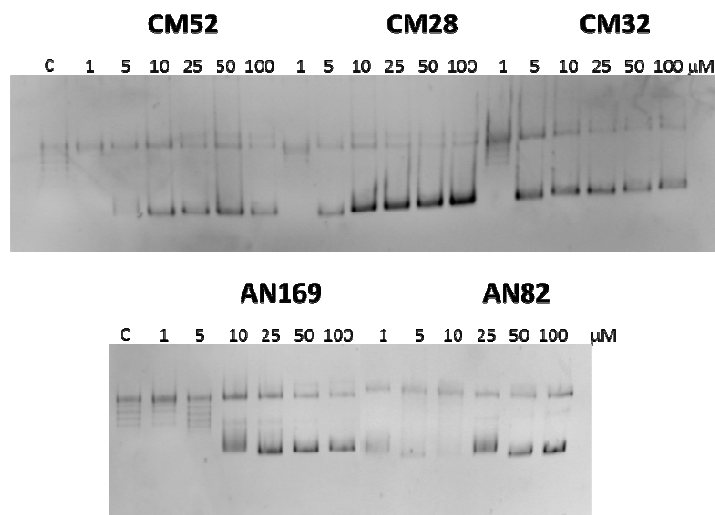


Fig. 4.8: topoisomerase II α inhibition assay (1% agarose in 1X TAE:10/1mM Tris-EDTA, 0.1% acetic acid pH 8.0).

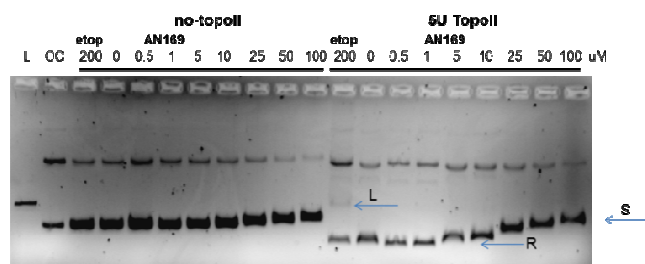


Fig. 4.9: topoisomerase II α cleavage assay for compound AN169. 200 μ M etoposide was used as positive control. L, OC and S refer respectively to the linear, open circular and supercoiled plasmid.

COMPOUND	Taq INHIBITION	IC ₅₀ Telomerase	IC ₅₀ Topoisomerase
CM52	>0.5 μ M	/	8 μ M
CM32	>0.25 μ M	/	1.5 μ M
CM28	>0.5 μ M	/	7 μ M
AN82	>10 μ M	/	18 μ M
AN169	>40 μ M	8 μ M	12 μ M

Table 4II: IC₅₀ of the selected compounds towards taq polymerase, telomerase and topoisomerase II α inhibition.

Once characterized the binding on double strand DNA and the consequent enzyme inhibition, the G4 was investigated as potential target for these naphthalene diimides derivatives.

As anticipated, focusing on these compounds, a high difference in terms of thermal stability regardless the symmetrical and asymmetrical derivatives was observed. In particular, comparing compounds CM52 and AN169, it emerged that the asymmetrical substituent has a higher stabilizing potential on G4 (fig. 4.5).

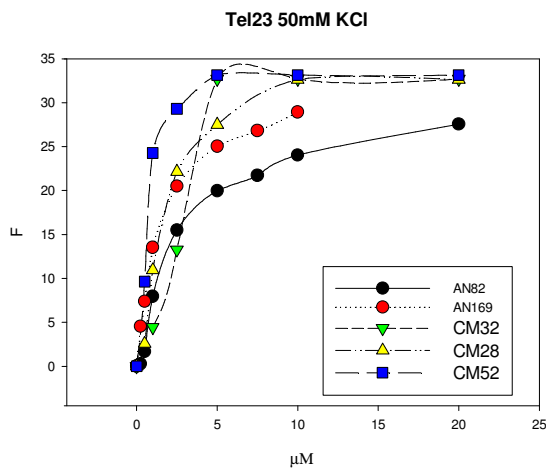


Fig. 4.5: fluorescence melting experiment performed on telomeric sequence annealed in 50mM KCl containing buffer.

This evidence prompted us to investigate the selectivity of the selected compounds towards different G4 template, in particular the oncogenic sequences previously characterized.

From our data, it emerged a clear difference among the two groups with the asymmetrical NDIs being always the best performers (fig. 4.6). Interestingly, this modulation is maximally evident on EGFR-272 probably due to its tendency to fold into a unique, well defined parallel G4 structure thus suggesting it as a preferential target for the asymmetrical binders.

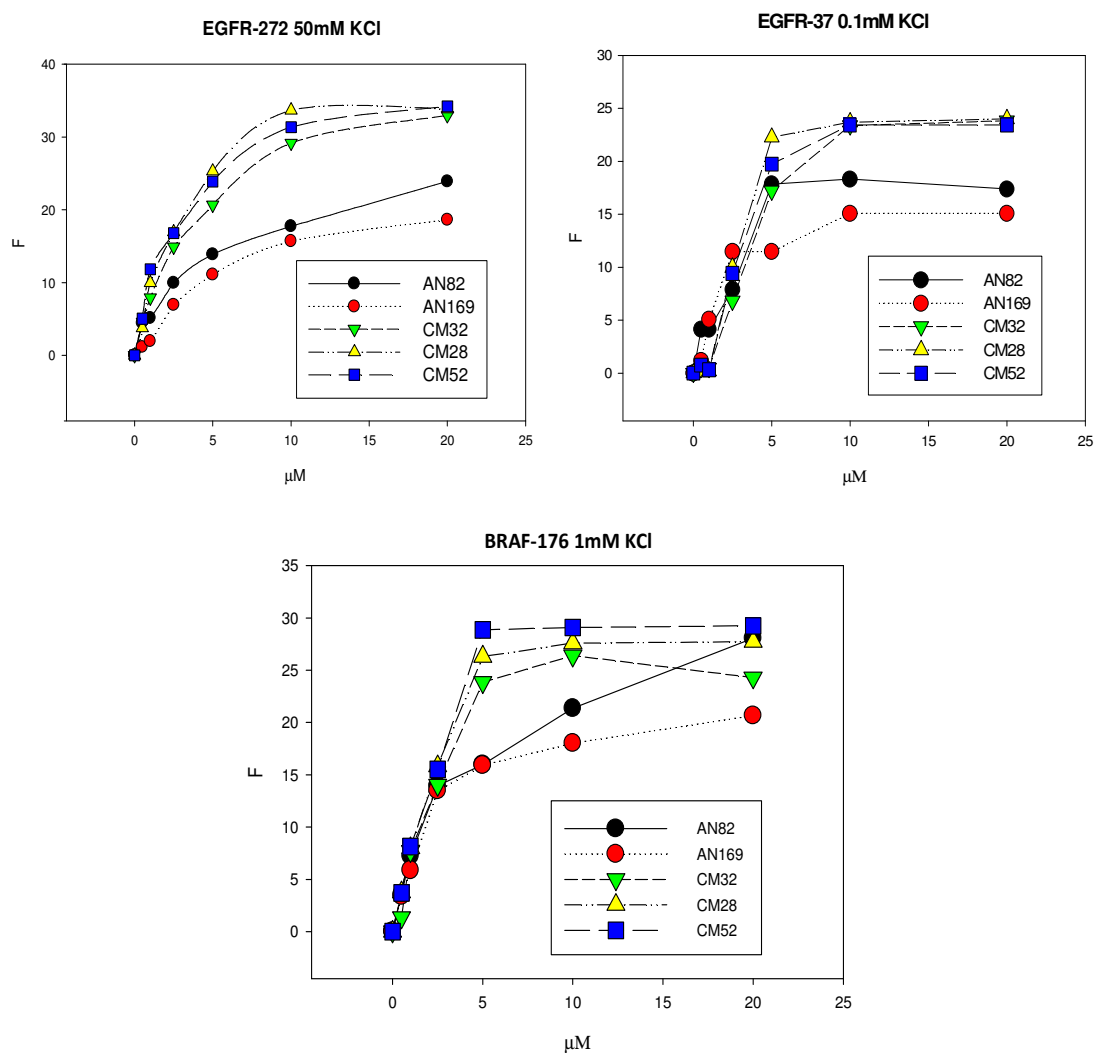


Fig. 4.6: fluorescence melting experiments on EGFR-272 50mM KCl, BRAF-176 on 1mM KCl, EGFR-37 0.1mM KCl, with selected NDIs.

Considering these preliminary results, we decided to better characterize the stabilizing effect on EGFR-272 G4 template with reference to the one on telomere. As model compounds we choose CM52 and AN169 (fig. 4.7).

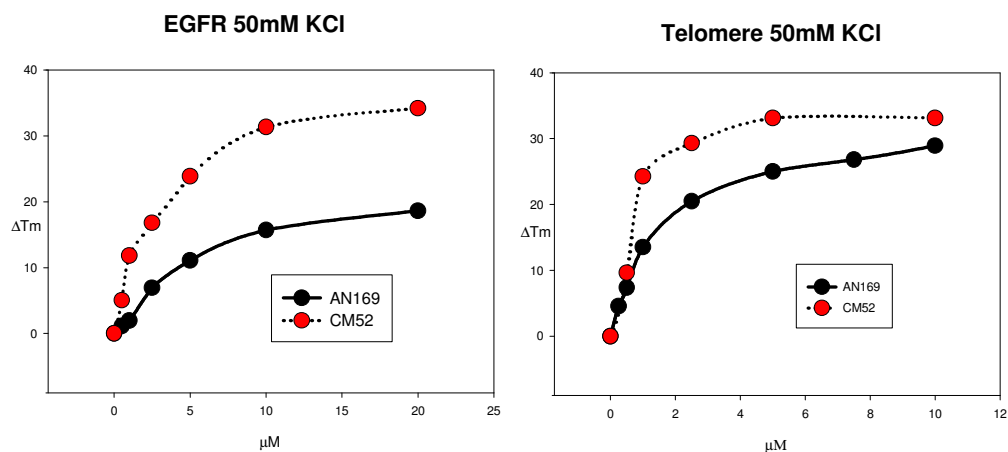


Fig. 4.7: melting temperature differences (ΔT_m) of the selected naphthalene diimides substituents on telomere and double stranded DNA. Data refer to 2,5 μM concentration compound.

Fluorescence melting data evidenced that the preferred template is the telomere for both compounds. However, a higher difference in terms of thermal stabilization occurs on EGFR-272. Indeed, on EGFR-272 the asymmetrical CM52 induces a higher ΔT_m with respect to the symmetrical AN169. Thus, we pointed our attention on the binding mechanism of the two NDIs on EGFR-272 by investigating the binding modes and the binding affinities.

The dichroic spectrum is strictly related to the topology of the DNA molecule, for this reason circular dichroism was useful to outline if upon binding EGFR-272, the selected derivatives alter its folding. We first annealed the oligonucleotide in buffer at 50mM KCl concentration. Indeed, as previously outlined, at this experimental conditions, the EGFR-272 is partially parallel folded and the effect of the compound on the sequence can be appreciated. When the oligonucleotide was titrated with tested compounds, its dichroic spectrum was significantly switched from a parallel type shape towards an antiparallel one by suggesting a specific interaction of the compound on the G4 molecule (fig. 4.8). Parkinson et al. (Parkinson et al., 2008), previously reported that the topology of the telomeric G4 is not perturbed by the NDI interaction. Probably it is due to the parallel and hybrid type conformation assumed by the telomere which best fit with the NDI interaction mode.

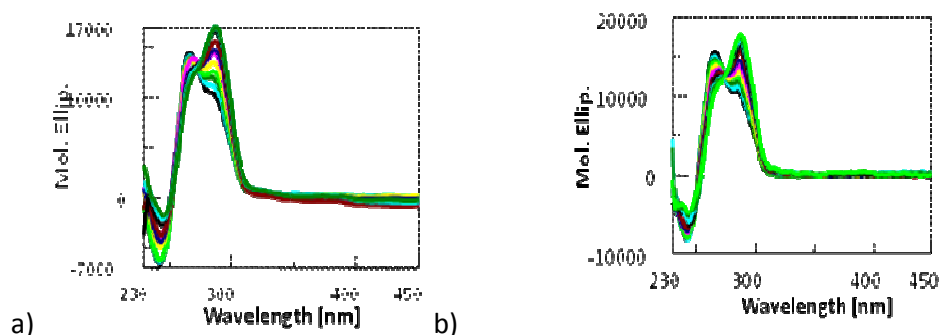


Fig.4.8: circular dichroism titration of single stranded EGFR-272 by CM52 (a) and AN169 (b).

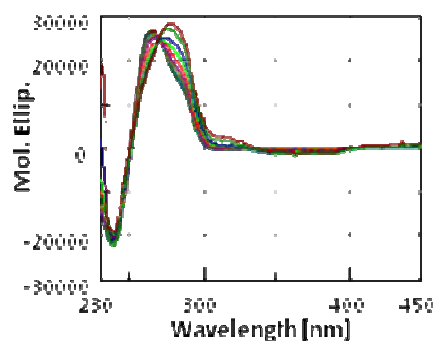


Fig. 4.9: circular dichroism titration of double stranded EGFR-272 by CM52.

Comparing the dichroic spectra of the two derivatives on EGFR-272 it emerged that both act with the same binding mode and efficiency on the single strand (fig. 4.8).

Considering that oncogene promoters are double stranded sequences (fig. 4.9), and considering also the role of G4 as a switching off mechanism of gene transcription, we investigated also the interactions of CM52 and AN169 on the double stranded EGFR-272. Again the double strand spectrum was perturbed in favor of the antiparallel G4 signature by both compounds, even with lower efficiency with respect the single stranded oligonucleotide (fig. 4.10). This is a relevant result as we previously outlined that the G4 of EGFR doesn't form from the double strand at physiological temperature. Thus the role of ligand as conformational switchers was confirmed (fig. 4.11).

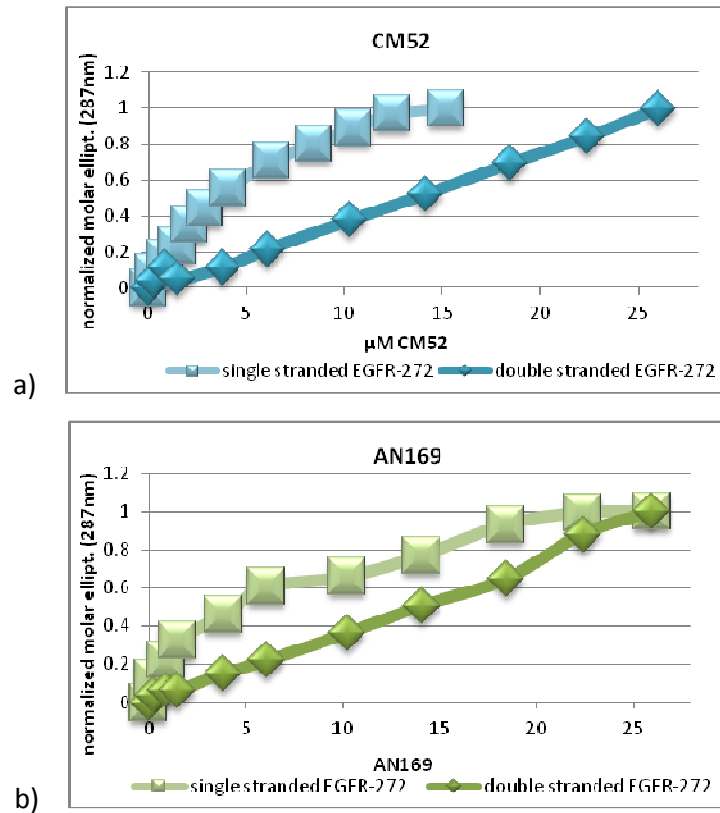


Fig. 4.10: normalized molar ellipticity for single stranded and double stranded EGFR-272 in function of increasing compounds concentrations, CM52 (a) and AN169 (b).

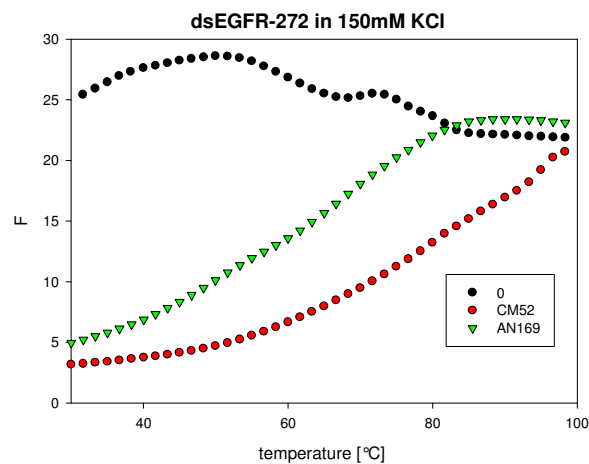


Fig. 4.11: melting curves of the double stranded EGFR-272 (previously annealed in 150mM KCl) at 20 μM concentration of CM52 and AN169.

To better highlight different affinities of our ligand for the target G4, the binding affinities were investigated according to different analytical techniques. As first approach we performed ITC (Isothermal Titration Calorimetry) titration which can provide full characterization of the thermodynamic parameters (Buurma and Haq, 2007) (Giancola and Pagano, 2013; Murat et al., 2011) (Freyer and Lewis, 2008). From isothermal titration calorimetry analysis it emerged that both compounds bind EGFR as well as telomeric G4 according to an exothermic process. Focussing on the binding affinity constants, a higher affinity constant of CM52 on EGFR-272 was found (fig. 4.11).

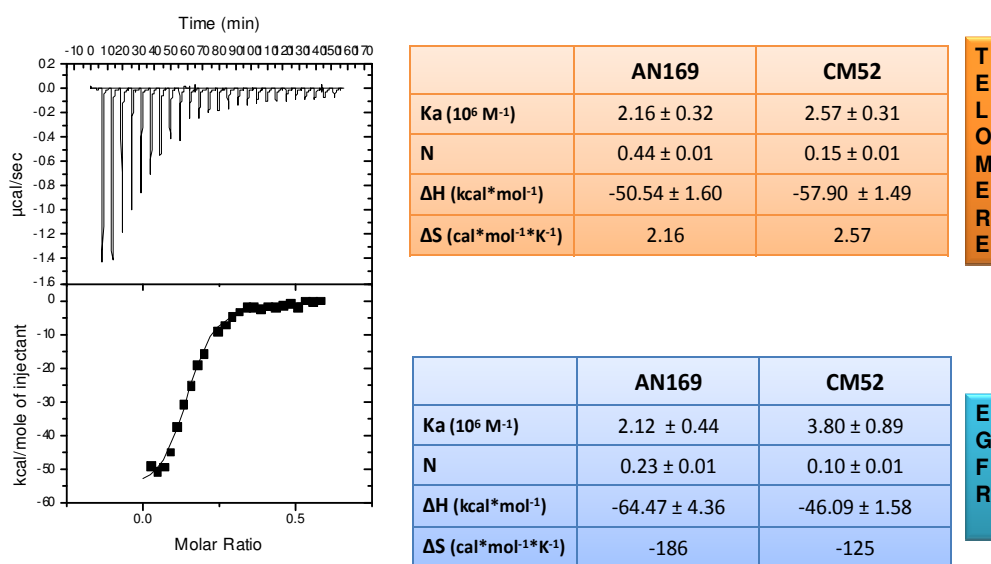
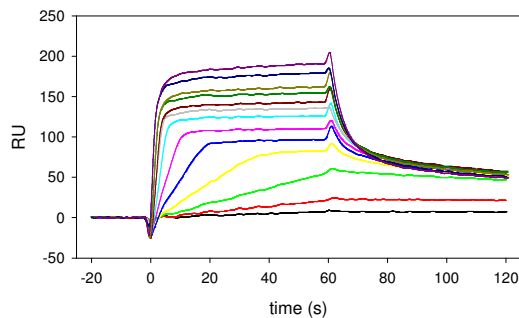


Fig. 4.12: thermodynamic parameters obtained through ITC experiment.

This preliminary result, strictly agrees with previous thermal analysis. However, ΔT_m and K_a well relate. A high binding stoichiometry was always found even it was lower in the case of AN169 with telomere. This suggests the occurrence of aspecific bindings on DNA which is probably related to the polyaminic lateral chains potentially involved in ionic pairs with the DNA molecule. Consistently, AN169 has a lower stoichiometry which we can relate to the absence of the polyamine lateral chain. The high analyte concentrations required by ITC

analysis, reinforced the formation of the aspecific binding events and they hinder to appreciate the thermodynamic parameters due to the analyte-ligand interaction. To overcome this problems we shifted to surface plasmon resonance microfluidic system, in which the non-static system analysis and the lower analyte-ligand concentrations required, should reduce the issues of binding aspecificity(Daghestani and Day, 2010)(Murat et al., 2011). By observing the RU/RU_{max} vs ligand concentration curves (fig. 4.14), the occurrence of non-specific bindings is again confirmed at high ligand concentrations, but SPR analysis allowed to calculate the affinity constant by fitting data with the one binding site + non-specific binding equation (fig. 4.13).



$$y = (B_{\max} * x) / (K_d + x) + N_s * x$$

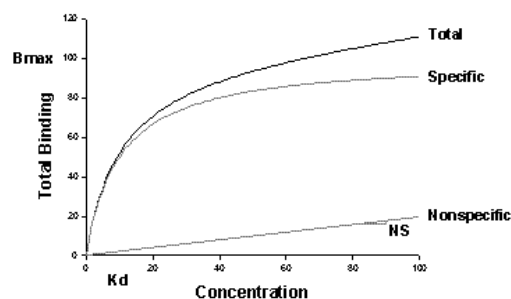


Fig. 4.13: model and equation of the one site saturation+ non specific binding. $N (>0)$ indicate the aspecificity and is the slope of nonspecific binding line

K_a values obtained are reported on table 4III. Also in this case, data are concordant with thermal and ITC analysis by outlining the higher affinity of CM52 and AN169 for telomere

with respect the EGFR-272 sequence but again the higher difference in terms of affinity among the two derivatives on EGFR-272.

	K _a	
	TELOMERE	EGFR-272
AN169	$(2,28 \pm 0,06) \cdot 10^6 \text{ M}^{-1}$	$(0,56 \pm 0,26) \cdot 10^6 \text{ M}^{-1}$
CM52	$(2,76 \pm 0,07) \cdot 10^6 \text{ M}^{-1}$	$(1,08 \pm 0,11) \cdot 10^6 \text{ M}^{-1}$

Table 4III: K_a values obtained by SPR experiment on telomere and EGFR-272 for AN169 and CM52.

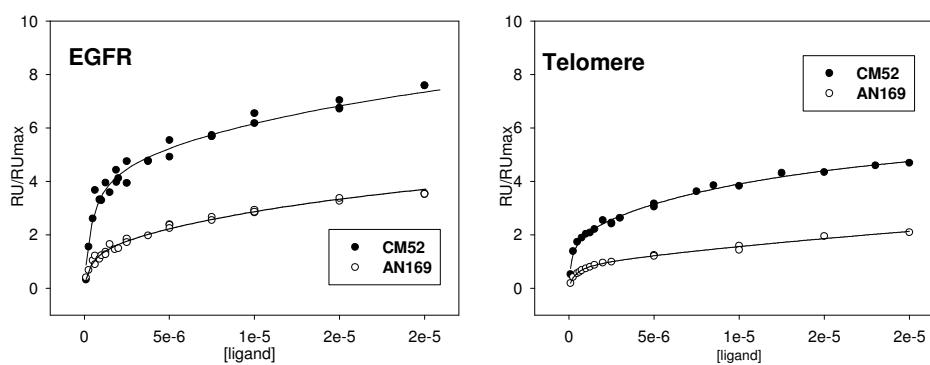


Fig. 4.14: RU/RU_{max} plotted in function of ligand concentrations (CM52 and AN 169). Data were obtained by SPR experiment by testing compounds on telomere and EGFR-272 sequences.

4.4 CONCLUSIONS

All herein tested naphthalene diimides derivatives differ in terms of affinity towards the G4 and the double stranded form but they are all able to bind with a significant affinity the double strand DNA. This aspect correlates with their interference with some nuclear proteins (topoisomerase II α , *taq* polymerase, telomerase) which are involved in cell replicative processes. Indeed, the inhibition of proteins activity was found to be higher for those compounds which bind with more efficiency the double strand. This can be attributed not to a direct interference with the protein, but rather to the binding with DNA.

Despite this drawback, we observed a different binding behavior among the symmetrical and asymmetrical substituents, which was maximal in the case of EGFR-272. In particular, considering CM52 and AN169 as representatives for each class of derivatives, not a significant difference in terms of telomere recognition has been observed. Conversely, on EGFR-272, the symmetry significantly reduce the binding. This was fully confirmed, by comparing the experimental binding affinities towards telomere and EGFR-272. Thus we can conclude that in this case the more efficient binding of CM52 towards EGFR does not derive from a preferential recognition of the oncogenic sequence in comparison to the telomeric one but, conversely to a conservation. One explanation for this lack of selectivity of CM52 might be related the occurrence of aspecific bindings probably due to the high degree of protonation sites in its polyaminic lateral chain. Indeed, our data showed that both compounds share the same binding mode on the single stranded EGFR-272 shifting the parallel G4 conformation towards an antiparallel one. This can explain the not large differences in EGFR affinity emerges among the two compounds.

Finally, we can conclude that, although not a specific ligand has been found to target EGFR-272 sequence, we can assess that the asymmetry of the naphthalene diimide derivatives can be the right approach to develop a selective and high affine compound towards this target. Moreover, the herein collected result, confirmed that also if not efficient rearrangement of the double strand form of EGFR towards non-canonical structures can be observed in physiological conditions, this process is largely speed up in the presence of proper ligands, thus confirming the potential relevance of this approach in anticancer chemotherapy.

Chapter 5

EXPERIMENTAL SECTION

Oligonucleotides were purchased lyophilized from *Metabion International* AG (German), resuspended in milliQ water and then purified by electrophoretic technique:

BRAF-176: 5'-GGG GGT GCG GGG GGG AGC GGG GGA AGG GGG-3';

EGRF-272: 5'-GGG GAC CGG GTC CAG AGG GGC AGT GCT GGG-3';

EGRF-37: 5'-GGG GAG GCA GGG CGG GAG GAG GAG GG-3';

HTS: 5'-AGG GTT AGG GTT AGG GTT AGG G-3';

Tel24: 5'-TTG GGT TAG GGT TAG GGT TAG GGA-3';

Tel26: 5'-AAA GGG TTA GGG TTA GGG TTA GGG AA-3';

wTel26 :5'-TTA GGG TTA GGG TTA GGG TTA GGG AT-3'.

Oligonucleotide sequences labelled with fluorescein at 3' and Dabcyl at 5' termini were purchased lyophilized from *Eurogentech* (Belgium) as HPLC purified products.

Calf thymus DNA (ctDNA) was purchased by *Sigma Aldrich* (USA) and used with no further purification.

Antraquinones were synthesized by Prof. Mauro Freccero from *Università degli Studi di Pavia*.

Naphthalene diimides were synthesized by Prof. Vincenzo Tumiatti and Prof. Anna Minarini from *Università degli Studi di Bologna*.

5.1 Spectroscopic techniques

5.1.1 Fluorescence melting assay

Fluorescence melting curves were determined in a Roche LightCycler480 (λ_{exc} 488 nm, λ_{em} 520 nm) in a total reaction volume of 20 μl containing 0.5 μM of the tested sequences and increasing ligands concentrations in LiP buffer (10 mM LiOH, 50 mM

KCl, pH 7.0 with H₃PO₄). In a typical experiment the oligonucleotide was first denatured by heating to 95° at a rate of 0.1°C s⁻¹, maintained at 95°C for 5 min and then annealed by cooling to 30°C at a rate of 0.1°C s⁻¹. Then, samples were maintained at 30° for 5 min before being slowly heated to 95°C (1°C/ min) and annealed at a rate of 1°C/min. When double stranded DNA was used, the leading strand and the complementary one were mixed at equimolar concentrations, heated to 95°C for 5 min, and then cooled to room temperature overnight before use. Recordings were taken during both the annealing and melting steps.

T_m values were determined from the first derivatives of the melting profiles using the Roche LightCycler software. Each curve was repeated at least three times and errors were ± 0.4°C. ΔT_m were calculated by subtracting the T_m value recorded in the presence of the ligand from the corresponding value in the absence of ligand.

5.1.2 Circular dichroism (CD) measurements

Circular dichroism spectra from 230 to 320 nm were recorded using 10 mm path length cells on a Jasco J 810 spectropolarimeter equipped with a NESLAB temperature controller and interfaced to a PC 100 in 10 mM Tris-HCl, 1 mM EDTA pH 8.0. Before data acquisition, oligo solutions (ca. 4 μM) were heated at 95°C for 5 min and left to cool overnight at room temperature. The reported spectrum of each sample represents the average of 3 scans recorded with 1-nm step resolution. Observed ellipticities were converted to mean residue ellipticity [θ] = deg x cm² x dmol⁻¹ (Mol. Ellip.).

5.1.3 Thermal difference spectrum (TDS)

The thermal difference spectrum was obtained by subtracting of the oligonucleotide UV-spectra at 25°C from the one recorded at 95 °C (below and above the oligo melting temperature respectively). The experiments were performed in 10 mM NaCacodilate/100 mM KCl pH7.0 for the G-rich strand and in 10 mM NaCacodilate/100 mM KCl pH5.0 and pH3.0 for the C-rich strand. All the resulting thermal difference

spectra have been normalized to the value of 1 at the maximal intensity. Before the thermal difference spectra, the CD signal of the oligo in the same buffer has been recorded.

5.1.4 Surface Plasmon Resonance (SPR)

Two streptavidine sensor chips were immobilized with 50nM Biotin- TEG-human telomeric sequence (HTS) and Biotin-TEG-EGFR-272 in 1M NaCl, 50mM NaOH and 50mM NaOH,

1M NaCl in 50%isopropanol in a 0.22 μ m filtered buffer, containing Tris (10mM)and KCl 50mM, 0.025% P20. Before immobilization step, oligos were first annealed in the same buffer by boiling for 5 minutes and slowly cooled at room temperature. SPR analysis was conducted with progressive dilutions of compounds CM52 and AN169 in the running buffer (50 μ M-0). As the compounds were solved in dimethyle sulfoxyde (DMSO), also the contribution of the organic solvent was evaluated and then subtracted. Two start up cycles were performed before each analysis and a 60 seconds time of contact was applied. Finally, a 30 seconds regeneration was performed with 10mM glycine pH 2.5 followed by a 60 seconds stabilization period before each injection.

5.1.5 Isothermal Titration Calorimetry (ITC)

Oligonucleotides were purified and dialyzed in (0.22 μ m) Tris (10mM)and KCl 50mM pH 7.5. The exact concentration of DNA to use as analyte was first calculated through a UV-vis measurement using a NanoDrop 1000 (Thermo Scientific). Ligands CM52 and AN169 were weighted and solved in the same buffer and the exact concentration has been calculated through weight method. Samples were degassed before each experiment and as a control a water-water titration has been performed. Furthermore, a titration buffer on ligand has been performed to calculate the heat of dilution and to be subtracted at each titration curve. Data were analyzed using a

one site binding fitting equation reporting the exact concentrations of ligands and analyte.

5.2 Electrophoretic techniques

5.2.1 Electromobility Shift Assay

³²P end-labelled single-stranded oligonucleotides were obtained by incubation the oligonucleotides with T4 polynucleotide Kinase (*M-Medical S.r.l.*, Italy) and [γ -³²P] ATP (*Perkin Elmer S.p.a.*, Italy) for 30 min at 37°C. The enzyme was then removed by extraction with phenol/chloroform/isoamyl alcohol (25:24:1) (*Sigma Aldrich*). A mixture of purified labeled and unlabelled oligonucleotides (total final concentration 1 μ M) was heated to 95°C for 5 min in 10 mM Tris, 1 mM EDTA, pH 8.0 buffer at increasing KCl concentrations (0-2,5-10-50-150-300mM) and let to cool overnight at room temperature. The folding of the starting material was monitored by native 20% polyacrylamide gel electrophoresis in 0.5X TBE (44.5 mM Tris base, 44.5 mM boric acid and 1 mM Na₂EDTA) added of KCl (10 mM in buffer and in gels). Resolved bands were visualized and quantified on a Phosphor Imager (STORM 840, *Pharmacia Biotech Amersham*).

5.2.2 Topoisomerase inhibition

0.125 μ g of pBR322 (*Inspiralis*) were incubated with increasing concentrations (0.5-100 μ M) of tested compounds (CM28, CM32, CM52, AN82 and AN169) for 1hour at 37 °C n the presence/absence of 1 U of human topoisomerase II (*Fermentas*) in the required buffer. Reaction products were resolved on a 1% agarose gel prepared in 1X TAE (10/1mM Tris-EDTA, 0.1% acetic acid pH 8.0). When the topoisomerase activity was tested, the gel and the buffer were added of 0.1% SDS (Sodium Dodecyl Sulphate, *Sigma*) and a 200mM NaCl (*Sigma*) loading buffer (10/1 mM Tris-EDTA; 200mM NaCl, bromophenol blue, xylene cyanol, pH 8.0) was used. The electrophoretic run was

conducted at 80 Volts for 90 minutes and then the bands were stained with ethidium bromide (0.5 µg/mL in H₂O).

5.2.3 DNA polymerase stop assay

$\gamma^{32}\text{P}$ labeled HT4 primer (96 nM) was annealed with 48nM non-labeled template HT4 in 100/500mM Tris-HCl/KCl. Then, primer elongation process was initiated by adding Taq polymerase (2.5 U) and 100 µM dNTPs, in the presence of increasing ligand concentration. The reaction was incubated at 55°C for 30 minutes and Taq polymerase inactivated with a fast cooling in ice. Samples were dried and then loaded into a 12% denaturing polyacrylamide sequencing gel.

5.2.4 Taq polymerase inhibition assay

The 1065-906 sequence of plasmid pBR322 was amplified through 25 cycles PCR by using the primers 3'GYRA (5' TGA GGA TCC GCC TGG ACA GCA TGG 3') and 5'GYRA (5' GTC GAA TTC TCG GCG AGA AGC AGG 3') at 0.5 µM concentration in the presence/absence of increasing concentrations of tested derivatives (0; 0.25; 0.5; 1; 2.5; 5; 10; 20 µM). The reaction was carried out in a *Eppendorf* thermocycler and the following PCR cycles were performed: 30 seconds at 94°C; 30 seconds at 58°C and 30 seconds at 72 °C. 0.25 µg of plasmid and 2 U of *Taq* polymerase enzyme (*Fermentas*) were used. PCR products were resolved on a 2% agarose gel in 0.5 X TBE (70 Volts for 1.5 hours) and stained by ethidium bromide.

5.2.5 Telomerase Repeat Amplification Protocol (TRAP) assay

Telomerase activity was assayed using a modified Telomere Repeat Amplification Protocol (TRAP) assay. 100ng of the substrate TS (5' AAT CCG TCG AGC AGA GTT 3') was elongated by telomerase by incubation of the reaction mixture with 1.5 µg of protein extract at 37°C for 30min in the presence/absence of increasing drug concentrations. Then products amplification was performed by addition of 100ng of

reverse primer ACX (5' GCG CGG CTT ACC CTT ACC CTT ACC CTA ACC 3') and 2U *Taq* polymerase. Then, PCR amplification step was performed (33 cycles of 30 sec at 92°C, 30 sec at 58°C and 45 sec at 72 °C). The reaction products were loaded onto a 10% polyacrylamide gel in 0.5X TBE and visualized after Sybr Green I staining.

5.2.6 Adducts formation

2.5 μ M of 3'FAM-labelled 22 bases scramble oligonucleotide (5'GGA TGT GAG TGT GAG TGT GAG G 3') was incubated at increasing concentrations of AQs compounds (0-0.1-1-10-50 μ m) for 1hour at room temperature. The reaction was also conducted at 50°C for 24hours, at pH 9.5 for 24 hours and for 1 hour in the presence of 2mM DTT (dithiotreitol). Samples were then loaded on a denaturing 20% polyacrilamide sequencing gel with 7M urea in 1X TBE. Resolved bands were visualized and quantified on a Phosphor Imager (STORM 840, *Pharmacia Biotech Amersham*).

5.2.7 DMS-treatment and LM-PCR (Ligation Mediated-Polymerase Chain Reaction)

Del4 plasmid was extracted from transformed *E.coli* (overnight at 37°C in LB Broth, then extraction procedure was applied). 2 μ g of the extracted plasmid was then incubated overnight at 37°C with and without 200mM KCl. The plasmid was then treated with DMS (4:1 DMS:Ethanol vol/vol) for 10min at room temperature in the presence of 1 μ L of calf thymus DNA (ctDNA) (0,1 μ g/ μ L). Also a control without DMS treatment was considered. The reaction was then stopped with 18 μ L of a mixture of 3M β -mercaptoethanol: H₂O: NaOAc (1:6:7 vol/vol). DNA was then ethanol precipitated and dried. The pellet was resuspended in 10% piperidine and incubated for 30min at 90°C, then precipitated again. After this step, LM-PCR was carried on. PCR (Polymerase Chain Reaction) and T4 ligation reaction steps are reported in detail below.

As regards in vivo DMS treatment, we used two different cell lines: HEK-293 and Raji cells. The first ones for Del4 transfection and the latter ones to work on genomic DNA. In this case, we induced G4 by treating cells with a G4 stabilizer, GQC-05 and a G4 inducer SN-38. $2,5 * 10^5$ cells/well of HEK-293 were plated in a 6 wells plate then transfect with Del4 plasmid ($2\mu\text{g/w}$). They were treated with $1\mu\text{M}$ GQC-05 and $2\mu\text{M}$ SN-38 and incubated for 24hours. Raji cells were treated in solution. A 0.1% DMS diluted in the growth media (DMEM for HEK and RPMI for Raji) was added at each well and incubated for 5min and 10min for HEK-293 (Human Embryonic Kidney) and Raji (Burkitt's lymphoma). The reaction was stopped by diluting 10-fold the cell suspension in cold PBS (phosphate saline buffer) and cells collected by centrifugation. DNA was then extracted and quantified. An agarose gel was performed as a check for the DMS reaction. The average fragment size should be <500 nucleotides as it means that DMS treatment has been efficient [genes and development 2014]. Then, piperidine treatment and LM-PCR was carried on. For each step of PCR, a $1\mu\text{M}$ primer final concentration was used.

PCR1: 5min at 95°C

5min at 64°C

10min at 72°C

$1\mu\text{M}$ P1

1X final concentration Dream Taq mix

$45\mu\text{g}$ DNA

T4 catalyzed ligation reaction:

$15\mu\text{L}$ PCR1 products

$15\mu\text{L}$ annealed primers $25\mu\text{M}$ (linker short and P2f)

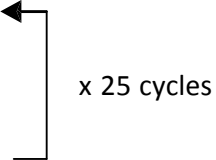
1X final concentration ligation buffer

2 μL T4 ligase

Final volume of 40 μL reached with water.

PCR2: 5min at 95°C

30sec 95°C
30sec 56°C
15sec 72°C



x 25 cycles

4°C ∞

8 μL PCR1 products

1 μM final P2f

1 μM final P2r

1X Dream Taq mix

PCR3: 3.5min 95°C

2min 60°C

10min 72°C

1min 95°C

2min 60°C

10min 72°C

9 μL PCR2 products

1 μM final P3xR/ P3x2

1X Dream Taq mix

After LM-PCR, the amplified DNA was again precipitated, dried, resuspended in 8 μL 80% formamide loading dye, heated at 95°C for 5min and fast cooled in ice.

Samples were then loaded in a 16% denaturing sequencing gel (7M urea, 1X Tris-Borate-EDTA).

5.3 Cellular experiments.

5.3.1 Protein extract preparation

An aliquot of 5×10^6 HeLa cells in exponential phase of growth was pelleted and lysed for 30 minutes in ice using 100 μ L of 0.5% of the zwitterionic detergent CHAPS (3-[[3-Cholamidopropyl]dimethylammonio]-1-propanesulfonate), 1mM EGTA (ethylene glycol tetraacetic acid), 25% 2-mercaptoethanol, 1.74% PMSF (phenylmethylsulfonyl fluoride) and 10% w/w glycerol. The lysate was centrifuged at 13000 rpm for 30 minutes at 4°C and the supernatant collected, quantified by Bradford assay and stored at -80°C.

5.3.2 MTT assay

HeLa (human hepitelial) cell line was maintained in DMEM medium supplemented with 10% heat-inactivated foetal bovine serum, 50 U/ml of penicillin G and 50 μ g/ml of streptomycin, at 37°C in humidified atmosphere and 5% of CO₂. To evaluate toxic profiles of the potential antitelomeric compounds, MTT assays were performed as described: cells were plated in 96 well plates at 10.000 cells/well, and cultured overnight. Afterwards, compounds were added in triplicate and plates were incubated in presence of the drug for 72 hours. At the end of this period, MTT was added to a final concentration of 0.8 mg/ml, and two additional hours of incubation were performed. After that, medium was aspirated carefully and 150 μ l of DMSO were added per well. Soluble formazan salts were homogenated by manual pipetting and absorbance at 540 nm was read (*BioRad Microplate Reader 680*). Curves consisted in 8 serial dilutions in triplicate in each case (from 0 μ M to 100 μ M), the DMSO contribution was subtracted and results were analyzed as sigmoid dose-response curves. MTT assay was performed on anthraquinones derivatives.

References

1. Alcaro, S.; Musetti, C.; Distinto, S.; Casatti, M.; Zagotto, G.; Artese, A.; Parrotta, L.; Moraca, F.; Costa, G.; Ortuso, F.; Maccioni, E.; Sissi, C. Identification and characterization of new DNA G-quadruplex binders selected by a combination of ligand and structure-based virtual screening approaches. *J. Med. Chem.* **2013**, *56*, 843-855.
2. Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R. A.; Yang, D. Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Res.* **2006**, *34*, 2723-2735.
3. Arteaga, C. L. The epidermal growth factor receptor: from mutant oncogene in nonhuman cancers to therapeutic target in human neoplasia. *J. Clin. Oncol.* **2001**, *19*, 32S-40S.
4. Bacolla, A.; Wells, R. D. Non-B DNA conformations, genomic rearrangements, and human disease. *J. Biol. Chem.* **2004**, *279*, 47411-47414.
5. Balasubramanian, S.; Hurley, L. H.; Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discov.* **2011**, *10*, 261-275.
6. Bandele, O. J.; Osheroff, N. Cleavage of plasmid DNA by eukaryotic topoisomerase II. *Methods Mol. Biol.* **2009**, *582*, 39-47.
7. Bhavsar-Jog, Y. P.; Van Dornshuld, E.; Brooks, T. A.; Tschumper, G. S.; Wadkins, R. M. Epigenetic modification, dehydration, and molecular crowding effects on the thermodynamics of i-motif structure formation from C-rich DNA. *Biochemistry* **2014**, *53*, 1586-1594.
8. Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **2013**, *5*, 182-186.
9. Biffi, G.; Tannahill, D.; Miller, J.; Howat, W. J.; Balasubramanian, S. Elevated levels of G-quadruplex formation in human stomach and liver cancer tissues. *PLoS One* **2014**, *9*, e102711.
10. Biroccio, A.; Porru, M.; Rizzo, A.; Salvati, E.; D'Angelo, C.; Orlandi, A.; Passeri, D.; Franceschin, M.; Stevens, M. F.; Gilson, E.; Beretta, G.; Zupi, G.; Pisano, C.; Zunino, F.; Leonetti, C. DNA damage persistence as determinant of tumor sensitivity to the combination of Topo I inhibitors and telomere-targeting agents. *Clin. Cancer Res.* **2011**, *17*, 2227-2236.
11. Brazda, V.; Laister, R. C.; Jagelska, E. B.; Arrowsmith, C. Cruciform structures are a common DNA feature important for regulating biological processes. *BMC Mol. Biol.* **2011**, *12*, 33-2199-12-33.

12. Brooks, T. A.; Hurley, L. H. Targeting MYC Expression through G-Quadruplexes. *Genes Cancer*. **2010**, *1*, 641-649.
13. Brooks, T. A.; Hurley, L. H. The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics. *Nat. Rev. Cancer*. **2009**, *9*, 849-861.
14. Brown, R. V.; Danford, F. L.; Gokhale, V.; Hurley, L. H.; Brooks, T. A. Demonstration that drug-targeted down-regulation of MYC in non-Hodgkins lymphoma is directly mediated through the promoter G-quadruplex. *J. Biol. Chem*. **2011**, *286*, 41018-41027.
15. Burden, D. A.; Osheroff, N. Mechanism of action of eukaryotic topoisomerase II and drugs targeted to the enzyme. *Biochim. Biophys. Acta* **1998**, *1400*, 139-154.
16. Buscaglia, R.; Gray, R. D.; Chaires, J. B. Thermodynamic characterization of human telomere quadruplex unfolding. *Biopolymers* **2013**, *99*, 1006-1018.
17. Buscaglia, R.; Miller, M. C.; Dean, W. L.; Gray, R. D.; Lane, A. N.; Trent, J. O.; Chaires, J. B. Polyethylene glycol binding alters human telomere G-quadruplex structure by conformational selection. *Nucleic Acids Res*. **2013**, *41*, 7934-7946.
18. Buurma, N. J.; Haq, I. Advances in the analysis of isothermal titration calorimetry data for ligand-DNA interactions. *Methods* **2007**, *42*, 162-172.
19. Cairns, D.; Michalitsi, E.; Jenkins, T. C.; Mackay, S. P. Molecular modelling and cytotoxicity of substituted anthraquinones as inhibitors of human telomerase. *Bioorg. Med. Chem*. **2002**, *10*, 803-807.
20. Cech, T. R. Beginning to understand the end of the chromosome. *Cell* **2004**, *116*, 273-279.
21. Chung, W. J.; Heddi, B.; Hamon, F.; Teulade-Fichou, M. P.; Phan, A. T. Solution structure of a G-quadruplex bound to the bisquinolinium compound Phen-DC(3). *Angew. Chem. Int. Ed Engl*. **2014**, *53*, 999-1002.
22. Clark, G. R.; Pytel, P. D.; Squire, C. J.; Neidle, S. Structure of the first parallel DNA quadruplex-drug complex. *J. Am. Chem. Soc*. **2003**, *125*, 4066-4067.
23. Collie, G. W.; Parkinson, G. N. The application of DNA and RNA G-quadruplexes to therapeutic medicines. *Chem. Soc. Rev*. **2011**, *40*, 5867-5892.

24. Collie, G. W.; Promontorio, R.; Hampel, S. M.; Micco, M.; Neidle, S.; Parkinson, G. N. Structural basis for telomeric G-quadruplex targeting by naphthalene diimide ligands. *J. Am. Chem. Soc.* **2012**, *134*, 2723-2731.
25. Daghestani, H. N.; Day, B. W. Theory and applications of surface plasmon resonance, resonant mirror, resonant waveguide grating, and dual polarization interferometry biosensors. *Sensors (Basel)* **2010**, *10*, 9630-9646.
26. Dai, J.; Carver, M.; Punchihewa, C.; Jones, R. A.; Yang, D. 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.* **2007**, *35*, 4927-4940.
27. Dailey, M. M.; Miller, M. C.; Bates, P. J.; Lane, A. N.; Trent, J. O. Resolution and characterization of the structural polymorphism of a single quadruplex-forming sequence. *Nucleic Acids Res.* **2010**, *38*, 4877-4888.
28. Darby, R. A.; Sollogoub, M.; McKeen, C.; Brown, L.; Risitano, A.; Brown, N.; Barton, C.; Brown, T.; Fox, K. R. High throughput measurement of duplex, triplex and quadruplex melting curves using molecular beacons and a LightCycler. *Nucleic Acids Res.* **2002**, *30*, e39.
29. Day, H. A.; Huguin, C.; Waller, Z. A. Silver cations fold i-motif at neutral pH. *Chem. Commun. (Camb)* **2013**, *49*, 7696-7698.
30. Day, H. A.; Pavlou, P.; Waller, Z. A. i-Motif DNA: structure, stability and targeting with ligands. *Bioorg. Med. Chem.* **2014**, *22*, 4407-4418.
31. de Rosa, M.; de Sanctis, D.; Rosario, A. L.; Archer, M.; Rich, A.; Athanasiadis, A.; Carrondo, M. A. Crystal structure of a junction between two Z-DNA helices. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 9088-9092.
32. Di Antonio, M.; Biffi, G.; Mariani, A.; Raiber, E. A.; Rodriguez, R.; Balasubramanian, S. Selective RNA versus DNA G-quadruplex targeting by in situ click chemistry. *Angew. Chem. Int. Ed Engl.* **2012**, *51*, 11073-11078.
33. Di Antonio, M.; Doria, F.; Mella, M.; Merli, D.; Profumo, A.; Freccero, M. Novel naphthalene diimides as activatable precursors of bisalkylating agents, by reduction and base catalysis. *J. Org. Chem.* **2007**, *72*, 8354-8360.
34. Doria, F.; Nadai, M.; Folini, M.; Di Antonio, M.; Germani, L.; Percivalle, C.; Sissi, C.; Zaffaroni, N.; Alcaro, S.; Artese, A.; Richter, S. N.; Freccero, M. Hybrid ligand-alkylating agents targeting telomeric G-quadruplex structures. *Org. Biomol. Chem.* **2012**, *10*, 2798-2806.

35. Doria, F.; Nadai, M.; Folini, M.; Scalabrin, M.; Germani, L.; Sattin, G.; Mella, M.; Palumbo, M.; Zaffaroni, N.; Fabris, D.; Freccero, M.; Richter, S. N. Targeting loop adenines in G-quadruplex by a selective oxirane. *Chemistry* **2013**, *19*, 78-81.
36. Frank-Kamenetskii, M. D.; Mirkin, S. M. Triplex DNA structures. *Annu. Rev. Biochem.* **1995**, *64*, 65-95.
37. Freyer, M. W.; Lewis, E. A. Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions. *Methods Cell Biol.* **2008**, *84*, 79-113.
38. Gatto, B.; Zagotto, G.; Sissi, C.; Cera, C.; Uriarte, E.; Palu, G.; Capranico, G.; Palumbo, M. Peptidyl anthraquinones as potential antineoplastic drugs: synthesis, DNA binding, redox cycling, and biological activity. *J. Med. Chem.* **1996**, *39*, 3114-3122.
39. Gellert, M.; Lipsett, M. N.; Davies, D. R. Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. U. S. A.* **1962**, *48*, 2013-2018.
40. Giancola, C.; Pagano, B. Energetics of ligand binding to G-quadruplexes. *Top. Curr. Chem.* **2013**, *330*, 211-242.
41. Gorsche, R.; Jovanovic, B.; Gudynaite-Savitch, L.; Mach, R. L.; Mach-Aigner, A. R. A highly sensitive in vivo footprinting technique for condition-dependent identification of cis elements. *Nucleic Acids Res.* **2014**, *42*, e1.
42. Grand, C. L.; Han, H.; Munoz, R. M.; Weitman, S.; Von Hoff, D. D.; Hurley, L. H.; Bearss, D. J. The cationic porphyrin TMPyP4 down-regulates c-MYC and human telomerase reverse transcriptase expression and inhibits tumor growth in vivo. *Mol. Cancer. Ther.* **2002**, *1*, 565-573.
43. Grand, C. L.; Powell, T. J.; Nagle, R. B.; Bearss, D. J.; Tye, D.; Gleason-Guzman, M.; Hurley, L. H. Mutations in the G-quadruplex silencer element and their relationship to c-MYC overexpression, NM23 repression, and therapeutic rescue. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 516.
44. Greider, C. W.; Blackburn, E. H. A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. *Nature* **1989**, *337*, 331-337.
45. Guedin, A.; Alberti, P.; Mergny, J. L. Stability of intramolecular quadruplexes: sequence effects in the central loop. *Nucleic Acids Res.* **2009**, *37*, 5559-5567.

46. Guedin, A.; Gros, J.; Alberti, P.; Mergny, J. L. How long is too long? Effects of loop size on G-quadruplex stability. *Nucleic Acids Res.* **2010**, *38*, 7858-7868.
47. Haider, S. M.; Neidle, S.; Parkinson, G. N. A structural analysis of G-quadruplex/ligand interactions. *Biochimie* **2011**, *93*, 1239-1251.
48. Hampel, S. M.; Pepe, A.; Greulich-Bode, K. M.; Malhotra, S. V.; Reszka, A. P.; Veith, S.; Boukamp, P.; Neidle, S. Mechanism of the antiproliferative activity of some naphthalene diimide G-quadruplex ligands. *Mol. Pharmacol.* **2013**, *83*, 470-480.
49. Han, H.; Hurley, L. H.; Salazar, M. A DNA polymerase stop assay for G-quadruplex-interactive compounds. *Nucleic Acids Res.* **1999**, *27*, 537-542.
50. Hanahan, D.; Weinberg, R. A. The hallmarks of cancer. *Cell* **2000**, *100*, 57-70.
51. Hazel, P.; Huppert, J.; Balasubramanian, S.; Neidle, S. Loop-length-dependent folding of G-quadruplexes. *J. Am. Chem. Soc.* **2004**, *126*, 16405-16415.
52. http://www.chem.cmu.edu/groups/army/index_files/Page855.html .
53. Huppert, J. L.; Balasubramanian, S. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* **2007**, *35*, 406-413.
54. Huppert, J. L.; Balasubramanian, S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* **2005**, *33*, 2908-2916.
55. Huppert, J. L.; Bugaut, A.; Kumari, S.; Balasubramanian, S. G-quadruplexes: the beginning and end of UTRs. *Nucleic Acids Res.* **2008**, *36*, 6260-6268.
56. Jain, A. K.; Bhattacharya, S. Groove binding ligands for the interaction with parallel-stranded ps-duplex DNA and triplex DNA. *Bioconjug. Chem.* **2010**, *21*, 1389-1403.
57. Kang, H. J.; Kendrick, S.; Hecht, S. M.; Hurley, L. H. 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.* **2014**, *136*, 4172-4185.
58. Kendrick, S.; Hurley, L. H. The role of G-quadruplex/i-motif secondary structures as cis-acting regulatory elements. *Pure Appl. Chem.* **2010**, *82*, 1609-1621.

59. Keniry, M. A.; Owen, E. A. Insight into the molecular recognition of spermine by DNA quadruplexes from an NMR study of the association of spermine with the thrombin-binding aptamer. *J. Mol. Recognit.* **2013**, *26*, 308-317.
60. Kufe DW, Pollock RE, Weichselbaum RR, et al. Mechanism of oncogene activation. *Holland-Frei Cancer Medicine. 6th edition* **2003**.
61. Lane, A. N.; Chaires, J. B.; Gray, R. D.; Trent, J. O. Stability and kinetics of G-quadruplex structures. *Nucleic Acids Res.* **2008**, *36*, 5482-5515.
62. Lannan, F. M.; Mamajanov, I.; Hud, N. V. Human telomere sequence DNA in water-free and high-viscosity solvents: G-quadruplex folding governed by Kramers rate theory. *J. Am. Chem. Soc.* **2012**, *134*, 15324-15330.
63. Li, J.; Correia, J. J.; Wang, L.; Trent, J. O.; Chaires, J. B. Not so crystal clear: the structure of the human telomere G-quadruplex in solution differs from that present in a crystal. *Nucleic Acids Res.* **2005**, *33*, 4649-4659.
64. Li, W.; Wu, P.; Ohmichi, T.; Sugimoto, N. Characterization and thermodynamic properties of quadruplex/duplex competition. *FEBS Lett.* **2002**, *526*, 77-81.
65. Lim, K. W.; Amrane, S.; Bouaziz, S.; Xu, W.; Mu, Y.; Patel, D. J.; Luu, K. N.; Phan, A. T. Structure of the human telomere in K⁺ solution: a stable basket-type G-quadruplex with only two G-tetrad layers. *J. Am. Chem. Soc.* **2009**, *131*, 4301-4309.
66. Lin, P. H.; Chen, R. H.; Lee, C. H.; Chang, Y.; Chen, C. S.; Chen, W. Y. Studies of the binding mechanism between aptamers and thrombin by circular dichroism, surface plasmon resonance and isothermal titration calorimetry. *Colloids Surf. B Biointerfaces* **2011**, *88*, 552-558.
67. Lipps, H. J.; Rhodes, D. G-quadruplex structures: in vivo evidence and function. *Trends Cell Biol.* **2009**, *19*, 414-422.
68. Liu, Y.; Collar, C. J.; Kumar, A.; Stephens, C. E.; Boykin, D. W.; Wilson, W. D. Heterocyclic diamidine interactions at AT base pairs in the DNA minor groove: effects of heterocycle differences, DNA AT sequence and length. *J Phys Chem B* **2008**, *112*, 11809-11818.
69. Mergny, J. L.; Lacroix, L. UV Melting of G-Quadruplexes. *Curr. Protoc. Nucleic Acid Chem.* **2009**, *Chapter 17*, Unit 17.1.

70. Mergny, J. L.; Li, J.; Lacroix, L.; Amrane, S.; Chaires, J. B. Thermal difference spectra: a specific signature for nucleic acid structures. *Nucleic Acids Res.* **2005**, *33*, e138.
71. Milelli, A.; Tumiatti, V.; Micco, M.; Rosini, M.; Zuccari, G.; Raffaghello, L.; Bianchi, G.; Pistoia, V.; Fernando Diaz, J.; Pera, B.; Trigili, C.; Barasoain, I.; Musetti, C.; Toniolo, M.; Sissi, C.; Alcaro, S.; Moraca, F.; Zini, M.; Stefanelli, C.; Minarini, A. Structure-activity relationships of novel substituted naphthalene diimides as anticancer agents. *Eur. J. Med. Chem.* **2012**, *57*, 417-428.
72. Miller, M. C.; Buscaglia, R.; Chaires, J. B.; Lane, A. N.; Trent, J. O. Hydration is a major determinant of the G-quadruplex stability and conformation of the human telomere 3' sequence of d(AG3(TTAG3)3). *J. Am. Chem. Soc.* **2010**, *132*, 17105-17107.
73. Miyoshi, D.; Sugimoto, N. Molecular crowding effects on structure and stability of DNA. *Biochimie* **2008**, *90*, 1040-1051.
74. Modica, E.; Zanaletti, R.; Freccero, M.; Mella, M. Alkylation of amino acids and glutathione in water by o-quinone methide. Reactivity and selectivity. *J. Org. Chem.* **2001**, *66*, 41-52.
75. Monneret, C. Recent developments in the field of antitumour anthracyclines. *Eur. J. Med. Chem.* **2001**, *36*, 483-493.
76. Murat, P.; Singh, Y.; Defrancq, E. Methods for investigating G-quadruplex DNA/ligand interactions. *Chem. Soc. Rev.* **2011**, *40*, 5293-5307.
77. Nadai, M.; Doria, F.; Di Antonio, M.; Sattin, G.; Germani, L.; Percivalle, C.; Palumbo, M.; Richter, S. N.; Freccero, M. Naphthalene diimide scaffolds with dual reversible and covalent interaction properties towards G-quadruplex. *Biochimie* **2011**, *93*, 1328-1340.
78. Ndlebe, T.; Schuster, G. B. Long-distance radical cation transport in DNA: Horizontal charge hopping in a dimeric quadruplex. *Org. Biomol. Chem.* **2006**, *4*, 4015-4021.
79. Neidle, S.; Balasubramanian, S. Quadruplex Nucleic Acids. **2006**, *1*, 1-30.
80. Neidle, S.; Parkinson, G. N. Quadruplex DNA crystal structures and drug design. *Biochimie* **2008**, *90*, 1184-1196.
81. Paramasivan, S.; Rujan, I.; Bolton, P. H. Circular dichroism of quadruplex DNAs: applications to structure, cation effects and ligand binding. *Methods* **2007**, *43*, 324-331.

82. Parkinson, G. N.; Cuenca, F.; Neidle, S. Topology conservation and loop flexibility in quadruplex-drug recognition: crystal structures of inter- and intramolecular telomeric DNA quadruplex-drug complexes. *J. Mol. Biol.* **2008**, *381*, 1145-1156.
83. Parkinson, G. N.; Lee, M. P.; Neidle, S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **2002**, *417*, 876-880.
84. Patel, D. J.; Phan, A. T.; Kuryavyi, V. Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res.* **2007**, *35*, 7429-7455.
85. Perry, P. J.; Read, M. A.; Davies, R. T.; Gowan, S. M.; Reszka, A. P.; Wood, A. A.; Kelland, L. R.; Neidle, S. 2,7-Disubstituted amidofluorenone derivatives as inhibitors of human telomerase. *J. Med. Chem.* **1999**, *42*, 2679-2684.
86. Perry, P. J.; Reszka, A. P.; Wood, A. A.; Read, M. A.; Gowan, S. M.; Dosanjh, H. S.; Trent, J. O.; Jenkins, T. C.; Kelland, L. R.; Neidle, S. Human telomerase inhibition by regioisomeric disubstituted amidoanthracene-9,10-diones. *J. Med. Chem.* **1998**, *41*, 4873-4884.
87. Petraccone, L.; Malafronte, A.; Amato, J.; Giancola, C. G-quadruplexes from human telomeric DNA: how many conformations in PEG containing solutions? *J Phys Chem B* **2012**, *116*, 2294-2305.
88. Petraccone, L.; Pagano, B.; Giancola, C. Studying the effect of crowding and dehydration on DNA G-quadruplexes. *Methods* **2012**, *57*, 76-83.
89. Pfeifer, G. P.; Tanguay, R. L.; Steigerwald, S. D.; Riggs, A. D. In vivo footprint and methylation analysis by PCR-aided genomic sequencing: comparison of active and inactive X chromosomal DNA at the CpG island and promoter of human PGK-1. *Genes Dev.* **1990**, *4*, 1277-1287.
90. Pommier, Y.; Leo, E.; Zhang, H.; Marchand, C. DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. *Chem. Biol.* **2010**, *17*, 421-433.
91. Randazzo, A.; Spada, G. P.; da Silva, M. W. Circular dichroism of quadruplex structures. *Top. Curr. Chem.* **2013**, *330*, 67-86.
92. Risitano, A.; Fox, K. R. Stability of intramolecular DNA quadruplexes: comparison with DNA duplexes. *Biochemistry* **2003**, *42*, 6507-6513.

93. Rodriguez, R.; Miller, K. M.; Forment, J. V.; Bradshaw, C. R.; Nikan, M.; Britton, S.; Oelschlaegel, T.; Xhemalce, B.; Balasubramanian, S.; Jackson, S. P. Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nat. Chem. Biol.* **2012**, *8*, 301-310.
94. Rodriguez, R.; Muller, S.; Yeoman, J. A.; Trentesaux, C.; Riou, J. F.; Balasubramanian, S. A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. *J. Am. Chem. Soc.* **2008**, *130*, 15758-15759.
95. Sacca, B.; Lacroix, L.; Mergny, J. L. The effect of chemical modifications on the thermal stability of different G-quadruplex-forming oligonucleotides. *Nucleic Acids Res.* **2005**, *33*, 1182-1192.
96. Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11593-11598.
97. Sissi, C.; Palumbo, M. Telomeric G-quadruplex architecture and interactions with potential drugs. *Curr. Pharm. Des.* **2014**, *20*, 6489-6509.
98. Strachan, T.; Read, A. P. **1999**.
99. Sun, D.; Hurley, L. H. Biochemical techniques for the characterization of G-quadruplex structures: EMSA, DMS footprinting, and DNA polymerase stop assay. *Methods Mol. Biol.* **2010**, *608*, 65-79.
100. Sun, D.; Thompson, B.; Cathers, B. E.; Salazar, M.; Kerwin, S. M.; Trent, J. O.; Jenkins, T. C.; Neidle, S.; Hurley, L. H. Inhibition of human telomerase by a G-quadruplex-interactive compound. *J. Med. Chem.* **1997**, *40*, 2113-2116.
101. Sun, H.; Xiang, J.; Liu, Y.; Li, L.; Li, Q.; Xu, G.; Tang, Y. A stabilizing and denaturing dual-effect for natural polyamines interacting with G-quadruplexes depending on concentration. *Biochimie* **2011**, *93*, 1351-1356.
102. Sun, W.; Yang, J. Functional mechanisms for human tumor suppressors. *J. Cancer.* **2010**, *1*, 136-140.
103. Tanious, F. A.; Nguyen, B.; Wilson, W. D. Biosensor-surface plasmon resonance methods for quantitative analysis of biomolecular interactions. *Methods Cell Biol.* **2008**, *84*, 53-77.

104. van Dijk, E. H.; Myles, D. J.; van der Veen, M. H.; Hummelen, J. C. Synthesis and properties of an anthraquinone-based redox switch for molecular electronics. *Org. Lett.* **2006**, *8*, 2333-2336.
105. Watson, J. D.; Crick, F. H. Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. 1953. *Rev. Invest. Clin.* **2003**, *55*, 108-109.
106. Webba da Silva, M. Geometric formalism for DNA quadruplex folding. *Chemistry* **2007**, *13*, 9738-9745.
107. Webba da Silva, M.; Trajkovski, M.; Sannohe, Y.; Ma'ani Hessari, N.; Sugiyama, H.; Plavec, J. Design of a G-quadruplex topology through glycosidic bond angles. *Angew. Chem. Int. Ed Engl.* **2009**, *48*, 9167-9170.
108. Wei, D.; Todd, A. K.; Zloh, M.; Gunaratnam, M.; Parkinson, G. N.; Neidle, S. Crystal structure of a promoter sequence in the B-raf gene reveals an intertwined dimer quadruplex. *J. Am. Chem. Soc.* **2013**, *135*, 19319-19329.
109. Yaku, H.; Fujimoto, T.; Murashima, T.; Miyoshi, D.; Sugimoto, N. Phthalocyanines: a new class of G-quadruplex-ligands with many potential applications. *Chem. Commun. (Camb)* **2012**, *48*, 6203-6216.
110. Zagotto, G.; Ricci, A.; Vasquez, E.; Sandoli, A.; Benedetti, S.; Palumbo, M.; Sissi, C. Tuning G-quadruplex vs double-stranded DNA recognition in regioisomeric lysyl-peptidyl-anthraquinone conjugates. *Bioconjug. Chem.* **2011**, *22*, 2126-2135.
111. Zagotto, G.; Sissi, C.; Lucatello, L.; Pivetta, C.; Cadamuro, S. A.; Fox, K. R.; Neidle, S.; Palumbo, M. Aminoacyl-anthraquinone conjugates as telomerase inhibitors: synthesis, biophysical and biological evaluation. *J. Med. Chem.* **2008**, *51*, 5566-5574.
112. Zagotto, G.; Sissi, C.; Moro, S.; Dal Ben, D.; Parkinson, G. N.; Fox, K. R.; Neidle, S.; Palumbo, M. Amide bond direction modulates G-quadruplex recognition and telomerase inhibition by 2,6 and 2,7 bis-substituted anthracenedione derivatives. *Bioorg. Med. Chem.* **2008**, *16*, 354-361.
113. Zhao, J.; Bacolla, A.; Wang, G.; Vasquez, K. M. Non-B DNA structure-induced genetic instability and evolution. *Cell Mol. Life Sci.* **2010**, *67*, 43-62.
114. Zhou, J.; Wei, C.; Jia, G.; Wang, X.; Feng, Z.; Li, C. Formation of i-motif structure at neutral and slightly alkaline pH. *Mol. Biosyst* **2010**, *6*, 580-586.

-
115. Zikich, D.; Liu, K.; Sagiv, L.; Porath, D.; Kotlyar, A. I-motif nanospheres: unusual self-assembly of long cytosine strands. *Small* **2011**, *7*, 1029-1034.
 116. Zvereva, M. I.; Shcherbakova, D. M.; Dontsova, O. A. Telomerase: structure, functions, and activity regulation. *Biochemistry (Mosc)* **2010**, *75*, 1563-1583.

Appendix

Aryl ethynyl anthraquinones: a useful platform for targeting telomeric G-quadruplex structures†

Cite this: *Org. Biomol. Chem.*, 2014, **12**, 3744

Claudia Percivalle,^a Claudia Sissi,^b Maria Laura Greco,^b Caterina Musetti,^b Angelica Mariani,^a Anna Artese,^c Giosuè Costa,^c Maria Lucia Perrone,^a Stefano Alcaro^c and Mauro Freccero*^a

Received 28th January 2014,
Accepted 4th April 2014

DOI: 10.1039/c4ob00220b

www.rsc.org/obc

Aryl ethynyl anthraquinones have been synthesized by Sonogashira cross-coupling and evaluated as telomeric G-quadruplex ligands, by the FRET melting assay, circular dichroism, the DNA synthesis arrest assay and molecular docking. Both the binding properties and G-quadruplex vs. duplex selectivity are controlled by the structures of the aryl ethynyl moieties.

Introduction

Small molecule-mediated DNA targeting represents one of the most effective approaches for the development of chemotherapeutics. The ability of DNA to fold into highly stable secondary structures could be exploited for the design of anticancer agents interacting with nucleic acids in a sequence or structural selective fashion.^{1,2} One such target is represented by G-quadruplex (G4) DNA and RNA motifs.^{3,4} G4 is a four-stranded nucleic acid structure that can be formed in guanine-rich nucleic acid sequences *via* Hoogsteen hydrogen bond formation and cation coordination.⁵ G4 stabilization has been proposed to interfere with important biological processes for cellular homeostasis, such as DNA damage response activation,^{6–8} oncogene expression^{3,9–11} and genomic stability.¹² Putative quadruplex sequences (PQS) are highly spread in the genome¹³ and transcriptome,¹⁴ including gene promoter regions or gene bodies and telomeres, providing these structures with the potential to act as regulatory elements of different processes. A general lack of evidence of the formation and the existence of G4 *in vivo* and its real biological functions made the G4 relevance as therapeutic targets controversial.¹⁵ However, the existence of G4 structures in cells has recently

been demonstrated by means of immuno-fluorescence staining with an engineered structure-specific antibody¹⁶ and by Chromatin Immuno-Precipitation (ChIP-Seq).⁷ The potential therapeutic opportunities offered by the targeting of these structures prompted the design of a large number of ligands that specifically interact with the terminal tetrads, G4 loops and grooves.¹⁷ In the last two decades, several selective G4 ligands have been reported and in most of the cases they share a large planar aromatic surface that provide the ligands with a π -stacking surface for binding with the external tetrads of the G4.^{17,18} Cationic side chains, at physiological pH, further increase the ligand binding properties, providing an additional electrostatic interaction with the phosphate backbone.¹⁷ Dissecting the function of a specific G4 family over the others could be achieved by developing small molecule ligands that can discriminate not only over duplex DNA but also over different G4 architectures.¹⁹ The development of these compounds would provide unprecedented tools to analyze cells for the functions of G4s present in a specific genomic region.

Anthraquinone derivatives (AQs) represent an interesting scaffold to develop selective and multifunctional G4 ligands, with many potential applications, because of their well characterised DNA-binding properties,^{20,21} fairly low redox potential and their ability to act as photosensitizers by one-electron oxidation.²² Structurally, AQs are strictly related to the anthracycline antibiotics like doxorubicin and daunomycin.^{23–26}

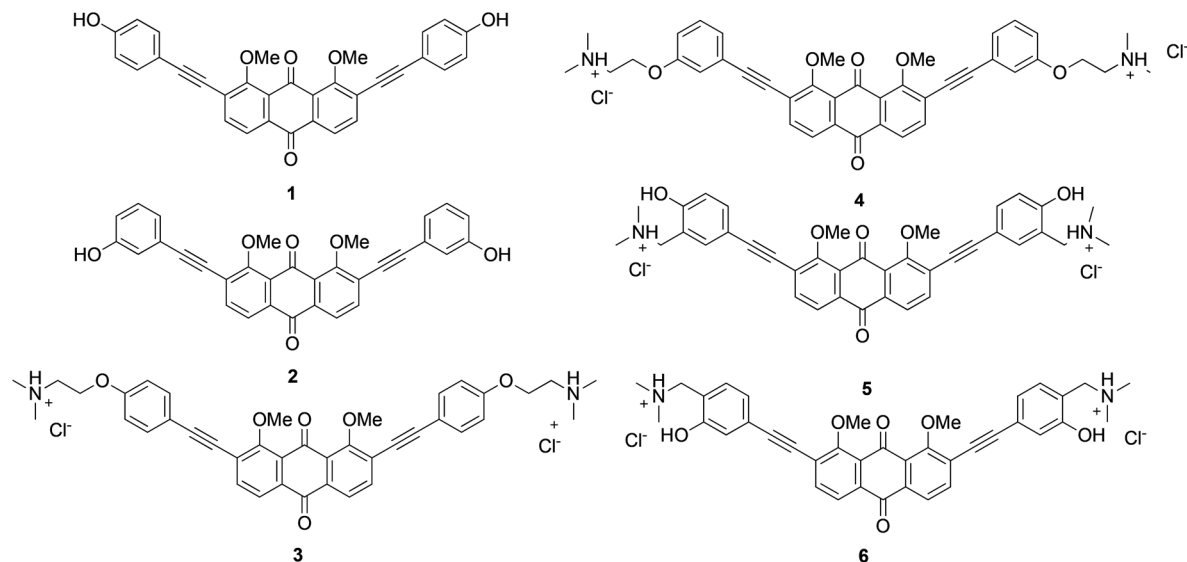
It has been shown that doxorubicin and daunomycin can interact with telomeric DNA *via* G4 stabilization, mediated by the anthraquinone scaffold and demonstrated by the crystal structure of a complex between the telomeric G4 DNA and daunomycin.²⁷ With the aim to optimize G4 recognition the synthesis of several 1,4-1,5-1,8-2,6- and 2,7-difunctionalized amidoanthracene-9,10-diones has been performed and the resulting compounds have been tested as G4 ligands.^{28–31} The five different regio-isomers showed different abilities to recog-

^aDipartimento di Chimica, Università di Pavia, V.le Taramelli 10, 27100 Pavia, Italy. E-mail: claudia.percivalle01@universitadipavia.it, angy88@tiscali.it, marialucia.perrone01@universitadipavia.it, mauro.freccero@unipv.it

^bDepartment of Pharmaceutical and Pharmacological Sciences, v. Marzolo 5, 35131 Padova, Italy. E-mail: claudia.sissi@unipd.it, marialaura.greco@studenti.unipd.it, caterina.musetti@gmail.com

^cDipartimento di Scienze della Salute, Università di Catanzaro, Campus "Salvatore Venuta", Viale Europa, 88100 Catanzaro, Italy. E-mail: alcaro@unicz.it, artese@unicz.it, gcosta@unicz.it

† Electronic supplementary information (ESI) available: Experimental details, procedures and materials, NMR spectra, CDs, computational data, IFD scores and best poses. See DOI: 10.1039/c4ob00220b



Scheme 1 Aryl ethynyl anthraquinones (AQs) synthesized and evaluated as G4 ligands.

nize G4 telomeric structures according to the nature and the position of the substituent side chains. Consistently, conjugation of the anthraquinone core with amino sugars³² or amino acids^{33,34} was applied to modulate their G4 binding properties. One such example is a neomycin–anthraquinone conjugate that exhibits a nanomolar affinity for telomeric G4 DNA, which is 1000-fold higher when compared to its constituent units.³² This higher affinity is ascribed to the dual binding mode of the conjugate which can interact with the grooves (neomycin) and with the guanines of the G4 (anthraquinone) *via* π -stacking interactions.

For AQ-amino acid conjugates, the combination of a basic amino acid (Lys) with a more hydrophobic residue (Phe) has provided a better G4 selectivity *versus* the duplex DNA.³⁴

Unlike the large majority of G4 ligands, AQs exhibit interesting redox properties, as they easily generate radical anions and di-anions by bio-compatible reduction.³⁵ We demonstrated that formation of stable radical anions can be exploited to generate alkylating agents such as Quinone Methides (QMs, generated from *o*- or *p*-benzyl substituted phenols).^{36,37} We anticipate that AQs could be similarly exploited for *in situ* generation of QMs at G4 sites, thus enabling G4 covalent targeting.³⁸ Alkylation has been proposed as an alternative approach to physically lock the DNA G4 in its folded conformation, enabling the investigation of the biological implications associated with G4s stabilization.^{38–42}

Effective electronic conjugation between the AQ core and the aryl moiety (embedding the QM precursor) will ensure generation of the alkylating QM under reductive conditions. Moreover, we postulated that increasing the aromatic surface and the structural constraints by introducing aryl side chains would have been beneficial for both the AQs G4 binding properties and for G4 *vs.* duplex selectivity. Therefore, we explored conjugation of suitable QM precursors to the AQ scaffold introducing ethynyl spacers by means of the Sonogashira cross-

coupling. Our synthetic effort resulted in a small library of aryl ethynyl anthraquinones (AQs, 1–6, Scheme 1). It involved the symmetric functionalization of the anthraquinone core at 2 and 7 positions with chemically diverse aryl moieties, such as negatively charged phenolates arising from 1 and 2, positively charged secondary amines (3,4) and zwitterionic Mannich bases (5,6).

Then, all the ligands here synthesized have been tested for their G4 binding properties in comparison with double stranded DNA. As G4 model sequences, we selected the human telomeric DNA. Telomeres consist of a hexameric nucleotide repeat unit d(TTAGGG) and several four repeat sequences are currently extensively used as mimics, since they allow exploration of different G4 conformations.

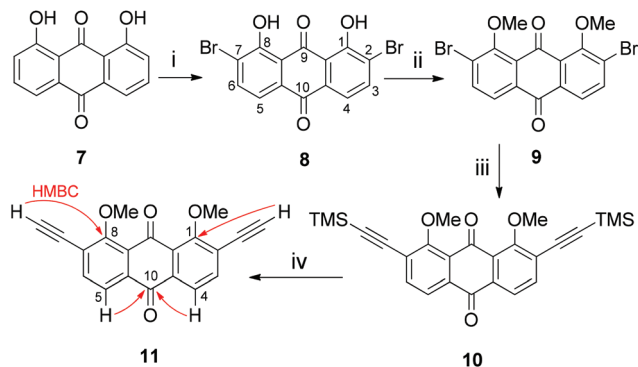
Due to the differential binding properties observed on the selected DNA substrates, the chemical versatility of the Sonogashira cross-coupling involved in the preparation of these compounds, and their redox properties,³⁵ we believe that our aryl ethynyl anthraquinones represent a promising platform for the development of a new generation of multifunctional G4 interacting ligands.

Results and discussion

Chemistry

Final products 1–6 were synthesized starting from the commercially available 1,8-dihydroxyanthraquinone 7. The bromination reaction was performed at r.t. using NBS in dichloromethane and $\text{NH}(\text{iPr})_2$. The high reactivity of the hydroxyanthraquinone under these conditions resulted in a poorly selective bromination of 7, affording the anthraquinone 8 (Scheme 2) as the major product of the mixture (30% yield).

The structure of the most abundant stereoisomer 8 has been tentatively assigned according to literature data, which



Scheme 2 (i) NBS, DCM, $\text{NH}(\text{iPr})_2$, r.t., 4 h (yield 30%); (ii) Me_2SO_4 , K_2CO_3 , acetone, reflux, overnight (yield 70%); (iii) TMSA, anhydrous THF, 20 mol% $\text{PdCl}_2(\text{PPh}_3)_2$, 20 mol% CuI ; TEA, reflux 16 h, N_2 (yield 60%); (iv) MeOH-DCM 5 : 1, K_2CO_3 , 1 h, 0 °C, N_2 (yield 98%).

suggest that direct bromination of the unprotected 1,8-dihydroxyanthraquinone such as aloe-emodin and chrysophanol analogues takes place at the desired 2- and 7-positions (see Scheme 2 for numbering), in the presence of a catalytic amount of a secondary amine.^{43–45} The presence of the two OH groups is not compatible with the Sonogashira cross-coupling reaction, therefore these groups were protected as methyl ethers. This step was conducted directly on the bromination crude, which was extremely challenging to purify under standard chromatographic conditions. This crude was suspended in acetone and heated to reflux overnight in the presence of dimethyl sulfate and K_2CO_3 . The chromatographic purification of the resulting dimethyl ether **9** was much more efficient (70%) and straightforward than **8**. Unfortunately, the unambiguous assignment of the corrected bromination regioselectivity could not be achieved by NMR through ^1H , ^{13}C -HMBC (Heteronuclear Multiple Bond Correlation) experiments as the chemical shifts of the two carbonyls were too

close to each other (180.9 vs. 181.8 ppm). Therefore, such a task was carried out on the further synthesized anthraquinone **11**. A first Sonogashira cross-coupling was conducted with compound **9** in the presence of a large excess (10 : 1) of trimethylsilylacetylene (TMSA) in anhydrous THF, TEA and $\text{Pd}(\text{PPh}_3)_2\text{Cl}_2$ and CuI (20 mol% each). These reaction conditions provided **10** in reasonably good yield (60%). The yield of this key step was significantly affected by the sequence of the reactants addition. Adding TMSA immediately after TEA was found to provide the best reaction yield (60%). Deprotection of TMS groups was achieved quantitatively using K_2CO_3 in MeOH-DCM at 0 °C, affording the bis-terminal alkyne **11**, which was used without further purification. This synthetic strategy provided a facile route to the synthesis of the building block **11** in only 4 steps and fairly good yields. The functionalization of the anthraquinone core at 2- and 7-positions was finally and unambiguously assigned using HMBC interactions of the terminal alkyne hydrogens (3.56 ppm) with the quaternary carbons on the methoxy substituents (C-1 and C-8, 161.7 ppm). In addition, both H-4 and H-5 exhibit the HMBC interaction with the most de-shielded carbonyl C-10 (ESI†).

With this building block in hand we sought to investigate convenient synthetic strategies for the preparation of the final products: **1–6**. The aryl iodides (**14–19**, Table 1) required for the Sonogashira reaction were synthesized starting from the *p*- and *m*-iodophenol (ESI, Scheme S1†). For the synthesis of compounds **1** and **2** a protection/deprotection strategy of the phenol was required, as expected, while for the Mannich base the cross-coupling could be carried out using the free phenol derivatives (**18**, **19**). Such an unexpected difference may be related to the formation of an intramolecular H-bond within the Mannich bases **18** and **19**, which could introduce a sort of “self-protective” effect on the phenol OH acidity, recovering the typical reactivity of an “OH-free” aryl iodide.⁴⁶ The yields of the Sonogashira coupling with Mannich bases (**18**, **19**) are

Table 1 Sonogashira cross-couplings, yielding AQs **3–6**, **12** and **13**

AQ	Aryl iodide	X	Y	CuI mol%	$\text{Pd}(\text{PPh}_3)_2\text{Cl}_2$ mol%
12	14	H	OCOCH_3	10	5
13	15	OCOCH_3	H	10	5
3	16	H	$\text{OCH}_2\text{CH}_2\text{NMe}_2$	10	10
4	17	$\text{OCH}_2\text{CH}_2\text{NMe}_2$	H	10	10
5	18	CH_2NMe_2	OH	10	10
6	19	OH	CH_2NMe_2	10	10

still lower compared to the reaction with iodoacetylphenol (14, 15), but the opportunity to avoid the protection and deprotection steps justified somehow the choice of our synthetic strategy.

The Sonogashira coupling conditions between the bis-terminal alkyne 11 and the aryl iodides have been optimized with respect to the solvent, base and catalysts (THF anhydrous, TEA, Pd(PPh₃)₂Cl₂, CuI) for each single aryl iodide substrate (see the Experimental section for the procedure and yields). The final products 3–6 were purified as bis-hydrochloride salts by reverse phase HPLC followed by trifluoroacetate/chloride exchange. Compounds 12 and 13 required an additional deprotection step, which was performed in aqueous methanol in the presence of K₂CO₃ at r.t. HPLC purification afforded the final products 1 and 2 (ESI, Scheme S2†).

The new ligands 1–6 were characterized by absorption spectroscopy and their molar extinction coefficients were calculated in 10 mM Tris, 50 mM KCl at pH 7.5 (Experimental section).

The cationic 3–6 showed a linear correlation between absorption and concentration of up to 50 μM thus confirming good solubility and the lack of extensive aggregation. The only exception was provided by derivatives 1 and 2, which evidenced a relevant deviation starting from the 15 μM ligand concentration followed by precipitation at higher levels. Therefore, all the subsequent analyses for these two derivatives were performed at concentrations lower than 15 μM.

Fluorescence melting assay

To assess the potential of the AQs 1–6 to stabilize peculiar G4 topologies we screened them by fluorescence melting using DNA telomeric sequences properly labelled at the 5'-end with a quencher (dabcyl) and at the 3'-end with a fluorophore (fluorescein).⁴⁷ An increase of the oligonucleotide melting temperature upon addition of the tested compound relies on the ability of the ligands to stabilize the DNA G4 folded structure. Since the human telomeric G4 is characterized by a large conformational flexibility, the analysis was performed under different conditions and with different sequences known to promote distinct folding: HTS (d[AG₃(T₂AG₃)₃T]) which in the presence of K⁺ folds mainly in a population of prevalently hybrid conformations, whereas in Na⁺ it assumes a defined antiparallel folding,⁴⁸ and Tel24 [d(T₂AG₃)₄] which adopts a hybrid-1 folding in K⁺ containing solutions.⁴⁹ The same analysis was additionally performed using a double stranded random sequence (dsDNA) to check for duplex *vs.* quadruplex selectivity.

To summarize our results we report the variation of the oligonucleotide melting temperature as a function of ligand concentration (Fig. 1).

Among the tested ligands only 1 and 2 did not induce any modification of the melting profile of the tested DNA sequences. This sustained the fundamental requirement of protonable groups in the side chain to grant effective nucleic acid recognition. Although the ΔT_m values remain quite low at 1 μM ligand concentration, all the other compounds (AQs 3–6) stabilized the G4 forms. In particular, a sigmoidal correlation

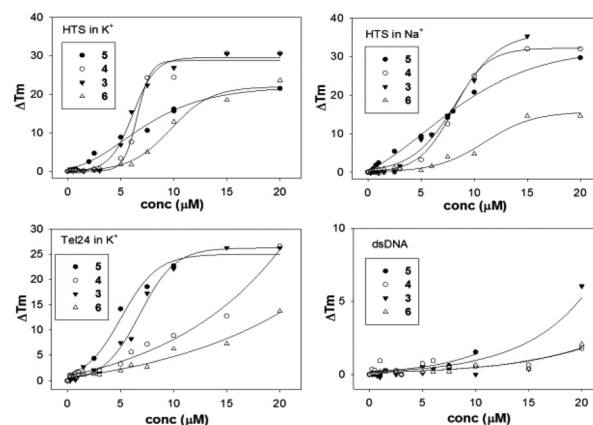


Fig. 1 Variation of the melting temperature of HTS, Tel24 and double stranded DNA (dsDNA) by increasing concentrations of the tested ligands (3–6) in K⁺ or Na⁺ containing buffer.

emerged between the observed ΔT_m and ligand concentrations, which suggested the presence of cooperative binding events. Among the active derivatives, at low micromolar ligand concentration, 5 and 6 turned out to be the most and the least effective, respectively. Conversely, at concentrations higher than 5 μM 3 and 4 behaved as better ligands for Tel22 in K⁺. If we compare the behavior of each anthraquinone derivative for the different tested G4 targets, we did not observe prominent selectivity for any of them. The only exception was 4, which was less active on Tel24. Thus, the presence of oxygen in the *meta* position on the aromatic ring of the side chains seems to negatively perturb the DNA recognition process. Interestingly, the regio-isomers 5 and 6 showed very different binding profiles and this can suggest a peculiar binding interaction for the *para* isomer 5 with the G4 structures. Finally, all tested compounds almost did not affect the thermal stability of the double stranded DNA (Fig. 1), indicating their use as potential G4 selective binders.

Circular dichroism

CD titrations were performed to investigate the ability of the novel compounds to induce structural modifications to the tested G4. Thus, the study was performed using the same oligonucleotide sequences used for thermal stabilization experiments. Moreover, we extended our analysis to wtTel26: d[(T₂AG₃)₄T₂] which, in the presence of K⁺, folds into a hybrid-2 type of arrangement.⁵⁰

The recorded dichroic spectra of the oligonucleotides in the presence of potassium are all characterized by two positive bands, one centered at 290 nm and the other at 265–268 nm which reflect the principal 3 + 1 hybrid arrangement assumed. Conversely, in the presence of sodium, the dichroic spectrum of Tel22 shows a negative band at 260 nm and a positive band at 290 nm, typical of the antiparallel conformation signature identified by NMR spectroscopy.⁵¹

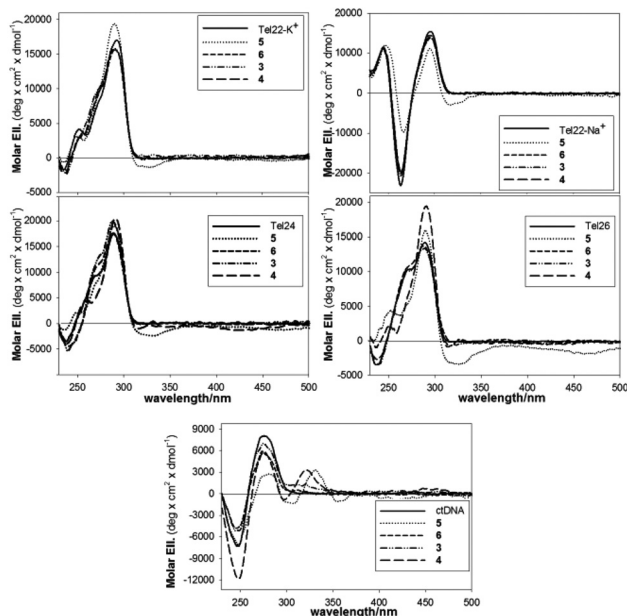


Fig. 2 Circular dichroism spectra of DNA templates (4 μM strand concentrations) alone (solid lines) or upon addition of 4 equivalents of tested AQs 3–6 recorded in 10 mM TRIS, 50 mM KCl or NaCl, pH 7.0, 25 $^{\circ}\text{C}$.

Variations of the intensity of the dichroic features of all tested G4 folded DNA sequences were detected upon addition of the ligands (Fig. 2).

This confirmed the occurrence of a DNA–ligand interaction, which does not affect the G4 topology to a large extent. Interestingly, the most relevant CD variations occurred generally with 5. In this instance, induced dichroic bands (ICD) in the ligand absorption range were also observed (Fig. 2). Such contribution should derive from the insertion of the ligand chromophore into the chiral environment provided by the nucleic acid. Since this effect is a function of the mutual orientation of the AQ chromophore and DNA, we can assume that its presence/lack among the tested derivatives is linked to a significant repositioning of the ligand in the complex as a consequence of the side chain nature and position. Thus, distinct binding modes for 5 *vs.* 3, 4 and 6 can be further inferred.

When the DNA substrate was arranged into a double helix, the most prominent effect was reduction of the 275 and 245 nm DNA dichroic bands. This should exclude the occurrence of an efficient intercalation binding mode for these ligands, since this process usually causes an increment of these optical contributions. This result is in agreement with the above reported lack of thermal stabilization induced by the tested ligands in this nucleic acid conformation.

Enzymatic assays

The above described results were collected to evaluate the recognition of a G4 structure by the tested ligands. Additionally, we analyzed whether they can promote G4 folding generating species sufficiently stable to interfere with enzymes devoted to the processing of the nucleic acid. Thus, the DNA polymerase

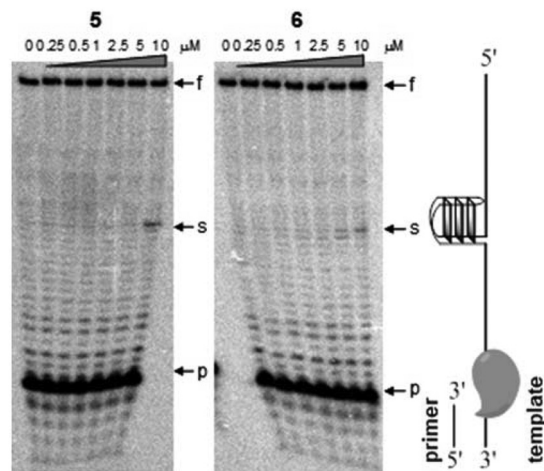


Fig. 3 Taq polymerase stop assay performed with increasing concentrations of 5 and 6 on the HT4-temp template in the presence of 50 mM K^+ . f, s and p refer to full length product, truncated product and primers respectively.

stop assay was performed using a template containing a four-repeat human telomeric sequence (HT4-temp). In a typical experiment, if the compounds under investigation promote G4 formation by the template, the formation of truncated products due to the collision of polymerase with the folded G-rich tract (see the cartoon on the right side of gel reported in Fig. 3) appears.⁵²

The results summarized in Fig. 3 showed that increasing concentrations of the tested ligands lead to a slight decrease of the intensity of the band relative to the fully processed oligonucleotide, which is more evident for 5. At the same time, the tested compounds blocked, by some other means, the primer extension by DNA polymerase starting from 5 to 10 μM , the same concentration range in which they stabilized the G-quadruplex as determined by the melting assay. The observed stop occurs at a well-defined site corresponding to the template G-rich stretch. Interestingly, 3 and 6 tend to arrest the enzyme at position-1 with reference to the G-rich tract at 5 μM concentration. This behavior is not shared by 5, thus further sustaining the different binding mode of this derivative.

Conversely, the same reaction performed on a DNA sequence not G-rich (HT4sc-temp) failed to evidence any interference in the enzymatic activity by tested ligands up to 40 μM . This result correlates with the increased ability of the tested compound to recognize G4 over other nucleic acid arrangements.

Evaluation of the best fitting ligand by docking

The conformational polymorphism of the DNA human telomeric repeat sequence Tel22 prompted us to generate poses of our anthraquinone derivatives using Induced Fit Docking (IFD)⁵³ simulations following our recent experience,⁵⁴ in order to take into account the target flexibility and to optimize the network of DNA–ligand interactions as compared to rigid docking. In particular, as already reported in our recent model-

Table 2 Evaluation of the Induced Fit docking consensus score, calculated for AQs 1–6 towards the six G4 folds

AQ	Consensus score (kcal mol ⁻¹)
1	-6.42
2	-6.79
3	-8.84
4	-9.01
5	-9.29
6	-9.00

ing work,⁵⁵ we included in our study the Protein Data Bank [The Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB); <http://www.rcsb.org/pdb>] entries with the codes 1KF1,⁵⁶ 143D,⁵¹ 2HY9,⁵⁷ 2JPZ,⁵⁸ 2JSL⁵⁹ and 2JSM⁵⁹ taking into account all available X-ray and NMR telomeric structures. The binding energy (IFD score) related to the docking generated ensembles indicated a different theoretical ligand affinity toward the six used G4 folds (Table S1†).

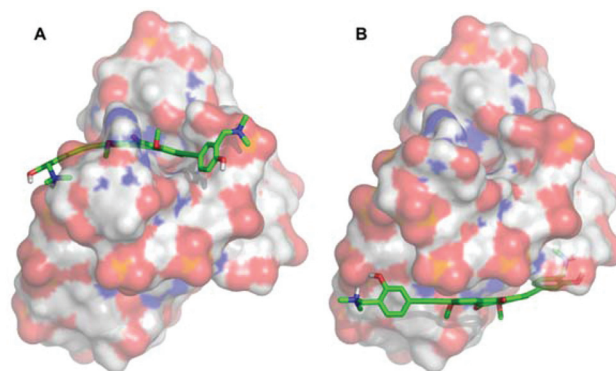
Recent studies indicated the hybrid-type intramolecular G4 structures as the major conformations formed in human telomeric sequences in K⁺ containing solution, with a dynamic equilibrium between hybrid-1 and hybrid-2 folds.^{57,58,60–63} However in K⁺ solution the parallel structure is also found. Remarkably, we obtained better docking results with these G4 folds. These data are in agreement with the human telomeric stabilization reported in Fig. 1 which highlighted a favourable contribution of K⁺ in G4 stabilization by tested ligands. Since this cation is much more abundant than Na⁺ in cellular environments, such a finding highlights the preference toward the physiologically relevant G4 conformations.

Among the analyzed compounds, 5 showed the best average affinity (consensus score, Table 2) with respect to the others, in particular against 1KF1, 2JSM and 2JSL models (Table S1†).

Interestingly, solution studies evidenced a striking difference between the two regio-isomers 5 and 6. This experimental observation is in agreement with our theoretical results, since 5 showed an improved affinity compared to 6 in almost all the considered folds. Such a finding is particularly evident in the recognition of these ligands towards the G4 2JSM hybrid-1 model, as indicated in Fig. 4. Specifically 5 was better embedded in the DNA structure, since it is accommodated in a kind of internal pocket and is involved in a wide stacking interaction network. By contrast 6 is able to recognize only the bottom site of the 2JSM model, probably due to the different position of the phenolic hydroxyl moiety.

Moreover IFD simulations revealed the ability of 5 to establish one pivotal hydrogen bond between its hydroxyl group with the phosphate oxygen of guanine at position 9 and another one between the hydrogen atom of its amino moiety with the phosphate oxygen of guanine at position 2, thus allowing the ligand to better anchor to the G4 structure.

The best poses of the studied anthraquinone derivatives in complex with all the G4 considered folds are reported in the ESI (Fig. S2–S35†).

**Fig. 4** Best pose of (A) 5 and (B) 6 against 2JSM hybrid-1 model of the DNA human telomeric repeat sequence *d*[AG₃(T₂AG₃)₃]. 5 and 6 are indicated as green carbon stick representation, while the DNA is shown as transparent surface. Nonpolar hydrogen atoms are omitted for sake of clarity.

Conclusions

In conclusion, we reported the synthesis of several aryl ethynyl anthraquinones (1–6) *via* the optimized Sonogashira cross-couplings. This synthetic protocol is flexible and can be exploited to introduce further chemical diversity into the ethynyl-AQ scaffold. The ethynyl-AQ derivatives demonstrated to interact and stabilize G4 structures of the telomeric DNA sequence. Their binding properties and quadruplex *vs.* duplex selectivity have been characterized by FRET melting, CD, stop assay as well as IF docking experiments. Compound 5 was the most effective ligand according to all of the assays performed. Varying the relative position of the substituents on the phenolic aromatic ring from *para* to *ortho* (5 and 6, respectively) provided compounds with quite different G4 binding and stabilization properties. Our investigation highlights that structural positioning of Mannich bases is crucial for efficient G4 binding. Although our best candidate (5) is not the most efficient G4 binder among the AQ derivatives tested so far,^{28–31} it lacks significant intercalation into the double helix. This finding highlights the potential to exploit 5 and its analogues as precursors of alkylating QMs, targeting G4s. Although further structural refinement is required to increase the affinity of this scaffold towards telomeric G4-DNA, these preliminary results are encouraging. We are currently working on the development of a larger aryl (QM precursor) ethynyl anthraquinone library as G4 bi-modal ligands, acting on a selective reversible recognition and subsequent alkylation upon reductive activation.

Experimental section

Synthesis of 2,7-dibromo-1,8-dihydroxyanthracene-9,10-dione (8)

10 g of 1,8-dihydroxyanthraquinone 7 (0.042 mol) were dissolved in 300 ml of dichloromethane and 7.35 ml of diisopropylamine NH(iPr)₂ were added under stirring. A solution of

N-bromosuccinimide (18.5 g, 0.104 mol in 800 ml of DCM) was added dropwise over 30 min and the mixture was stirred for 4–6 h at room temperature. After that time the reaction was quenched in slightly acidic water (500 ml, 1% HCl) and the organic layer was separated. The aqueous solution was then washed with DCM (2 × 250 ml); the organic phases were recombined and dried over Na₂SO₄. The solvent was removed under vacuum to afford an orange solid. The crude product can be used directly for the next protection step. To determine the reaction yield and characterize the product, **8** was isolated by column chromatography in cyclohexane–toluene 1:1 affording a yellow-orange solid. Yield 30%. Mp >300 °C. ¹H-NMR (300 MHz, DMSO-d₆): δ 7.60 (d, 2H, *J* = 7.9 Hz), 8.17 (d, 2H, *J* = 7.9 Hz), 12.41 (br s, 2H).

Anal. found: C, 42.2; H, 1.6. Calc. for C₁₄H₆Br₂O₄: C, 42.2; H, 1.5%.

Synthesis of 2,7-dibromo-1,8-dimethoxyanthracene-9,10-dione (**9**)

18.0 g of the crude product **8** were suspended in acetone (900 ml) and K₂CO₃ (19.1 g, 0.128 mol) and Me₂SO₄ (44 ml, 0.461 mol) was added. The mixture was heated to reflux overnight while stirring under argon (18–20 h). During the reaction the mixture became dark while at the end of it a yellow solid crashed out. After 20 h the suspension was cooled down and the solvent was removed under reduced pressure. The crude product was dissolved in DCM (200 ml) and an ammonia aqueous solution (5%) was added. The biphasic mixture was stirred at room temperature for 1 h. After this period the organic phase was separated and washed twice (2 × 200 ml) with an acidic aqueous solution (1% HCl), while the aqueous phase was washed with DCM to recover all the product traces (2 × 200 ml). The organic phases were then recombined, dried over Na₂SO₄ and the solvent removed under vacuum to afford a brown solid. The crude product was purified by flash chromatography (MPLC) with a cyclohexane–ethyl acetate gradient (TLC eluent cyclohexane–ethyl acetate 7:3) affording a yellow solid. Yield 70%. Mp >300 °C. ¹H-NMR (300 MHz, CDCl₃): δ 4.07 (s, 6H), 7.90–7.99 (m, 4H). ¹³C-NMR (CDCl₃): δ 62.3, 123.7, 127.5, 128.8, 133.6, 138.1, 156.8, 180.9, 181.8. Anal. found: C, 45.0; H, 2.4. Calc. for C₁₆H₁₀Br₂O₄: C, 45.1; H, 2.4%.

Synthesis of 1,8-dimethoxy-2,7-bis(trimethylsilyl)ethynylanthracene-9,10-dione (**10**)

3.0 g of **9** (7.30 mmol) were dissolved in anhydrous THF (400 ml) and then in order Pd(PPh₃)₂Cl₂ (20 mol%, 1.02 g, 1.46 mmol) and CuI (20 mol%, 0.278 g, 1.46 mmol) were added under stirring while bubbling the solution with an argon flow. 10.1 ml of TEA (10 equiv., 73 mmol) were added followed immediately by 10.4 ml of TMSA (10 equiv., 73 mmol). The stirred solution was heated to reflux for 16 h under an argon atmosphere. After this period the dark solution was cooled and poured in 200 ml of water. The mixture was then extracted with DCM (3 × 250 ml); the organic phase was collected and dried over Na₂SO₄. The solvent was removed

under reduced pressure affording the crude product as a dark solid. The crude was purified by flash chromatography with a hexane–ethyl acetate gradient (TLC eluent hexane–ethyl acetate, 9:1) to give **10** as a yellow product. Yield 60%. Mp >300 °C. ¹H-NMR (300 MHz, CDCl₃): δ 0.31 (s, 18H), 4.13 (s, 6H), 7.74 (d, 2H, *J* = 8.0 Hz), 7.94 (d, 2H, *J* = 8.0 Hz).

¹³C-NMR (CDCl₃): δ -0.4, 62.0, 99.7, 104.5, 122.0, 125.5, 128.4, 133.8, 137.7, 161.4, 181.6, 182.1. Anal. found: C, 67.8; H, 6.2. Calc. for C₂₆H₂₈O₄Si₂: C, 67.8; H, 6.1%.

Synthesis of 2,7-diethynyl-1,8-dimethoxyanthracene-9,10-dione (**11**)

1.13 g of **10** (2.46 mmol) was dissolved in a MeOH–DCM 5:1 mixture (167:33 ml) and the solution obtained was cooled at 0 °C. 0.60 g of K₂CO₃ was added and the mixture was stirred at 0 °C for 1 h under Ar. The solution was then allowed to reach room temperature and quenched in 100 ml of water. The aqueous solution was extracted with DCM (3 × 150 ml) and the organic phase was collected and dried over Na₂SO₄. The solvent was then removed under vacuum to afford quantitatively **11** as a yellow solid. Yield ≥98%. Mp >300 °C. ¹H-NMR (300 MHz, CDCl₃): δ 3.56 (s, 2H), 4.13 (s, 6H), 7.79 (d, 2H, *J* = 8.0 Hz), 7.97 (d, 2H, *J* = 8.0 Hz). ¹³C-NMR (CDCl₃): δ 62.5, 78.6, 85.8, 122.2, 124.7, 128.3, 133.2, 138.2, 161.7, 181.4, 182.0. Anal. found: C, 75.8; H, 3.8. Calc. for C₂₀H₁₂O₄: C, 75.9; H, 3.8%.

General procedure for the synthesis of 12–13

The corresponding aryl iodide **14** or **15** (2.0 mmol, 0.53 g) was dissolved in anhydrous THF (60 ml) and Pd(PPh₃)₂Cl₂ (5 mol%, 70.4 mg, 0.10 mmol) and CuI (10 mol%, 38.2 mg, 0.2 mmol) were added under stirring while bubbling the solution with an argon flow. TEA (0.56 ml, 2 equiv., 4.0 mmol) was then added immediately followed by a solution of **11** (0.4 equiv., 254 mg, 0.80 mmol) in a degassed anhydrous THF solution (15 ml).

The mixture was heated to reflux for 5 h under Ar, and then cooled down and poured into water (100 ml). The aqueous solution was extracted with DCM (3 × 100 ml) and the organic phases were collected and dried over Na₂SO₄. The solvent was removed under reduced pressure to afford an orange solid. The crude products were purified by flash chromatography (eluent: cyclohexane–acetate) to give **12** (17%) and **13** (20%).

2,7-Bis(*p*-acetoxylphenyl)ethynyl-1,8-dimethoxyanthracene-9,10-dione (**12**)

Orange powder. Yield 17%. Mp >300 °C. ¹H-NMR (300 MHz, CDCl₃): δ 2.34 (s, 6H), 4.20 (s, 6H), 7.16 (d, 4H, *J* = 8.6 Hz), 7.62 (d, 4H, *J* = 8.6 Hz), 7.83 (d, 2H, *J* = 8.0 Hz), 8.02 (d, 2H, *J* = 8.0 Hz). ¹³C-NMR (CDCl₃): δ 21.0, 62.3, 84.8, 97.3, 120.1, 121.8, 122.3, 125.7, 128.4, 132.9, 133.7, 137.3, 151.1, 160.9, 168.9, 181.7, 182.0. Anal. found: C, 73.9; H, 4.1. Calc. for C₃₆H₂₄O₈: C, 74.0; H, 4.1%.

2,7-Bis(*m*-acetoxyphenyl)ethynyl)-1,8-dimethoxyanthracene-9,10-dione (13)

Yellow powder. Yield 20%. Mp >300 °C. $^1\text{H-NMR}$ (300 MHz, CDCl_3): δ 2.35 (s, 6H), 4.20 (s, 6H), 7.15–7.18 (m, 2H), 7.34–7.35 (m, 2H), 7.40–7.50 (m, 4H), 7.83 (d, 2H, $J = 8.0$ Hz), 8.03 (d, 2H, $J = 8.0$ Hz). $^{13}\text{C-NMR}$ (CDCl_3): δ 20.9, 62.4, 85.3, 97.0, 122.3, 122.6, 123.7, 124.7, 125.6, 128.4, 129.2, 129.5, 133.9, 137.4, 150.5, 169.1, 171.3, 181.7, 182.0. Anal. found: C, 74.1; H, 4.1. Calc. for $\text{C}_{36}\text{H}_{24}\text{O}_8$: C, 74.0; H, 4.1%.

General procedure for the synthesis of 1–2

Compounds **12** or **13** (0.59 g, 1 mmol) were dissolved in a mixture of MeOH–H₂O 4:1 (80:20 ml). K_2CO_3 was added (1.68 g, 12.1 mmol) and the mixture was stirred at room temperature under Ar for 5 h (**12**) or 20 h (**13**). After the indicated time, the solution was poured in 50 ml of water and methanol was eliminated by evaporation. The aqueous solution was then acidified with HCl 10% and extracted with CHCl_3 (3 × 250 ml). The organic phase was collected and dried over Na_2SO_4 . The solvent was removed under reduced pressure to afford the crude products.

Both crude products were purified by reverse phase HPLC (gradient H₂O + 0.1% TFA, CH_3CN) to afford the final products **1** (15%) and **2** (10%).

2,7-Bis(*p*-hydroxyphenyl)ethynyl)-1,8-dimethoxyanthracene-9,10-dione (1)

Orange needles. Yield 15%. Mp >300 °C. $A_{319}(\text{H}_2\text{O}) = 37\,300 \text{ M}^{-1} \text{ cm}^{-1}$. $^1\text{H-NMR}$ (300 MHz, DMSO-d_6): δ 4.07 (s, 6H), 6.86 (d, 4H, $J = 8.7$ Hz), 7.47 (d, 4H, $J = 8.7$ Hz), 7.90 (br s, 4H), 10.12 (br s, 2H). $^{13}\text{C-NMR}$ (DMSO-d_6): δ 61.8, 83.2, 98.8, 111.7, 116.0, 122.0, 125.3, 128.3, 133.2, 133.4, 137.0, 158.9, 159.7, 181.3, 181.5. Anal. found: C, 76.8; H, 3.9. Calc. for $\text{C}_{32}\text{H}_{20}\text{O}_6$: C, 76.8; H, 4.0%.

2,7-Bis(*m*-hydroxyphenyl)ethynyl)-1,8-dimethoxyanthracene-9,10-dione (2)

Yellow needles. Yield 10%. Mp >300 °C. $A_{315}(\text{H}_2\text{O}) = 35\,700 \text{ M}^{-1} \text{ cm}^{-1}$. $^1\text{H-NMR}$ (300 MHz, DMSO-d_6): δ 4.18 (s, 6H), 7.00 (d, 2H, $J = 8.0$ Hz), 7.10 (br s, 2H), 7.17 (d, 2H, $J = 8.0$ Hz), 7.39 (t, 2H, $J = 8.0$ Hz), 8.03 (d, 2H, $J = 8.0$ Hz), 8.08 (d, 2H, $J = 8.0$ Hz), 9.90 (s, 2H). $^{13}\text{C-NMR}$ (DMSO-d_6): δ 62.0, 84.2, 97.7, 111.3, 117.1, 117.8, 122.1, 122.4, 124.6, 128.3, 130.1, 133.7, 137.4, 157.5, 160.1, 181.3, 181.4. Anal. found: C, 76.7; H, 4.0. Calc. for $\text{C}_{32}\text{H}_{20}\text{O}_6$: C, 76.8; H, 4.0%.

General procedure for the synthesis of 3–6

0.28 mmol of the corresponding aryl iodide (**14**, **15**, **16**, **17**) was dissolved in 10 ml of anhydrous THF. $\text{Pd}(\text{PPh}_3)_2\text{Cl}_2$ (10 mol%, 20.0 mg, 0.28 mmol) and CuI (10 mol%, 5.4 mg, 0.28 mmol) were added under stirring while purging the solution with an argon flow. The mixture was heated at 50 °C and TEA (2 equiv., 78 μl , 0.57 mmol) was added.

The bis-ethynyl derivative **11** (0.33 equiv., 30 mg, 0.094 mmol) was dissolved in 10 ml of THF. This solution was

purged with Ar and added dropwise over 20 minutes in the reaction vessel containing the iodide and the catalysts at 50 °C. After all of the alkyne was added to the solution the reaction was stopped. The mixture was cooled to r.t. and then poured in 20 ml of water. The aqueous phase was extracted with DCM containing 20% of methanol (3 × 30 ml) to increase the solubility of the products in the organic phase.

The organic phase was dried over Na_2SO_4 and the solvent was removed under reduce pressure to afford the crude products.

Ethynyl-AQs (**3**, **4**, **5** and **6**) were purified by reverse phase HPLC (gradient H₂O + 0.1% TFA, CH_3CN). TFA salts were exchanged with HCl to afford bis-hydrochloride as final salt products. Due to the low solubility of the products the crude solid was subjected to a particular preparation method before the injection in preparative HPLC. In more detail, the crude was suspended in MeOH–H₂O (slightly acidic) 3:1, sonicated and heated at 60 °C for 10 min. The suspension was then filtered and injected directly in HPLC, while the solid was subjected to another treatment before being wasted.

2,7-Bis(4-(2-(dimethylamino)ethoxyphenyl)ethynyl)-1,8-dimethoxyanthracene-9,10-dione-2HCl (3)

Yellow oil. Yield 7%. $A_{310}(\text{H}_2\text{O}) = 29\,300 \text{ M}^{-1} \text{ cm}^{-1}$. $^1\text{H-NMR}$ (300 MHz, CD_3OD): δ 3.03 (s, 12H), 3.66 (t, 4H, $J = 4.6$), 4.18 (s, 6H), 4.44 (t, 4H, $J = 4.6$), 7.13 (d, 4H, $J = 8.9$ Hz), 7.62 (d, 4H, $J = 8.9$ Hz), 7.89 (d, 2H, $J = 8.0$ Hz), 8.02 (d, 2H, $J = 8.0$ Hz). $^{13}\text{C-NMR}$ (CD_3OD): δ 44.2, 58.0, 63.1, 63.5, 85.1, 99.3, 116.5, 117.5, 123.8, 127.6, 130.1, 134.9, 135.4, 138.8, 160.2, 162.1, 183.4, 183.8. Anal. found: C, 67.2; H, 5.7; N, 3.9. Calc. for $\text{C}_{40}\text{H}_{40}\text{Cl}_2\text{N}_2\text{O}_6$: C, 67.1; H, 5.6; Cl, 9.9; N, 3.9%.

2,7-Bis(3-(2-(dimethylamino)ethoxyphenyl)ethynyl)-1,8-dimethoxyanthracene-9,10-dione-2HCl (4)

Yellow oil. Yield 20%. $A_{300}(\text{H}_2\text{O}) = 34\,500 \text{ M}^{-1} \text{ cm}^{-1}$. $^1\text{H-NMR}$ (300 MHz, CD_3OD): δ 3.02 (s, 12H), 3.64 (t, 4H, $J = 4.6$ Hz), 4.13 (s, 6H), 4.41 (t, 4H, $J = 4.6$ Hz), 7.08 (d, 2H, $J = 8.2$ Hz), 7.21 (br s, 4H), 7.32–7.38 (m, 2H), 7.82 (d, 2H, $J = 8.0$ Hz), 7.92 (d, 2H, $J = 8.0$ Hz). $^{13}\text{C-NMR}$ (CD_3OD): δ 44.3, 58.0, 63.2, 63.6, 86.0, 98.8, 117.6, 118.9, 123.8, 125.3, 126.8, 127.0, 129.9, 131.4, 135.6, 139.1, 159.3, 162.3, 183.2, 183.5. Anal. Found: C, 67.1; H, 5.7; N, 4.0. Calc. for $\text{C}_{40}\text{H}_{40}\text{Cl}_2\text{N}_2\text{O}_6$: C, 67.1; H, 5.6; N, 3.9%.

2,7-Bis(3-((dimethylamino)methyl)-4-iodoxyphenyl)ethynyl)-1,8-dimethoxyanthracene-9,10-dione-2HCl (5)

Orange powder. Yield 15%. Mp >300 °C. $A_{326}(\text{H}_2\text{O}) = 28\,200 \text{ M}^{-1} \text{ cm}^{-1}$. $^1\text{H-NMR}$ (300 MHz, CD_3OD): δ 2.91 (s, 12H), 4.12 (s, 6H), 4.36 (s, 4H), 7.00 (d, 2H, $J = 8.4$ Hz), 7.54 (d, 2H, $J = 8.4$ Hz), 7.62 (s, 2H), 7.74 (d, 2H, $J = 8.0$ Hz), 7.88 (d, 2H, $J = 8.0$ Hz). $^{13}\text{C-NMR}$ (CD_3OD): δ 43.7, 58.0, 63.2, 85.2, 99.0, 115.6, 117.3, 118.9, 123.8, 127.4, 129.8, 135.2, 136.8, 137.4, 138.7, 159.3, 162.0, 183.2, 183.6. Anal. found: C, 66.5; H, 5.3; N, 4.0. Calc. for $\text{C}_{38}\text{H}_{36}\text{Cl}_2\text{N}_2\text{O}_6$: C, 66.4; H, 5.3; N, 4.1%.

2,7-Bis(4-((dimethylamino)methyl)-3-iodoxyphenyl)ethynyl)-1,8-dimethoxyanthracene-9,10-dione·2HCl (6)

Yellow powder. Yield 10%. Mp >300 °C. $A_{302} = 29\,300\text{ M}^{-1}\text{ cm}^{-1}$. $^1\text{H-NMR}$ (300 MHz, CD_3OD): δ 2.91 (s, 12H), 4.17 (s, 6H), 4.37 (s, 4H), 7.10–7.25 (m, 4H), 7.45 (d, 2H, $J = 8.2\text{ Hz}$), 7.91 (d, 2H, $J = 8.0\text{ Hz}$), 8.01 (d, 2H, $J = 8.0\text{ Hz}$). $^{13}\text{C-NMR}$ (CD_3OD): δ 43.7, 58.2, 63.3, 85.8, 98.2, 119.3, 119.5, 123.8, 124.8, 126.9, 127.4, 130.03, 134.2, 135.9, 139.1, 158.2, 162.5, 183.3, 183.6. Anal. found: C, 66.5; H, 5.3; N, 4.1. Calc. for $\text{C}_{38}\text{H}_{36}\text{Cl}_2\text{N}_2\text{O}_6$: C, 66.4; H, 5.3; N, 4.1%.

General procedures and synthesis of the arylido-derivatives 14–19 have been reported in the ESL.†

Circular dichroism measurements

Circular dichroism spectra were recorded on a Jasco J-810 spectropolarimeter equipped with a Peltier temperature controller in 10 mM Tris, 50 mM KCl at pH 7.5 using a 10 mm path-length cell. DNA substrates were the four-repeat human telomeric sequences Tel22: $\text{d}[\text{AG}_3(\text{T}_2\text{AG}_3)_3]$, Tel24: $\text{d}(\text{T}_2\text{AG}_3)_4$ and wtTel26: $\text{d}[(\text{T}_2\text{AG}_3)_4\text{T}_2]$ provided by Eurogentec. As double stranded DNA we used calf thymus DNA (ctDNA, Sigma). Before data acquisition, G4 forming solutions (4 μM strand concentration) were heated at 95 °C for 5 min and left to cool at room temperature overnight. The spectra of the nucleic acid alone and in the presence of increasing ligand concentrations (0–20 μM) were acquired. Each reported spectrum represents the average of 3 scans recorded with 1 nm step resolution. The observed CD signals were converted to the mean residue ellipticity $[\theta] = \text{deg cm}^{-2} \text{ dmol}^{-1}$ (Mol. Ellip.).

Fluorescence melting studies

Melting experiments were performed using a Roche LightCycler, using an excitation source at 488 nm and recording the fluorescence emission at 520 nm. Target DNA (Eurogentec) were the human telomeric sequence HTS $\text{d}[\text{AG}_3(\text{T}_2\text{AG}_3)_3\text{T}]$, Tel24 $\text{d}[(\text{T}_2\text{AG}_3)_4]$, and a 18 bp double stranded DNA (5'-GTGAGATACCGACAGAAG) labeled with Dabcyl at the 5' end and FAM at the 3' end. Samples contained 0.25 μM of target DNA and increasing concentrations of tested derivatives in 50 mM potassium buffer (10 mM LiOH; 50 mM KCl pH 7.5 with H_3PO_4). They were first heated to 95 °C at a rate of 0.1 °C s^{-1} , maintained at 95 °C for 5 min and then annealed by cooling to 30 °C at a rate of 0.1 °C s^{-1} . Then, samples were maintained at 30 °C for 5 min before being slowly heated to 95 °C (1 °C min^{-1}) and annealed at a rate of 1 °C min^{-1} . Recordings were acquired during both these melting and annealing steps to check for hysteresis. T_m values were determined from the first derivatives of the melting profiles using the Roche LightCycler software. Each curve was repeated at least three times and errors were ± 0.4 °C. ΔT_m values were calculated by subtracting the T_m value recorded in the presence of the ligand from the corresponding value in the absence of the ligand.

Polymerase stop assay

The DNA primer $\text{d}[\text{TAATACGACTCACTATAG}]$, the human telomeric template sequence HT4-temp $\text{d}[\text{TC}_2\text{A}_2\text{CTATGTATAC}(\text{T}_2\text{AG}_3)_4\text{ACATATCGATGA}_3\text{T}_2\text{GCTATAGTGAGTCGTATTA}]$ and the control template sequence HT4sc-temp $\text{d}[\text{TC}_2\text{A}_2\text{CTATGTATACT}_2\text{G}_2\text{ATGTGAGTGTG AGTGTGAG}_2\text{ACATATCGATGA}_3\text{T}_2\text{GCTATAGTGAGTCGTATTA}]$ were obtained from Eurogentec. The primer was initially 5'-labeled with ^{32}P and T4 polynucleotide kinase (Thermo Scientific), by incubating the reaction mixture at 37 °C for 30 min. The kinase activity was inactivated by heating the reaction mixture at 85 °C for 5 min, followed by two extractions with one volume of phenol- CHCl_3 (50 : 50). An equimolar mixture of the labeled primer and template (20 nM) had been annealed in the polymerase required buffer and subsequently, increasing ligand concentrations have been added. After incubation (30 min at r.t.) 2.5 U of Taq polymerase (Thermo Scientific) and 100 μM dNTP mixture were added to each sample and the resulting solutions were kept for 30 min at 55 °C. Reaction products were resolved by gel electrophoresis (12% polyacrylamide gel with 7 M urea) in $1\times$ TBE (89 mM Tris base, 89 mM boric acid, 2 mM Na_2EDTA). Gels were dried and resolved bands were visualized on a PhosphorImager (Amersham).

Docking experiments

In order to take into account the conformational polymorphism of the DNA human telomeric repeat sequence $\text{d}[\text{AG}_3(\text{T}_2\text{AG}_3)_3]$, we included in our study six PDB entries (codes 1KF1,⁵⁶ 143D,⁵¹ 2HY9,⁵⁷ 2JPZ,⁵⁸ 2JSL⁵⁹ and 2JSM⁵⁹) among X-ray and NMR structures, using all the conformations stored in each experimental structure. Initially, both ligands and DNA were pretreated. For ligand preparation, the 3D structures of all the studied compounds were generated with the Maestro Build Panel [Maestro, version 9.3; Schrödinger, LLC: New York, NY, 2012] and submitted to 20 000 iterations of energy minimization using the Polak–Ribiere Conjugated Gradient (PRCG) algorithm, OPLS2005⁶⁴ as a force field with the all atoms notation, and the implicit model of solvation GB/SA water.⁶⁵ Co-crystallized water molecules and counter ions were removed from the DNA X-ray structure. In their sequences, all the hybrid models presented head and tail caps, each formed by a different number of additional nucleotides. In particular, the hybrid NMR structures 2HY9 and 2JPZ resulted both formed by 26-mer, while in the hybrid models 2JSL and 2JSM they were reported sequences with, respectively, 25- and 23-mer. Thus, to obtain a similar analysis with respect to the first two models, the hybrid PDB structures were modified by deleting these caps, that is, considering them as conformational templates for the canonical 22-mer $\text{d}[\text{AG}_3(\text{T}_2\text{AG}_3)_3]$. The 47 experimental conformations stored in the six PDB models were energy-optimized exactly under the same conditions (force field, implicit solvation model, iterations and convergence criterion) adopted for the ligands. The energy minimization was performed until the rmsd of all heavy atoms was within 0.05 Å of the original PDB model. The

evaluation of the most stable conformations of the DNA structures, for each model, has been performed after the pre-treatment. Docking studies were carried out using IFD.⁵¹ An initial Glide SP docking of each ligand was carried out using a softened potential, a van der Waals radius scaling factor of 0.50 for receptor/ligand atoms, and a number of 20 poses per ligand to be energy minimized with the OPLS-AA force field.⁶⁶ The poses were saved for each ligand and submitted to the subsequent Prime side chain orientation prediction of residues with a distance cutoff of 5 Å around each ligand. After the prime minimization of the nucleobases and the ligand for each pose, a Glide SP redocking of each DNA–ligand complex structure within 30 kcal mol⁻¹ above the global minimum was performed. Finally, each output pose was estimated by the binding energy (IFD score) and visually examined.

All the 3D figures were obtained with PyMOL graphics and the modeling package, version 0.98 [Delano W. L. The PyMOL Molecular Graphics System, 2002. <http://www.pymol.org>].

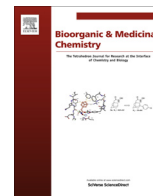
Acknowledgements

This research was supported by Grants from the Italian Ministry of Education, University and Research (MIUR): FIRB-IDEAS (code FIRB RBID082ATK_003), PRIN 2009 (code 2009MFRKZ8), and by AIRC, the Italian Association for Cancer Research (IG 2013 N.14708).

Notes and references

- J. Choi and T. Majima, *Chem. Soc. Rev.*, 2011, **40**, 5893–5909.
- J. Zhao, A. Bacolla, G. Wang and K. M. Vasquez, *Cell. Mol. Life Sci.*, 2010, **67**, 43–62.
- S. Balasubramanian, L. H. Hurley and S. Neidle, *Nat. Rev. Drug Discovery*, 2011, **10**, 261–275.
- M. L. Bochman, K. Paeschke and V. A. Zakian, *Nat. Rev. Genet.*, 2012, **13**, 770–780.
- J. T. Davis, *Angew. Chem., Int. Ed.*, 2004, **43**, 668–698.
- R. Rodriguez, S. Muller, J. A. Yeoman, C. Trentesaux, J. F. Riou and S. Balasubramanian, *J. Am. Chem. Soc.*, 2008, **130**, 15758–15759.
- R. Rodriguez, K. M. Miller, J. V. Forment, C. R. Bradshaw, M. Nikan, S. Britton, T. Oelschlaegel, B. Xhemalce, S. Balasubramanian and S. Jackson, *Nat. Chem. Biol.*, 2012, **8**, 301–310.
- C. Douarre, X. Mergui, A. Sidibe, D. Gomez, P. Alberti, P. Mailliet, C. Trentesaux and J. F. Riou, *Nucleic Acids Res.*, 2013, **41**, 3588–3599.
- T. A. Brooks and L. H. Hurley, *Genes Cancer*, 2010, **1**, 641–649.
- S. Kumari, A. Bugaut, J. L. Huppert and S. Balasubramanian, *Nat. Chem. Biol.*, 2007, **3**, 218–221.
- A. Bugaut and S. Balasubramanian, *Nucleic Acids Res.*, 2012, **40**, 4727–4741.
- K. Paeschke, M. L. Bochman, P. D. Garcia, P. Cejka, K. L. Friedman, S. C. Kowalczykowski and V. A. Zakian, *Nature*, 2013, **497**, 458–462; A. Piazza, A. Serero, J. B. Boulé, P. Legoix-Né, J. Lopes and A. Nicolas, *PLoS Genet.*, 2012, **8**, 1–17.
- J. L. Huppert and S. Balasubramanian, *Nucleic Acids Res.*, 2007, **35**, 406–413.
- J. L. Huppert, A. Bugaut, S. Kumari and S. Balasubramanian, *Nucleic Acids Res.*, 2008, **36**, 6260–6268.
- K. M. Miller and R. Rodriguez, *Expert Rev. Clin. Pharmacol.*, 2011, **4**, 139–142.
- G. Biffi, D. Tannahill, J. McCafferty and S. Balasubramanian, *Nat. Chem.*, 2013, **5**, 182–186.
- D. Monchaud and M. P. Teulade-Fichou, *Org. Biomol. Chem.*, 2008, **6**, 627–636.
- S. Neidle, *Curr. Opin. Struct. Biol.*, 2009, **19**, 239–250.
- M. Di Antonio, G. Biffi, A. Mariani, E. A. Raiber, R. Rodriguez and S. Balasubramanian, *Angew. Chem., Int. Ed.*, 2012, **51**, 11073–11078.
- D. Sun, B. Thompson, B. E. Cathers, M. Salazar, S. M. Kerwin, J. O. Trent, T. C. Jenkins, S. Neidle and L. H. Hurley, *J. Med. Chem.*, 1997, **40**, 2113–2116.
- D. Cairns, E. Michalitsi, T. C. Jenkins and S. P. Macka, *Bioorg. Med. Chem.*, 2002, **10**, 803–807.
- T. Ndlbe and G. B. Schuster, *Org. Biomol. Chem.*, 2006, **4**, 4015–4021.
- C. Monneret, *Eur. J. Med. Chem.*, 2001, **36**, 483–493.
- N. Li, Y. Ma, C. Yang, L. Guo and X. Yang, *Biophys. Chem.*, 2005, **116**, 199–205.
- Z. Hajihassan and A. Rabbani-Chadegani, *J. Biomed. Sci.*, 2009, **11**, 16–31.
- H. Suzuki, T. Ikeda, T. Yamagishi, S. Nakaike, S. Nakane and M. Ohsawa, *Mutat. Res.*, 1995, **328**, 151–161.
- G. R. Clark, P. D. Pytel, C. J. Squire and S. Neidle, *J. Am. Chem. Soc.*, 2003, **125**, 4066–4067.
- P. J. Perry, S. M. Gowan, A. P. Reszka, P. Polucci, T. C. Jenkins, L. R. Kelland and S. Neidle, *J. Med. Chem.*, 1998, **41**, 3253–3260.
- P. J. Perry, A. P. Reszka, A. A. Wood, M. A. Read, S. M. Gowan, H. S. Dosanjh, J. O. Trent, T. C. Jenkins, L. R. Kelland and S. Neidle, *J. Med. Chem.*, 1998, **41**, 4873–4884.
- P. J. Perry, M. A. Read, R. T. Davies, S. M. Gowan, A. P. Reszka, A. A. Wood, L. R. Kelland and S. Neidle, *J. Med. Chem.*, 1999, **42**, 2679–2684.
- G. Zagotto, C. Sissi, L. Lucatello, C. Pivetta, S. A. Cadamuro, K. R. Fox, S. Neidle and M. Palumbo, *J. Med. Chem.*, 2008, **51**, 5566–5574.
- N. Ranjan, E. Davis, L. Xue and D. P. Arya, *Chem. Commun.*, 2013, **49**, 5796–5798.
- G. Zagotto, C. Sissi, S. Moro, D. Dal Ben, G. N. Parkinson, K. R. Fox, S. Neidle and M. Palumbo, *Bioorg. Med. Chem.*, 2008, **16**, 354–361.
- G. Zagotto, A. Ricci, E. Vasquez, A. Sandoli, S. Benedetti, M. Palumbo and C. Sissi, *Bioconjugate Chem.*, 2011, **22**, 2126–2135.

- 35 E. H. van Dijk, D. J. T. Myles, M. H. van der Veen and J. C. Hummelen, *Org. Lett.*, 2006, **8**, 2333–2336.
- 36 E. Modica, R. Zanaletti, M. Freccero and M. Mella, *J. Org. Chem.*, 2001, **66**, 41–52.
- 37 M. Di Antonio, F. Doria, M. Mella, D. Merli, A. Profumo and M. Freccero, *J. Org. Chem.*, 2007, **72**, 8354–8360.
- 38 C. Percivalle, F. Doria and M. Freccero, *Curr. Org. Chem.*, 2014, **18**, 19–43.
- 39 M. Di Antonio, F. Doria, S. N. Richter, C. Bertipaglia, M. Mella, C. Sissi, M. Palumbo and M. Freccero, *J. Am. Chem. Soc.*, 2009, **131**, 13132–13141.
- 40 M. Nadai, F. Doria, M. Di Antonio, G. Sattin, L. Germani, C. Percivalle, M. Palumbo, S. N. Richter and M. Freccero, *Biochimie*, 2011, **93**, 1328–1340.
- 41 F. Doria, M. Nadai, M. Folini, M. Di Antonio, L. Germani, C. Percivalle, C. Sissi, N. Zaffaroni, S. Alcaro, A. Artese, S. N. Richter and M. Freccero, *Org. Biomol. Chem.*, 2012, **10**, 2798–2806.
- 42 F. Doria, M. Nadai, M. Folini, M. Scalabrin, L. Germani, G. Sattin, M. Mella, M. Palumbo, N. Zaffaroni, D. Fabris, M. Freccero and S. N. Richter, *Chem. – Eur. J.*, 2013, **19**, 78–81.
- 43 G. Bringmann and D. Menche, *Angew. Chem., Int. Ed.*, 2001, **40**, 1687–1690.
- 44 G. Bringmann, D. Menche, J. Kraus, J. Mühlbacher, K. Peters, E.-M. Peters, R. Brun, M. Bezabih and B. M. Abegaz, *J. Org. Chem.*, 2002, **67**, 5595–5610.
- 45 L. F. Tietze, K. M. Gericke and I. Schuberth, *Eur. J. Org. Chem.*, 2007, 4563–4577.
- 46 E. Negishi and L. Anastasia, *Chem. Rev.*, 2003, **103**, 1979–2017.
- 47 R. A. J. Darby, M. Sollogoub, C. McKeen, L. Brown, A. Risitano, N. Brown, C. Barton, T. Brown and K. R. Fox, *Nucleic Acids Res.*, 2002, **30**, e39.
- 48 J. M. Nicoludis, S. P. Barrett, J. L. Mergny and L. A. Yatsunyk, *Nucleic Acids Res.*, 2012, **40**, 5432–5447.
- 49 K. N. Luu, A. T. Phan, V. V. Kuryavyi, L. Lacroix and D. J. Patel, *J. Am. Chem. Soc.*, 2006, **128**, 9963–9970.
- 50 A. Ambrus, D. Chen, J. Dai, T. Bialis, R. A. Jones and D. Yang, *Nucleic Acids Res.*, 2006, **34**, 2723–2735.
- 51 Y. Wang and D. J. Patel, *Structure*, 1993, **1**, 263–282.
- 52 H. Han, L. H. Hurley and M. Salazar, *Nucleic Acids Res.*, 1999, **27**, 537–542.
- 53 W. Sherman, T. Day, M. P. Jacobson, R. A. Friesner and R. Farid, *J. Med. Chem.*, 2006, **49**, 534–553.
- 54 C. Varela, E. J. Tavares da Silva, C. Amaral, G. Correia da Silva, T. Baptista, S. Alcaro, G. Costa, R. A. Carvalho, N. A. Teixeira and F. M. Roleira, *J. Med. Chem.*, 2012, **55**, 3992–4002.
- 55 S. Alcaro, C. Musetti, S. Distinto, M. Casatti, G. Zagotto, A. Artese, L. Parrotta, F. Moraca, G. Costa, F. Ortuso, E. Maccioni and C. Sissi, *J. Med. Chem.*, 2013, **56**, 843–855.
- 56 G. N. Parkinson, M. P. Lee and S. Neidle, *Nature*, 2002, **417**, 876–880.
- 57 J. Dai, C. Punchihewa, A. Ambrus, D. Chen, R. A. Jones and D. Yang, *Nucleic Acids Res.*, 2007, **35**, 2440–2450.
- 58 J. Dai, M. Carver, C. Punchihewa, R. A. Jones and D. Yang, *Nucleic Acids Res.*, 2007, **35**, 4927–4940.
- 59 A. T. Phan, V. Kuryavyi, K. N. Luu and D. J. Patel, *Nucleic Acids Res.*, 2007, **35**, 6517–6525.
- 60 P. Alberti, P. Schmitt, C. H. Nguyen, C. Rivalle, M. Hoarau, D. S. Grierson and J. L. Mergny, *Bioorg. Med. Chem. Lett.*, 2002, **12**, 1071–1074.
- 61 J. L. Mergny, L. Lacroix, M. P. Teulade-Fichou, C. Hounsou, L. Guittat, M. Hoarau, P. B. Arimondo, J. P. Vigneron, J. M. Lehn, J. F. Riou, T. Garestier and C. Helene, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 3062–3067.
- 62 L. Guittat, A. De Cian, F. Rosu, V. Gabelica, E. De Pauw, E. Delfourne and J. L. Mergny, *Biochim. Biophys. Acta*, 2005, **1724**, 375–384.
- 63 A. Ambrus, D. Chen, J. Dai, T. Bialis, R. A. Jones and D. Yang, *Nucleic Acids Res.*, 2006, **34**, 2723–2735.
- 64 J. L. Banks, H. S. Beard, Y. Cao, A. E. Cho, W. Damm, R. Farid, A. K. Felts, T. A. Halgren, D. T. Mainz, J. R. Maple, R. Murphy, D. M. Philipp, M. P. Repasky, L. Y. Zhang, B. J. Berne, R. A. Friesner, E. Gallicchio and R. M. Levy, *J. Comput. Chem.*, 2005, **26**, 1752–1780.
- 65 W. C. Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J. Am. Chem. Soc.*, 1990, **112**, 6127–6129.
- 66 W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, 11225–11236.



Further SAR studies on bicyclic basic merbarone analogues as potent antiproliferative agents



Andrea Spallarossa^{a,*}, Chiara Rotolo^a, Claudia Sissi^b, Giuseppe Marson^b, Maria Laura Greco^b, Angelo Ranise^a, Paolo La Colla^c, Bernardetta Busonera^c, Roberta Loddo^c

^aDipartimento di Farmacia, Sezione di Chimica del Farmaco e del Prodotto Cosmetico, Università degli Studi di Genova, Viale Benedetto XV 3, I-16132 Genova, Italy

^bDipartimento di Scienze Farmaceutiche e Farmacologiche, Università di Padova, Viale Marzolo 5, I-35131 Padova, Italy

^cDipartimento di Scienze e Tecnologie Biomediche, Università di Cagliari, Cittadella Universitaria, S.S. 554, Km 4,500, I-09042 Monserrato, Cagliari, Italy

ARTICLE INFO

Article history:

Received 17 April 2013

Revised 26 August 2013

Accepted 27 August 2013

Available online 6 September 2013

Keywords:

Pyrimidopyrimidine derivatives

Topoisomerase II inhibitors

Antiproliferative agents

ABSTRACT

Pyrimidopyrimidine derivatives **1** were prepared as rigid thioanalogues of merbarone (a catalytic topoisomerase II inhibitor) and screened as antiproliferative agents against different tumor cell lines. A number of the synthesized compounds emerged as cytotoxic in cell-based assays (MT-4, HeLa and MCF-7 cells) at low micromolar concentrations. In a National Cancer Institute screening, selected member of the series showed a broad spectrum of antiproliferative activity against various tumours (melanoma, renal, CNS, colon and breast cancers). The acid–base and steric properties of the substituent at position 7 of the pyrimidopyrimidine scaffold deeply affected potency. Enzymatic assays evidenced that a subset of tested derivatives efficiently inhibit topoisomerase II α accordingly to merbarone mechanism of action. However this property does not fully rationalize the cytotoxicity data of the full ligand panel, suggesting that different target(s) should be additionally involved.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Merbarone (**Chart 1**) is a thiobarbituric catalytic topoisomerase type II (topoII) inhibitor identified by the National Cancer Institute (Bethesda, MD, USA) screening program. Merbarone blocks the topoII-mediated DNA cleavage with some selectivity toward the II α isoform.^{1–3} In *in vitro* assays, merbarone exhibits curative activity against L1210 leukemia and important activity against some other murine tumors.⁴ However, clinical tests toward many different tumor types showed a general lack of antitumor activity,^{5–8} potentially ascribable to merbarone high ionizability (and therefore poor bioavailability) at physiological pH.

With the aim to prepare merbarone analogues endowed with improved pharmacodynamic and pharmacokinetic properties, we previously synthesized compounds **I** (**Chart 1**). They can be considered as conformational constrained merbarone analogues in which the H-bonded pseudo-bicyclic structure of merbarone has been converted in a pyrimidopyrimidine moiety.⁹ In MT-4 cell-based assays, derivatives **Ib** and **Ie** showed IC₅₀ values of 3.3 and 3.6 μ M, respectively and were identified as the most active compounds of the series (**Chart 1**). Moreover, **Ib** showed submicromolar activity against leukemia cell lines in NCI screening. Finally, **Ib** emerged

to be more efficient than merbarone in inhibiting topoII catalytic activity.⁹

To extend the structure activity relationship (SAR) studies on compounds **I**, the aza-isosteres **1** were designed and synthesized (**Chart 2**). The novel derivatives **1** share the triphenyl substituted pyrimidopyrimidine substructure which was functionalized at position 7 with side chains endowed with different steric, conformational and acid/base properties. To better rationalize the SARs, the substituent at position 7 was formally fragmented into three parts: connecting group (CG), linker and terminal group (TG) (**Chart 2**). The nitrogen atom of CG can be either a monosubstituted (compounds **1a–I**, **Scheme 1**) or a disubstituted (**1m**) amine group or part of an hydrazino functionality (**1n**). The linker portion separating CG and TG varied from one carbon atom to three carbon atoms. In compound **1m**, the linker portion was conformationally restricted in a 6-membered ring whereas in **1n** the linker portion was missing being the CG directly connected to the TG. The TGs inserted include tertiary amines (**1a–f**, **1m**), basic heteroaromatic rings (pyridine for **1g**, **h**, **j**, **k**, imidazole for **1l**) and phenyl moieties (**1i** and **1n**).

2. Chemistry

Compounds **1a–n** were prepared starting from the key intermediate **A** (**Scheme 1**) obtained as previously reported.⁹ Briefly, the condensation of malononitrile (1 molar equiv) and

* Corresponding author. Tel.: +39 010 353 8364; fax: +39 010 353 8358.

E-mail address: andrea.spallarossa@unige.it (A. Spallarossa).

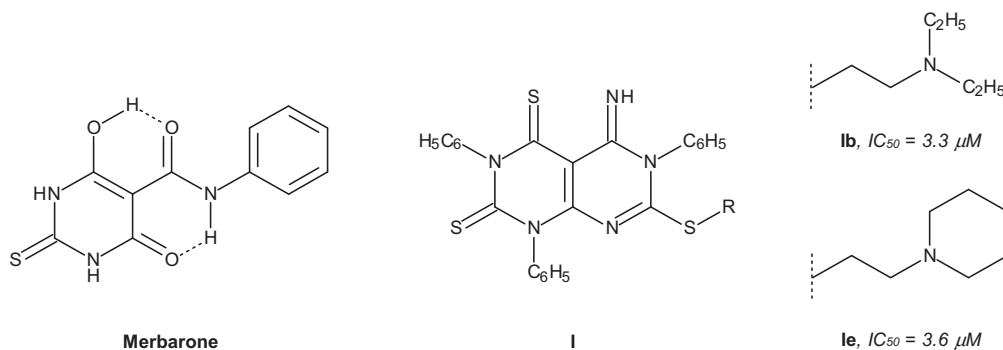


Chart 1. Molecular structures of merbarone and its rigid thioanalogues **1b** and **1e**.

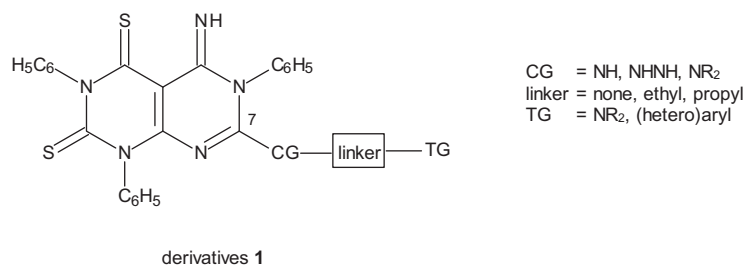


Chart 2. General structure of derivatives **1**. The three fragments (namely, connecting group (CG), linker and terminal group (TG)) in which the 7-substituent has been formally subdivided are indicated.

phenylisothiocyanate (3 molar equiv) in the presence of sodium hydride (1.5 molar equiv) led to the formation of the pyrimidopyrimidine scaffold which was then *S*-methylated at position 7 to afford product **A** in good yields.

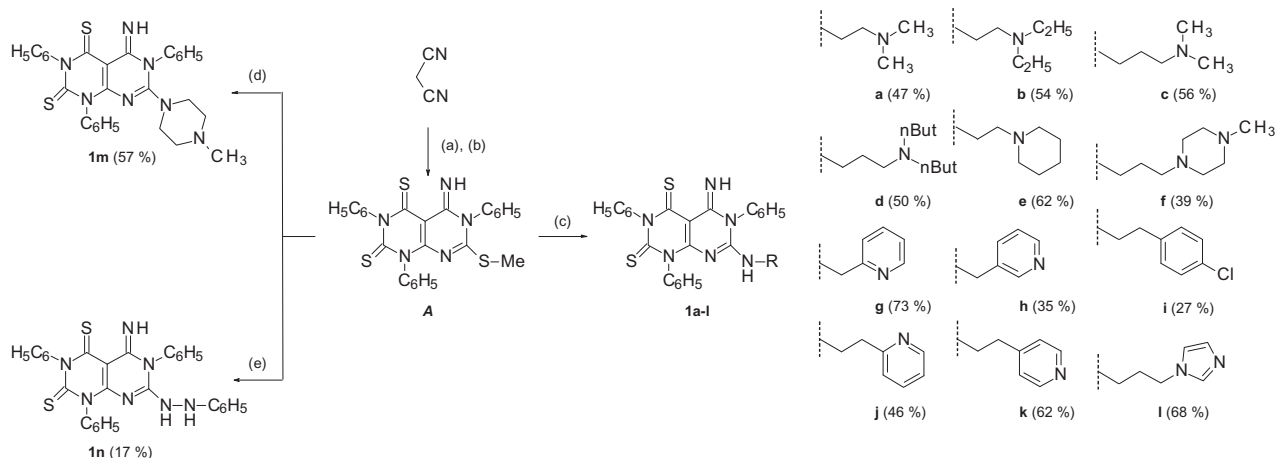
The desired compounds **1** were obtained by the displacement of **A** thiomethyl group with the nucleophiles **a–n** (Chart 3). To overcome the variable reactivity of **a–n** building blocks towards **A**, different synthetic procedures were set up. Thus, for amines **a–l** the reaction was carried out in dry DMF at rt either using a slight excess of the amine (1.2 equiv, for **a–f**) or by adding diisopropylethylamine (DIPEA, 1 equiv, for **g–l**). The steric bulkiness of *N*-methylpiperazine reduced the reactivity of the amine towards **A**. Therefore, more drastic solvent-free reaction conditions were adopted (120 °C for 5 h) and **1m** was obtained in good yield (57%). The hydrazo derivative **1n** was obtained in moderate yields

by refluxing the *S*-methyl intermediate and phenylhydrazine in a 1:1 mixture of acetonitrile and absolute ethanol.

3. Result and discussion

The antiproliferative properties of compounds **1a–n** were initially assessed in MT-4, HeLa and MCF-7 cell based assays (Table 1). Merbarone was considered as the reference compound.

The antiproliferative activity of compounds **1** appeared to be affected by the length of the linker and the nature of the terminal group (TG) and different trend were observed for the tested cell lines. Thus, the reduction of the TG steric hindrance did not lead to a consistent decrease in potency (compare **1a** and **1b**). The dimethylamino derivative **1a** was less active than its diethylamino congener **1b** against MT-4 cells but the two compounds showed



Scheme 1. Synthesis of derivatives **1a–n**. Yields are reported within brackets. Reagents and conditions: (a) C₆H₅NCS (1 equiv), NaH (1.5 equiv), dry DMF, 5 °C, then C₆H₅NCS (2 equiv), rt, 15 h; (b) CH₃I, NaHCO₃, rt, 12 h; (c) R-NH₂ (1.2 equiv), dry DMF, (DIPEA), rt, 12 h; (d) *N*-methylpiperazine, 120 °C, 5 h; (e) C₆H₅NHNH₂, ACN/EtOH, reflux, 10 h.

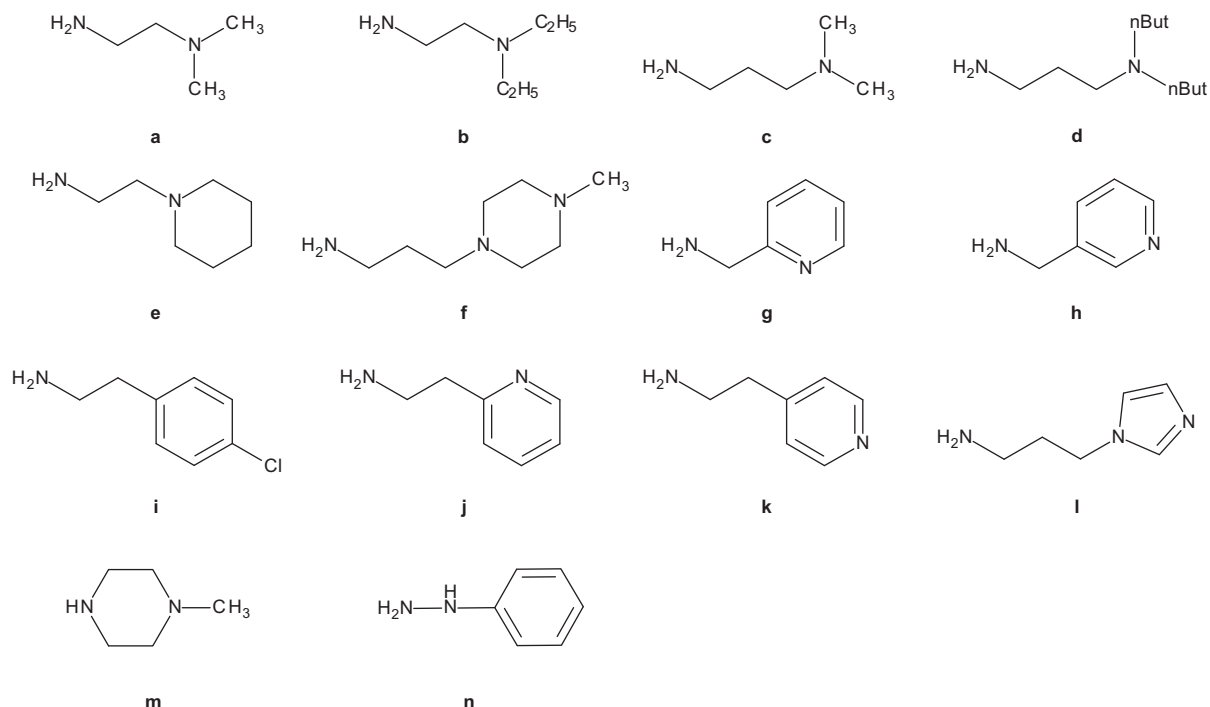


Chart 3. Amine and hydrazine building blocks used.

similar IC_{50} values in the HeLa and MCF-7 assays. Conversely, the elongation of the linker chain of one methylene unit raised up activity (compare **1a** and **1c**) which was further improved through the introduction of a bulky dibutylamino TG (compare **1d** and **1c**). Good antireplicative potency was also observed for derivative **1f** bearing the dibasic *N*-methylpiperazine TG. Interestingly, the direct anchoring of the *N*-methylpiperazine substructure on the pyrimido-pyrimidine scaffold cancelled activity (compare **1f** and **1m**). Even though with a reduction in potency in comparison with the MT-4 cell based assay, **1e** was found to be active also against HeLa and MCF-7 cell lines.

The 2-pyridyl derivative **1g** showed a micromolar antiproliferative activity against MT-4 cells but was ineffective against HeLa and MCF-7 cell lines. Conversely, its 3-pyridyl isomer **1h** showed an opposite cytotoxicity profile against the three considered cell lines. The elongation of the linker chain marginally effected the activity of the 2-pyridyl compounds against MT-4 cells whereas HeLa and MCF-7 activities were improved (compare **1g** and **1j**). The 4-pyridyl compound **1k** resulted to be slightly less potent than its congener **1j** against all the three cell lines. The 3-(1-imidazolyl)propyl congener **1l** paralleled the activity profile of the 2-pyridyl compound **1g** against all the considered cell lines.

Despite the lack of a basic TG, **1i** showed good micromolar activity against MT-4 and HeLa cells with a reduction in potency for the MCF-7 cell line. Conversely, the insertion of a phenylhydrazino substructure led to the inactive compound **1n** (data available for MT-4 cell only).

Furthermore, **1h** and **1i** showed similar IC_{50} values when tested as free bases or as hydrochloride salts thus indicating that the difference in activity among the cell lines cannot be ascribable to solubility issues. Noteworthy, the majority of the tested (namely, **1a–f**, **i**, **j**) compounds emerged to be more active than merbarone against all the considered cell lines (Table 1).

Compounds **1a**, **d**, **g** were selected by the National Cancer Institute (NCI) (Bethesda, MD, USA) for a screening on a panel of 60 different tumor cell-lines (Tables 2–4). Preliminarily, the derivatives were tested at a fixed concentration (10 μ M) and growth percent

inhibition was measured (Table 2). Compound **1g** did not show any significant antiproliferative activity (grow percentage range: 57.91–118.52%, data not shown) and therefore was discarded. Analogues **1a** and **1d** were further tested to assess the GI_{50} (measure of the growth inhibitory power), TGI (measure of the cytostatic activity) and LC_{50} (measure of the cytotoxic effect). As summarized in Table 3, both compounds showed antiproliferative activity against all the cell lines investigated (60 out of 60) and demonstrated to be cytostatic and cytotoxic at (sub)micromolar concentrations in a good number of cell lines (57–60 for **1a** and 56–57 for **1d**). The average GI_{50} values reported in Figure 1 indicate that both derivatives showed a widespread antiproliferative activity against all the considered cancer subpanels. In particular (Table 4), **1a** emerged as particularly effective against leukemia (K-562 cell line GI_{50} = 0.4 μ M; RPMI-8226 GI_{50} = 0.4 μ M) and melanoma (SK-MEL-28 cell line

Table 1
Cell based cytotoxicity for compounds 1

Compd	IC_{50}^a (μ M)		
	MT-4	HeLa	MCF-7
1a	8.4 \pm 0.2	12.6 \pm 1.1	16.0 \pm 2.1
1b	2.5 \pm 0.1	16.9 \pm 0.1	11.1 \pm 0.1
1c	1.7 \pm 0.05	7.8 \pm 0.5	5.5 \pm 0.5
1d	1.6 \pm 0.1	2.0 \pm 0.2 ^b	<1.0 ^b
1e	1.7 \pm 0.2	5.2 \pm 0.1 ^b	26.7 \pm 1.4 ^b
1f	2.0 \pm 0.3	8.5 \pm 0.3 ^b	5.7 \pm 0.6 ^b
1g	7.0 \pm 0.2	>50	>50
1h	>100	12.6 \pm 3.4 (14.6 \pm 1.4) ^b	24.3 \pm 0.8 ^b
1i	2.4 \pm 0.05	5.9 \pm 0.6 (5.6 \pm 0.9) ^b	25.6 \pm 1.3 ^b
1j	8.0 \pm 0.1	10.5 \pm 0.3	17.3 \pm 1.9
1k	16.0 \pm 1.0	15.4 \pm 2.4	17.8 \pm 0.8
1l	8.0 \pm 0.5	>100	>100
1m	>100	>50	>50
1n	>100	ND ^c	ND ^c
Merbarone	12 \pm 1	43.8 \pm 9.1	61.6 \pm 4.9

^a Data mean values for three separate experiments.

^b Activity of the hydrochloride salt.

^c Not determined.

Table 2Growth percent of different cancer cell lines in culture in the presence of **1a** and **1d** at the concentration of 10 μM ^a

Panel cell line	Compounds growth percent	
	1a	1d
<i>Leukemia</i>		
CCRF-CEM	−8.26	−78.41
HL-60 (TB)	36.60	−74.77
K-562	−11.39	−90.35
MOLT-4	3.92	−55.84
RPMI-8226	−4.62	−53.40
SR	34.73	−65.16
<i>Non-small cell lung cancer</i>		
A549/ATCC	35.35	−38.43
EKVX	39.76	−74.15
HOP-62	6.19	−99.55
HOP-92	n.d. ^b	−66.94
NCI-H226	82.30	−82.47
NCI-H23	67.64	−56.27
NCI-H322M	58.54	−57.77
NCI-H460	3.80	−95.68
NCI-H522	3.16	−54.78
<i>Colon cancer</i>		
COLO 205	6.65	−74.06
HCC-2998	−39.64	−91.46
HCT-116	−22.12	−24.97
HCT-15	18.84	−73.34
HT29	−35.69	−63.34
KM12	0.20	−100.00
SW-620	−59.11	−100.00
<i>CNS cancer</i>		
SF-268	24.00	−94.01
SF-295	−61.52	−100.00
SF-539	−86.90	−97.73
SNB-19	26.14	−87.44
SNB-75	79.04	−94.22
U251	2.10	−100.00
<i>Melanoma</i>		
LOX IMVI	−61.37	−89.69
MALME-3M	−89.51	−91.52
M14	−84.40	−86.29
MDA-M-435	−64.55	−100.00
SK-MEL-2	n.d. ^b	n.d. ^b
SK-MEL-28	−100.00	−100.00
SK-MEL-5	−98.77	−100.00
UACC-257	−74.63	−73.12
UACC-62	−90.31	−99.42
<i>Ovarian cancer</i>		
IGROV1	36.27	−70.87
OVCAR-3	−48.73	−100.00
OVCAR-4	41.73	−82.03
OVCAR-5	63.10	−60.25
OVCAR-8	33.57	−36.07
NCI/ADR-RES	58.75	−71.14
SK-OV-3	76.30	−86.52
<i>Renal cancer</i>		
786-0	−83.54	−100.00
A498	40.29	−100.00
ACHN	28.27	−95.78
CAKI-1	n.d. ^b	n.d. ^b
RXF 393	−39.10	−67.67
SN12C	52.27	−99.22
TK-10	47.73	−89.58
UO-31	3.10	−80.92
<i>Prostate cancer</i>		
PC-3	31.04	−70.77
DU-145	44.99	−100.00
<i>Breast cancer</i>		
MCF7	17.87	−88.05
MDA-MB-231/ATCC	−12.96	−89.07
HS 578T	12.31	−64.57
BT-549	4.45	−90.56

Table 2 (continued)

Panel cell line	Compounds growth percent	
	1a	1d
T-47D	−19.38	−62.09
MDA-MB-468	−34.62	−89.92

^a Data obtained from NCI's in vitro disease-oriented human tumor cell lines screen.^b Not tested.

$\text{GI}_{50} = 0.6 \mu\text{M}$) cancers; **1d** showed submicromolar GI_{50} values against a large number of cell lines (Table 4) being its average GI_{50} value less than 1 μM for all the considered cancer subpanels exception made for ovarian cancer (average $\text{GI}_{50} = 1.2 \mu\text{M}$, Fig. 1).

In order to assess whether derivatives **1** targeted topo-II, enzymatic assays on isolated human topoll α enzyme were carried out on selected members of the series. Merbarone was used as reference drug (Table 5). Derivatives **1h** and **1i** (tested either as free bases or as hydrochloride salts) were found to inhibit the enzyme more efficiently than the reference drug, being the IC_{50} values in the low micromolar concentration range. Similarly, the piperidine derivative **1e** emerged to be active against topoll α enzyme with an IC_{50} value of 3.0 μM . In the required buffer, above 50 μM ligand aggregation was verified for compounds **1d** and **1f**, thus preventing the analysis at the higher tested concentrations. Nevertheless, a well reproducible limited enzyme inhibition (10–15%) around 30 μM was observed suggesting a poor activity on the enzyme.

Spectrophotometric and electrophoretic analyses showed that neither **1h** nor **1i** were able to bind DNA or to stabilize cleavage complex thus supporting a common inhibition mechanism towards topoll α (Supplementary Figs. S1 and S2). Interestingly, these findings were in line with previous data on merbarone.¹

To address if topoisomerases of different types can be equally sensitive to our derivatives, compound **1i** was further tested against *Escherichia coli* Topoisomerase IV, *E. coli* Topoisomerase I and human Topoisomerase I. The compound was not able to inhibit the catalytic activity of the above listed enzymes (data not shown).

If the efficiency on topoll α exerted by **1e**, **1h** and **1i** correlated with the observed antiproliferative activity, not the same occurred with **1c**, **1d** and **1f**. Indeed, these compounds, which share a propyl linker between the bicyclic scaffold and the basic TG, were not significantly active against topoll α (Table 5). In order to shed some light on the mechanism of action, a COMPARE analysis was performed on compound **1d**.^{10–16} This bioinformatic tool correlates the antiproliferative profiles of two agents calculating a Pearson correlation coefficient (PCC). A high PCC suggests that the two molecules share a similar antiproliferative mechanism. Thus, the individual **1d** screening results at the GI_{50} end point were used as probes to search the NCI standard agent compound set containing 171 agents with confirmed mechanism(s) of action. A significant ($p < 0.01$) correlations involved **1d** and spirogermanium and the estrogen receptor antagonist tamoxifen (correlation coefficients: 0.650 and 0.643, respectively, Table 6). Interestingly, a recent investigation,¹⁷ pointed out the ability of **1d** to bind to the nonclassical estrogen receptor GPER/GPR30 thus supporting the indication provided by the COMPARE analysis.

4. Conclusion

The data acquired on the antiproliferative activity of derivatives **1** showed that the biological profile of the prepared compounds is deeply affected by the nature of the substituent at position 7 of the

Table 3
Anticancer activity of **1a** and **1d**^a

Compound	Investigated	Number (No) of human tumor cell lines ^b Giving positive GI ₅₀ , TGI and LC ₅₀					
		GI ₅₀ ^c (μM)		TGI ^d (μM)		LC ₅₀ ^e (μM)	
		No	Range	No	Range	No	Range
1a	60	60	0.4–16.8	60	2.1–42.1	57	5.3–91.0
1d	60	60	0.1–1.9	57	0.3–3.8	56	0.6–36.8

^a Data obtained from NCI's in vitro disease-oriented human tumor cell lines screen.

^b The table shows the number of cell lines against which each compound was screened, the number of lines against which it gave a positive GI₅₀, or TGI or LC₅₀ value (<100 μM) and the corresponding concentration range.

^c Compound concentration that produces 50% growth inhibition.

^d Compound concentration that produces total growth inhibition.

^e Compound concentration that produces 50% cytotoxic effect.

pyrimido-pyrimidine scaffold. In cell-based assays, derivatives **1d–f** emerged as the most promising compounds in terms of either antiproliferative potency or spectrum of action. In enzymatic assays, topoll α was identified as a potential macromolecular target for derivatives **1e, h, i**. Noteworthy, compounds **1d** and **1f** sharing a ω -aminopropyl chain were avoided of any topoll α activity thus suggesting different mechanisms of action within the same chemical class.

5. Experimental protocols

5.1. Chemistry

All building blocks used are commercially available. Amines (**a–m**, Chart 3), phenylhydrazine **n**, phenylisothiocyanate and 60% sodium hydride dispersion in mineral oil were purchased by Chiminord and Aldrich Chemical, Milan (Italy). Intermediate **A** was prepared as previously reported.¹⁸ Solvents were reagent grade. DMF was dried on molecular sieves (5 Å 1/16" inch pellets). Unless otherwise stated, all commercial reagents were used without further purification. Organic solutions were dried over anhydrous sodium sulphate. Thin layer chromatography (TLC) system for routine monitoring the course of reactions and confirming the purity of analytical samples employed aluminium-backed silica gel plates (Merck DC-Alufolien Kieselgel 60 F₂₅₄): CHCl₃ was used as developing solvent and detection of spots was made by UV light and/or by iodine vapours. Yields were not optimized. Melting points were determined on a Fisher-Johns apparatus and are uncorrected. IR spectra were recorded on a Perkin Elmer 398 spectrometer as KBr discs. ¹H NMR spectra (200 μMHz) were recorded in CDCl₃ on a Varian Gemini 200 instrument. Chemical shifts were reported in δ (ppm) units relative to the internal standard tetramethylsilane, and the splitting patterns were described as follows: s (singlet), t (triplet) and m (multiplet). The first order values reported for coupling constants *J* were given in Hz. Elemental analyses were performed by an EA1110 Elemental Analyser (Fison-Instruments, Milan); all compounds were analyzed for C, H, N and S and the analytical results were within $\pm 0.4\%$ of the theoretical values.

5.1.1. General procedure for the synthesis of compounds **1a–l**

Intermediate **A** (0.97 g, 2 mmol) was added in a single portion at rt to a stirred solution of the proper amine (**a–l**, 2.4 mmol) in dry DMF (10 mL). For derivatives **1a–f**, the mixtures were stirred overnight at rt and then heated at 60 °C for 3 h. For compounds **1g–l**, DIPEA (0.19 mL, 2 mmol) was added to the mixtures after 6 h at rt. Stirring was prolonged at rt for 6 h and then mixtures were heated at 60 °C for 3 h. After addition of cold water (50 mL), a solid precipitated. The raw solid was collected by filtration and purified by crystallization from DCM and absolute ethanol. Compounds **1d–f, h, i** were converted into the corresponding

hydrochloride salts by using a 2 N HCl solution in absolute ethanol. Hydrochlorides purity was assessed by elemental analyses.

5.1.1.1. 7-[[2-(Dimethylamino)ethyl]amino]-5-imino-1,3,6-triphenyl-5,6-dihydropyrimido[4,5-*d*]pyrimidine-2,4(1*H*,3*H*)-dithione (**1a**).

Mp: 234–236 °C; yield: 47%. IR (KBr) cm⁻¹: 3186; 1625; 1575. ¹H NMR (CDCl₃) δ : 1.94 (s, 6H, 2 CH₃); 2.66–2.97 (m, 2H, CH₂N(CH₃)₂); 3.23–3.70 (m, 2H, CH₂N); 6.20–8.30 (m, 16H, 15 arom H and amine NH, exchangeable); 11.5 (bs, 1H, imine NH, exchangeable). Anal. Calcd for C₂₇H₂₅N₇S₂: C, 63.38; H, 4.92; N, 19.16; S, 12.63. Found: C, 63.40; H, 4.98; N, 19.27; S, 12.52.

5.1.1.2. 7-[[2-(Diethylamino)ethyl]amino]-5-imino-1,3,6-triphenyl-5,6-dihydropyrimido[4,5-*d*]pyrimidine-2,4(1*H*,3*H*)-dithione (**1b**).

Mp: 153–155 °C; yield: 54%. IR (KBr) cm⁻¹: 3279; 1627; 1574. ¹H NMR (CDCl₃) δ : 0.65 (t, *J* = 7.2 Hz, 6H, 2 CH₃); 1.96–2.36 (m, 6H, 3 CH₂N); 2.58–2.88 (m, 2H, CH₂NH); 6.68–8.13 (m, 16H, 15 arom H and amine NH, exchangeable); 11.43 (bs, 1H, imine NH, exchangeable). Anal. Calcd for C₂₈H₂₇N₇S₂: C, 63.97; H, 5.18; N, 18.65; S, 12.20. Found: C, 63.99; H, 5.27; N, 18.68; S, 12.05.

5.1.1.3. 7-[[3-(Dimethylamino)propyl]amino]-5-imino-1,3,6-triphenyl-5,6-dihydropyrimido[4,5-*d*]pyrimidine-2,4(1*H*,3*H*)-dithione (**1c**).

Mp: 217–220 °C; yield: 56%. IR (KBr) cm⁻¹: 3174; 1626; 1572. ¹H NMR (CDCl₃) δ : 1.08–1.34 (m, 2H, C–CH₂–C); 1.37–1.94 (m, 6H, 2 CH₃); 2.01–2.34 (m, 2H, CH₂N(CH₃)₂); 2.67–3.02 (m, 2H, CH₂N); 6.44–8.04 (m, 16H, 15 arom H and amine NH, exchangeable); 11.5 (bs, 1H, imine NH, exchangeable). Anal. Calcd for C₂₉H₂₉N₇S₂: C, 64.54; H, 5.42; N, 18.17; S, 11.08. Found: C, 64.74; H, 5.48; N, 17.94; S, 10.85.

5.1.1.4. 7-[[3-(Dibutylamino)propyl]amino]-5-imino-1,3,6-triphenyl-5,6-dihydropyrimido[4,5-*d*]pyrimidine-2,4(1*H*,3*H*)-dithione (**1d**).

Mp: 180–182 °C; yield: 50%. IR (KBr) cm⁻¹: 3189; 1624; 1568. ¹H NMR (CDCl₃) δ : 0.91 (t, *J* = 6.8 Hz, 6H, 2 CH₃); 1.25–1.32 (m, 4H, 2 CH₂Me); 1.36–1.60 (m, 6H, 3 C–CH₂–C); 2.12 (s, 1H, NH amine, exchangeable); 2.74–2.87 (m, 4H, 2 CH₂N but); 2.96–3.03 (m, 4H, 2 CH₂N); 7.14–7.20 e 7.34–7.45 e 7.66–7.69 (m, 15H, arom H); 9.25 (bs, 1H, amine NH, exchangeable); 12.88 (s, 1H, imine NH, exchangeable). Anal. Calcd for C₃₅H₄₁N₇S₂: C, 67.38; H, 6.62; N, 15.72; S, 10.28. Found: C, 67.38; H, 6.70; N, 15.68; S, 10.03. Anal. Calcd for C₃₅H₄₁N₇S₂·3HCl: C, 57.33; H, 6.05; N, 13.37; S, 8.75. Found: C, 57.73; H, 6.48; N, 13.37; S, 8.10.

5.1.1.5. 5-Imino-1,3,6-triphenyl-7-[(2-piperidin-1-ylethyl)amino]-5,6-dihydropyrimido[4,5-*d*]pyrimidine-2,4(1*H*,3*H*)-dithione (**1e**).

Mp: 245–247 °C; yield: 62%. IR (KBr) cm⁻¹: 3168; 1624; 1575. ¹H NMR (CDCl₃) δ : 0.64–1.50 (m, 6H, 3 C–CH₂–C) 1.83–2.53 (m, 6H, 3 CH₂Npip); 2.61–2.90 (m, 2H, CH₂N);

Table 4
GI₅₀, TGI and LC₅₀ values of compounds **1a** and **1d** against different cancer cell lines in culture^a

Panel cell line	Compound					
	1a			1d		
	GI ₅₀ ^b (μM)	TGI ^c (μM)	LC ₅₀ ^d (μM)	GI ₅₀ ^b (μM)	TGI ^c (μM)	LC ₅₀ ^d (μM)
<i>Leukemia</i>						
CCRF-CEM	1.4	3.6	9.2	0.1	0.3	0.6
HL-60 (TB)	1.5	3.3	7.1	0.2	—	—
K-562	0.4	2.2	6.8	0.2	—	—
MOLT-4	1.8	4.7	17.2	0.3	—	—
RPMI-8226	0.4	2.2	7.3	0.2	0.3	0.7
SR	1.4	3.3	8.2	0.3	1.5	7.0
<i>Non-small cell lung cancer</i>						
A549/ATCC	4.1	18.6	70.4	0.7	2.4	8.3
EKVX	12.1	32.2	85.8	0.4	1.9	4.7
HOP-62	2.9	10.1	55.6	1.2	2.6	5.4
HOP-92	1.8	10.3	60.1	0.2	0.8	3.5
NCI-H226	11.6	26.3	59.5	1.4	2.8	5.5
NCI-H23	16.8	42.1	—	1.6	3.4	8.4
NCI-H322M	6.2	20.7	53.8	1.4	3.2	7.4
NCI-H460	3.2	11.8	61.2	0.4	1.6	4.0
NCI-H522	2.1	5.4	27.6	0.2	1.3	5.1
<i>Colon cancer</i>						
COLO 205	1.6	3.8	9.1	0.2	1.4	6.0
HCC-2998	2.3	6.8	43.0	0.4	1.8	4.7
HCT-116	2.7	16.0	—	0.3	1.1	3.4
HCT-15	1.4	4.5	17.7	0.3	1.1	3.4
HT29	1.4	3.4	8.3	0.2	0.5	1.8
KM12	4.3	14.6	49.9	0.4	1.5	3.8
SW-620	2.1	5.1	18.8	0.4	1.4	3.7
<i>CNS cancer</i>						
SF-268	4.1	15.3	45.6	1.0	2.2	4.7
SF-295	2.9	8.3	51.1	1.2	2.7	5.8
SF-539	1.8	3.5	7.0	1.2	2.7	5.7
SNB-19	4.9	19.0	50.0	1.2	2.7	5.8
SNB-75	2.4	7.5	31.9	0.4	1.6	4.1
U251	1.5	3.4	7.8	0.3	1.2	3.5
<i>Melanoma</i>						
LOX IMVI	1.4	2.7	5.3	0.2	0.3	0.6
MALME-3M	2.6	5.0	9.5	0.2	0.5	1.5
M14	1.8	4.3	89.1	0.2	0.4	0.8
MDA-MB-435	1.9	4.7	22.1	0.2	0.5	1.5
SK-MEL-2	1.7	3.4	6.7	1.7	3.5	6.9
SK-MEL-28	0.6	2.1	5.8	0.2	0.3	0.7
SK-MEL-5	1.6	3.1	6.1	0.9	2.1	4.6
UACC-257	1.7	3.6	7.5	0.2	0.4	0.9
UACC-62	1.8	3.5	6.6	0.3	1.2	3.5
<i>Ovarian cancer</i>						
IGROV1	8.1	23.0	57.5	1.9	3.8	7.5
OVCAR-3	3.0	12.5	45.4	0.8	2.1	4.6
OVCAR-4	3.1	10.8	—	0.6	1.9	4.4
OVCAR-5	12.1	29.5	72.1	1.4	2.8	5.5
OVCAR-8	4.1	15.7	58.7	0.9	2.7	—
NCI/ADR-RES	10.5	31.0	91.0	1.6	3.3	8.1
SK-OV-3	10.3	24.1	56.7	1.5	2.9	5.8
<i>Renal cancer</i>						
786-0	1.4	4.3	21.4	0.2	0.4	0.8
A498	2.4	6.8	28.1	1.1	2.4	5.2
ACHN	3.2	11.3	34.1	0.5	1.6	4.1
CAKI-1	3.6	17.6	72.0	0.1	0.8	3.1
RXF 393	1.2	3.1	8.1	0.2	0.5	36.8
SN12C	6.2	19.9	48.0	1.2	2.4	4.9
TK-10	3.5	14.3	48.8	1.3	2.7	5.6
UO-31	1.6	3.9	9.6	0.2	0.5	1.1
<i>Prostate cancer</i>						
PC-3	4.0	17.9	51.9	0.6	1.9	4.6
DU-145	5.3	18.8	48.8	1.1	2.3	4.8
<i>Breast cancer</i>						

(continued on next page)

Table 4 (continued)

Panel cell line	Compound					
	1a			1d		
	GI ₅₀ ^b (μM)	TGI ^c (μM)	LC ₅₀ ^d (μM)	GI ₅₀ ^b (μM)	TGI ^c (μM)	LC ₅₀ ^d (μM)
MCF7	2.2	10.9	48.8	0.3	1.3	3.6
MDA-MB-231/ATCC	1.8	5.6	24.7	0.3	0.9	3.0
HS 578T	2.3	7.7	40.0	0.5	1.9	5.5
BT-549	5.3	21.3	62.1	1.5	3.2	6.7
T-47D	2.1	5.4	30.5	0.2	0.5	8.3
MDA-MB-468	1.2	2.7	6.0	0.2	0.3	0.6

^a Data obtained from NCI's in vitro disease-oriented human tumor cell lines screen.

^b Compound concentration that produces 50% growth inhibition.

^c Compound concentration that produces total growth inhibition. Only the values <100 μM are reported.

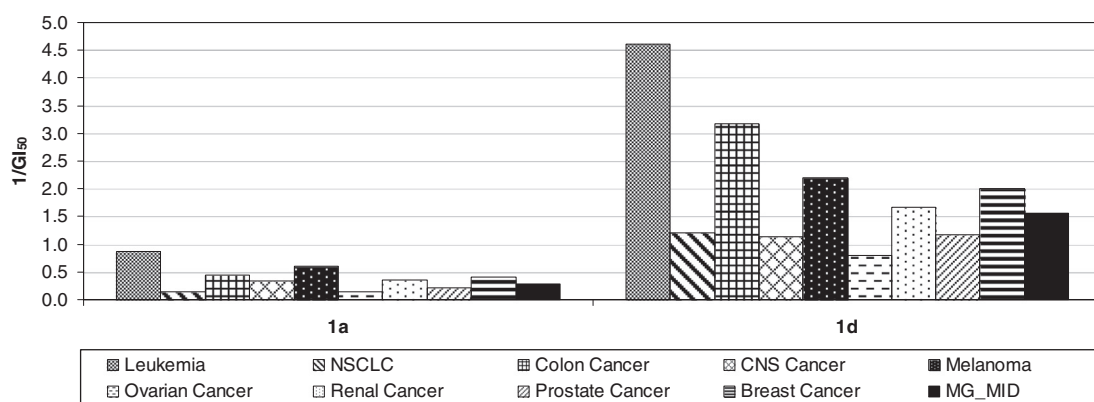


Figure 1. Mean Growth Inhibition (GI₅₀) values of 1a and 1d. For each compound, the average GI₅₀ (μM) concentration has been calculated both for each tumour subpanel and for all cell lines (Mean Graph Midpoint, MG_MID). The so obtained values were converted in the corresponding 1/GI₅₀ and plotted.

Table 5
Enzymatic activity human topollα

	IC ₅₀ (μM) ^a
1a	>100
1b	>100
1c	>100
1d	>32 ^b
1e	3.0 (3.0) ^b
1f	>35 ^b
1g	>100
1h	3.0 (3.0) ^b
1i	2.0 (2.0) ^b
1j	>100
1k	>100
Merbarone	32

^a Data mean values for three separate experiments.

^b Activity of the hydrochloride salt.

Table 6
Top ten results of the COMPARE analysis of compound 1d

Rank	Target vector descriptor	PCC	Common cell lines
1	Spirogermanium	0.650	56
2	Tamoxifen	0.643	60
3	Fluorodopan	0.610	43
4	Thaliciparine	0.600	50
5	Tamoxifen	0.567	57
6	Caracemide	0.563	58
7	L-Cysteine analogue	0.552	58
8	Thaliciparine	0.539	58
9	D-Tetrandrine	0.533	58
10	Dichloroallyl lawsone	0.512	59

6.22–8.30 (m, 16H, 15 arom H and amine NH, exchangeable); 11.37 (bs, 1H, imine NH, exchangeable). Anal. Calcd for C₃₁H₃₁N₇S₂: C, 65.81; H, 5.52; N, 17.33; S, 11.33. Found: C, 65.77; H, 5.87; N, 17.38; S, 11.03. Anal. Calcd for C₃₁H₃₁N₇S₂·3HCl: C, 55.15; H, 5.08; N, 14.52; S, 9.50. Found: C, 55.29; H, 5.40; N, 14.49; S, 9.11.

5.1.1.6. 5-Imino-7-[[3-(4-methylpiperazin-1-yl)propyl]amino]-1,3,6-triphenyl-5,6-dihydropyrimido[4,5-d]pyrimidine-2,4(1H,3H)-dithione (1f). Mp: 142–144 °C; yield: 39%. IR (KBr) cm⁻¹: 3168; 1624; 1575. ¹H NMR (CDCl₃) δ: 1.06–1.58 (m, 2H, C–CH₂–C); 2.25 (s, 3H, CH₃); 1.95–2.47 (m, 10H, 5 CH₂N); 3.20–3.72 (m, 2H, NHCH₂–C); 6.78–8.35 (m, 16H, 15 arom H and amine NH, exchangeable); 11.45 (bs, 1H, imine NH, exchangeable). Anal. Calcd for C₃₂H₃₄N₈S₂: C, 64.62; H, 5.76; N, 18.84; S, 10.78. Found: C, 64.62; H, 5.99; N, 18.62; S, 10.68. Anal. Calcd for C₃₂H₃₄N₈S₂·3HCl: C, 54.58; H, 5.30; N, 15.91; S, 9.11. Found: C, 54.95; H, 5.80; N, 15.91; S, 9.28.

5.1.1.7. 5-Imino-1,3,6-triphenyl-7-[(pyridin-2-ylmethyl)amino]-5,6-dihydropyrimido[4,5-d]pyrimidine-2,4(1H,3H)-dithione (1g). Mp: 260–262 °C; yield: 73%. IR (KBr) cm⁻¹: 3195; 1629; 1576. ¹H NMR (CDCl₃) δ: 4.04 (s, 2H, CH₂N); 6.55–8.56 (m, 21H, 19 arom H, imine NH and amine NH, exchangeable). Anal. Calcd for C₃₀H₂₃N₇S₂: C, 66.03; H, 4.25; N, 17.97; S, 11.75. Found: C, 66.04; H, 4.35; N, 18.25; S, 11.59.

5.1.1.8. 5-Imino-1,3,6-triphenyl-7-[(pyridin-3-ylmethyl)amino]-5,6-dihydropyrimido[4,5-d]pyrimidine-2,4(1H,3H)-dithione (1h). Mp: 264–266 °C; yield: 35%. IR (KBr) cm⁻¹: 3175; 1625; 1569. ¹H NMR (DMSO) δ 3.95 (s, 2H, CH₂N); 6.30–9.0 (m, 21H, 19 arom H, imine NH and amine NH, exchangeable). Anal. Calcd for

C₃₀H₂₃N₇S₂: C, 66.03; H, 4.25; N, 17.97; S, 11.75. Found: C, 66.17; H, 4.53; N, 18.05; S, 11.34. Anal. Calcd for C₃₀H₂₃N₇S₂·3 HCl: C, 55.01; H, 4.00; N, 14.97; S, 9.79. Found: C, 54.79; H, 4.36; N, 14.63; S, 9.51.

5.1.1.9. 7-[[2-(4-Chlorophenyl)ethyl]amino]-5-imino-1,3,6-triphenyl-5,6-dihydropyrimido[4,5-d]pyrimidine-2,4(1H,3H)-dithione (1i). Mp: 236–238 °C; yield: 27%. IR (KBr) cm⁻¹: 3329; 1644; 1569. ¹H NMR (CDCl₃) δ: 2.45–2.48 (m, 2H, C–CH₂–Ph); 6.66–6.70 and 7.06–7.69 (m, 19H, arom H); 8.80 (bs, 1H, amine NH, exchangeable); 12.90 (s, 1H, imine NH, exchangeable). Anal. Calcd for C₃₂H₂₅ClN₆S₂: C, 64.80; H, 4.25; N, 14.17; S, 10.81. Found: C, 64.84; H, 4.27; N, 14.21; S, 10.69. Anal. Calcd for C₃₂H₂₅ClN₆S₂·2HCl: C, 57.70; H, 4.09; N, 12.62; S, 9.36. Found: C, 57.81; H, 4.25; N, 12.30; S, 9.27.

5.1.1.10. 5-Imino-1,3,6-triphenyl-7-[(2-pyridin-2-ylethyl)amino]-5,6-dihydropyrimido[4,5-d]pyrimidine-2,4(1H,3H)-dithione (1j). Mp: 250–252 °C; yield: 46%. IR (KBr) cm⁻¹: 3195; 1625; 1571. ¹H NMR (DMSO) δ: 2.87–3.05 (m, 2H, CH₂N); 3.20–3.28 (m, 2H, CH₂N); 6.85–7.51 (m, 20H, 19 arom H and amine NH, exchangeable); 11.16 (bs, 1H, imine NH, exchangeable). Anal. Calcd for C₃₁H₂₅N₇S₂: C, 66.52; H, 4.50; N, 17.52; S, 11.46. Found: C, 66.55; H, 4.52; N, 17.56; S, 11.37.

5.1.1.11. 5-Imino-1,3,6-triphenyl-7-[(2-pyridin-4-ylethyl)amino]-5,6-dihydropyrimido[4,5-d]pyrimidine-2,4(1H,3H)-dithione (1k). Mp: 254–257 °C; yield: 62%. IR (KBr) cm⁻¹: 3156; 1624; 1575. ¹H NMR (CDCl₃) δ: 2.40 (bs, 1H, NH amine, exchangeable); 3.0–3.3 (m, 2H, CH₂–py); 3.60–3.80 (m, 2H, CH₂N); 6.55–7.74 e 8.22–8.58 (m, 20H, 19 arom H and amine NH, exchangeable); 11.87 (bs, 1H, imine NH, exchangeable). Anal. Calcd for C₃₁H₂₅N₇S₂: C, 66.52; H, 4.50; N, 17.52; S, 11.46. Found: C, 66.47; H, 4.88; N, 17.52; S, 10.62.

5.1.1.12. 7-[[3-(1-Imidazolil)propil]amino]-5-immino-1,3,6-trifenil-5,6-diidropirimido[4,5-d]pirimidin-2,4(1H,3H)-dithione (1l). Mp: 178–180 °C; yield: 68%. IR (KBr) cm⁻¹: 3155; 1622; 1572. ¹H NMR (DMSO) δ: 1.25–1.44 (m, 2H, C–CH₂–Ph); 2.41–2.43 (t, 2H, CH₂–imidaz); 2.60 (bs, 1H, NH amine, exchangeable); 3.27–3.43 (m, 2H, CH₂–N); 6.78 (s, 1H, H₂ imidaz); 6.93 (s, 2H, H_{4,5} imidaz); 7.06–7.16 and 7.27–7.46 (m, 16H, 15 arom. H and amine NH, exchangeable); 11.10 (bs, 1H, imine NH, exchangeable). Anal. Calcd for C₃₀H₂₆N₈S₂: C, 64.03; H, 4.66; N, 19.91; S, 11.40. Found: C, 64.05; H, 4.67; N, 19.94; S, 11.27.

5.1.2. Synthesis of 5-imino-7-(4-methylpiperazin-1-yl)-1,3,6-triphenyl-5,6-dihydropyrimido[4,5-d]pyrimidine-2,4(1H,3H)-dithione (1m)

A mixture of **A** (0.96 g, 2 mmol) and 4-methylpiperazine (5 mL) was heated at 120 °C for 6 h, prolonging stirring at rt overnight. The addition of cold water (50 mL) caused the precipitation of a solid which was collected by filtration and purified by crystallization from DCM and EtOH.

Mp: 262–264 °C; yield: 57%. IR (KBr) cm⁻¹: 3024; 1609; 1587. ¹H NMR (CDCl₃) δ: 2.26 (s, 3H, NCH₃); 1.81–3.06 (m, 7H, NCH₃ and 2 CH₂N); 3.16–3.86 (m, 4H, 2 CH₂N); 6.50–8.64 (m, 15H, arom. H); 13.53 (bs, 1H, imine NH, exchangeable). Anal. Calcd for C₂₉H₂₇N₇S₂: C, 64.76; H, 5.06; N, 18.23; S, 11.92. Found: C, 64.72; H, 5.01; N, 18.25; S, 11.87.

5.1.3. Synthesis of 5-imino-1,3,6-triphenyl-7-(2-phenylhydrazino)-5,6-dihydropyrimido[4,5-d]pyrimidine-2,4(1H,3H)-dithione (1n)

To an ACN/EtOH mixture (3:1, 20 mL) of **A** (0.96 g, 2 mmol), phenylhydrazine (3 mmol, 0.31 g) was added and the mixture

was refluxed for 10 h. The solvent was evaporated under reduced pressure and the oily residue was treated with water (100 mL) and extracted with DCM (2 × 20 mL). The pooled organic phases were washed with water (4 × 10 mL), dried over anhydrous Na₂SO₄ and filtered through a pad of Florisil. Evaporating in vacuo gave a residue which was crystallized from EtOH/DCM mixture.

Mp: 293–295 °C; yield: 17%. IR (KBr) cm⁻¹: 3424; 1617; 1584. ¹H NMR (CDCl₃) δ: 6.48–8.25 (m, 23H, 20 arom. H, imine NH and amine NH, exchangeable). Anal. Calcd for C₃₀H₂₃N₇S₂: C, 66.03; H, 4.25; N, 17.97; S, 11.75. Found: C, 66.14; H, 4.26; N, 17.98; S, 11.70.

5.2. Biology

5.2.1. Cell based assays

The biological evaluation of the compounds **1a–n** against MT-4 cells was performed according to the previously reported procedures.¹⁸ The screening against HeLa and MCF7 cells were carried out as previously reported.¹⁹

5.2.2. NCI-screening

The NCI high-flux anticancer drug screen utilized a panel of 60 human tumor cell lines in culture derived from nine cancer types (lung, colon, CNS, ovarian, renal, prostate and breast cancer, leukemia and melanoma).^{15,16,20} The compound were tested at 10-fold dilutions of five concentrations ranging from 10⁻⁴ to 10⁻⁸ M. According to the NCI protocol, cell lines were exposed to test agents in 96-well plates for the last 48 of a 72 h incubation and a sulforhodamine B (SRB) protein assay was used to estimate cell viability or growth. For each compound, the drug concentration required to produce 50% (GI₅₀) and total (TGI) growth inhibition, and 50% cytotoxic effect (LC₅₀) were obtained for 56–58 cell lines. Values were calculated for each of these parameters if the level activity was reached; if the effect was not reached or was exceeded, the value is expressed as greater or lesser than the maximum or minimum concentration tested.

5.2.3. Enzymatic assay

Supercoiled plasmid pBR322 (Fermentas, 0.15 μg) was incubated with 1 U Topoisomerases (Inspiralis) in the provided buffer for 1 h at 37 °C. Reaction products were resolved on 1% agarose gels in TBE 0.5X (45 mM Tris, 45 mM boric acid, 1 mM Na₂EDTA) and the DNA bands were visualized by ethidium bromide staining and photographed. The relative amounts of different DNA topoisomers were quantified using a Geliance 2000 apparatus. Each assay was performed at least in triplicate and error were found ±10%.

Acknowledgements

The authors are very grateful to Dr. V.L. Narayanan, Chief of Drug Synthesis and Chemistry branch, National Cancer Institute (Bethesda, MD, USA) for the in vitro evaluation of anticancer activity. This work was supported by the University of Genova (CUP D31J11001610005), University of Padova (PRAT CPDA114388) and MIUR (Ministero dell'Istruzione, dell'Università' e della Ricerca, PRIN#2009MFRKZ8_004 and PRIN#20105YY2HL_004). Fondazione Carige is gratefully acknowledged for financially supporting C.R. The authors thank Mr. O. Gagliardo for the microanalyses, Mr. F. Tuberoni, Dr. C. Rossi, Dr. M. Anzaldi and Dr. R. Raggio for the IR and NMR spectra.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bmc.2013.08.056>.

References and notes

1. Fortune, J. M.; Osheroff, N. J. *Biol. Chem.* **1998**, *273*, 17643.
2. Drake, F. H.; Hofmann, G. A.; Bartus, H. F.; Mattern, M. R.; Crooke, S. T.; Mirabelli, C. K. *Biochemistry* **1989**, *28*, 8154.
3. Drake, F. H.; Hofmann, G. A.; Mong, S. M.; Bartus, J. O.; Hertzberg, R. P.; Johnson, R. K.; Mattern, M. R.; Mirabelli, C. K. *Cancer Res.* **1989**, *49*, 2578.
4. Glover, A.; Chun, H. G.; Kleinman, L. M.; Cooney, D. A.; Plowman, J.; Grieshaber, C. K.; Malspeis, L.; Leyland-Jones, B. *Invest. New Drugs* **1987**, *5*, 137.
5. Dimaggio, J. J.; Warrel, R. P.; Muindi, J.; Stevens, Y. W.; Lee, S. J.; Lowenthal, D. A.; Haines, L.; Walsh, T. D.; Baltzer, L.; Yaldaei, S. *Cancer Res.* **1990**, *50*, 1151.
6. Kraut, E. H.; Fleming, T.; Macdonald, J. S.; Spiridonidis, C. H.; Bradof, J. E.; Baker, L. H. *Am. J. Clin. Oncol.* **1993**, *16*, 327.
7. Ajani, J. A.; Winn, R.; Baez, L.; Pollock, T.; Maher, T.; Hallinan-Fueger, B.; Newman, J. *Cancer Invest.* **1994**, *12*, 488.
8. Malik, U. R.; Dutcher, J. P.; Caliendo, G.; Lasala, P.; Mitnick, R.; Wiernik, P. H. *Med. Oncol.* **1997**, *14*, 159.
9. Ranise, A.; Spallarossa, A.; Schenone, S.; Bruno, O.; Bondavalli, F.; Pani, A.; Marongiu, M. E.; Mascia, V.; La Colla, P.; Loddo, R. *Bioorg. Med. Chem.* **2003**, *11*, 2575.
10. Boyd, M. R. In *Cancer: Principles and Practice of Oncology*; DeVita, V. T. J., Hellman, S., Rosenberg, S. A., Eds.; ; Lippincott: Philadelphia, 1989; Vol. 3, pp 1–12.
11. Boyd, M. R.; Paull, K. D.; Rubinstein, L. R. In *Cytotoxic Anticancer Drugs: Models and Concepts for Drug Discovery and Development*; Vlierote, F. A., Corbett, T. H., Baker, L. H., Eds.; Kluwer Academic: Hingham, 1992; pp 11–34.
12. Paull, K. D.; Shoemaker, R. H.; Hodes, L.; Monks, A.; Scudiero, D. A.; Rubinstein, L.; Plowman, J.; Boyd, M. R. *J. Natl. Cancer Inst.* **1989**, *81*, 1088.
13. Paull, K. D.; Lin, C. M.; Malspeis, L.; Hammel, E. *Cancer Res.* **1992**, *52*, 3892.
14. Bai, R.; Paull, K. D.; Herald, C. L.; Pettit, G. R.; Hamel, E. *J. Biol. Chem.* **1991**, *266*, 15882.
15. Monks, A.; Scudiero, D.; Skehan, P.; Shoemaker, R.; Paull, K.; Vistica, D.; Hose, C.; Langley, J.; Cronise, P.; Vaigro-Wolff, A.; Gray-Goodrich, M.; Campbell, H.; Mayo, J.; Boyd, M. *J. Natl. Cancer Inst.* **1991**, *83*, 757.
16. Weinstein, J. N.; Myers, T. G.; O'Connor, P. M.; Friend, S. H.; Fornance, A. J. J.; Kohn, K. W.; Fojo, T.; Bates, S. E.; Rubinstein, L. V.; Anderson, N. L.; Buolamwini, J. K.; van Osdol, W. W.; Monks, A. P.; Scudiero, D. A.; Sausville, E. A.; Zaharevitz, D. W.; Bunow, B.; Viswanadhan, V. N.; Johnson, G. S.; Wittes, R. E.; Paull, K. D. *Science* **1997**, *275*, 343.
17. Lappano, R.; Rosano, C.; Santolla, M. F.; Pupo, M.; Ponassi, M.; Spallarossa, A.; Ranise, A.; Maggiolini, M. *Curr. Cancer Drug Targets* **2012**, *12*, 531.
18. Ranise, A.; Spallarossa, A.; Schenone, S.; Bruno, O.; Bondavalli, F.; Vargiu, L.; Marceddu, T.; Mura, M.; La Colla, P.; Pani, A. *J. Med. Chem.* **2003**, *46*, 768.
19. Zagotto, G.; Ricci, A.; Vasquez, E.; Sandoli, A.; Benedetti, S.; Palumbo, M.; Sissi, C. *Bioconjugate Chem.* **2011**, *22*, 2126.
20. Grever, M. R.; Schepartz, S. A.; Chabner, B. *Sem. Oncol.* **1992**, *19*, 622.