Università degli studi di Roma
"La Sapienza"

FACOLTA' DI SCIENZE MATEMATICHE,
FISICHE E NATURALI

Dottorato di Ricerca in Scienze Chimiche
- Curriculum "Chimica macromolecolare e biologica"

XXII CICLO

"Synthesis and study of cyclic indole-based oligomers as new G-quadruplex ligands and potential anticancer agents"

Supervisori:
Prof. Armandodoriano Bianco
Prof. Giancarlo Ortaggi

Dottorando:
Dott. Luca Ginnari Satriani

Correlatore:
Prof.ssa Maria Savino

INDEX

1. Introduction.....3

1.1. G-quadruplex structures: variations on a theme.....3

1.2. G-quadruplexes: mainly applications……10

1.3. Targeting telomeres and telomerase……..14

1.3.1 Telomeres: structure and functions……14

1.3.2 Telomerase……16

1.4. G-quadruplex ligands as telomerase inhibitors……19

2. The aim of the thesis…….. 22

3. Risultats and discussion……...26

3.1. Indole-based macrocycles……26

3.1.1. Preparation of the monomer without chains……..28

3.1.2. Preparation of the monomer with the hydrosoluble basic chain……...…28

3.2. Coupling of two monomers in solution……29

3.2.1. Amidic coupling by EDAC……..31

3.2.2. Amidic coupling between acyl fluorides and N-silylamines……32

3.3. Triazatruxene derivatives……..34

3.3.1. Synthesis of AZATRUX (24a)……35

3.3.1.1. Cyclotrimerization of N-substitued-2-indolones……35

Preparation of N-substitued-2-indolone (18)……...36

Cyclotrimerization step……...…38

3.3.1.2. Synthesis of the unsubstituted triazatruxene core……..39

Hexabrominated indole trimer: formation, isolation and subsequent dehalogenation……40

Cyclotrymerization of indole and subsequent dehalogenation without intermediate purification……42

Cyclotrimerization of unsubstituted 2-indolone in $POCl_3$……..……43

3.3.1.3. Insertion of the basic hydrosoluble chains in two steps……44

3.3.2. Studies of selective interactions of AZATRUX for G-quadruplex DNA with respect to duplex DNA by ESI-MS experiments…...……...46

3.3.3. Spectroscopic studies on AZATRUX……...51
3.3.3.1. NMR spectroscopy……...51
3.3.3.2. UV-vis absorption spectroscopy ……56
3.3.3.3. Fluorescence emission spectroscopy ……...61
3.3.4. Synthesis of a series of tris-N-substituted triazatruxene derivatives…….....65
3.3.5. FRET assays of the triazatruxene derivatives on quadruplex and duplex DNA..…68
3.3.6. Comparing studies for the triazatruxene derivatives towards telomeric G-quadruplex and duplex DNA by ESI-MS……77
3.3.7 Cell biology studies on AZATRUX……79

4. Conclusions……...83

5. Materials and methods.....88
5.2. Compounds preparations and characterizations……88
5.2.1 General……..88
5.2.2 Indole-based macrocycles……..88
5.2.3 Triazatruxene derivatives……...95
5.2. Biophysical and biological assays……100
5.2.1. Molecular modelling……100
5.2.2. ESI-mass experiments……..101
5.2.3. NMR DOSY experiment……..105
5.2.4 UV/vis absorption and fluorescence spectroscopy……106
5.2.5. Fluorescence Resonance Energy Transfer (FRET) assays……108

6. References.....108

1. Introduction.

1.1. G-quadruplex structures: variations on a theme.

G-quadruplexes are a family of nucleic acids secondary structures stabilized by G-quartets (or G-tetrads) (Neidle, **2009**), coplanar quartets of guanines held together by a cyclic arrangement of eight unconventional hydrogen bonds (fig.1). Although G-quartets were observed the first time, in 1910 in concentrated solution of guanylic acid that formed a clear gel in aqueous solution upon cooling (Bang, **1910**), the structure of G-tetrads was not determined until 1962, when Gellert and Davies found to consist of tetrameric guanine bases (Gellert *et al*. **1962**), forming a novel hydrogen-bonded structure (Hoogsteen bonds) (fig.1). These tetrads may stack on each other like base pairs in a double helix, creating three-dimensional structures known as G-quadruplexes (fig.1).
These structures, which can be formed by G-rich strands of DNA (or RNA) as the telomeric sequences, are stabilized by stacking interactions between overlapping G-tetrads; moreover the presence of monovalent cations, in particular K^+ (and to a lesser extent NH_4^+ and Na^+), which are placed between two overlapped adjacent G-tetrads and complex the eight carbonyls of the eight gunanines presents (Huppert **2008**) (fig.1). Interestingly, many structural studies (by X-ray crystallography and NMR) have shown a great conformational diversity in different G-quadruplexes (Yang *et al*. **2008**), such as folding topologies, loop conformations, and capping structures, indicating that these structural motifs may be differentially regulated and targeted.
A structure classification is possible on the basis of:

1. *number of associated strands*:
   tetramolecular, bimolecular and intramolecular structures can be formed by the corresponding association of four, two or one strands (fig. 2);

2. *directionality 5' → 3' of the associated strands*:
   parallel and antiparallel structures can be formed by the same or the opposite direction of the four associated strands (fig. 3);

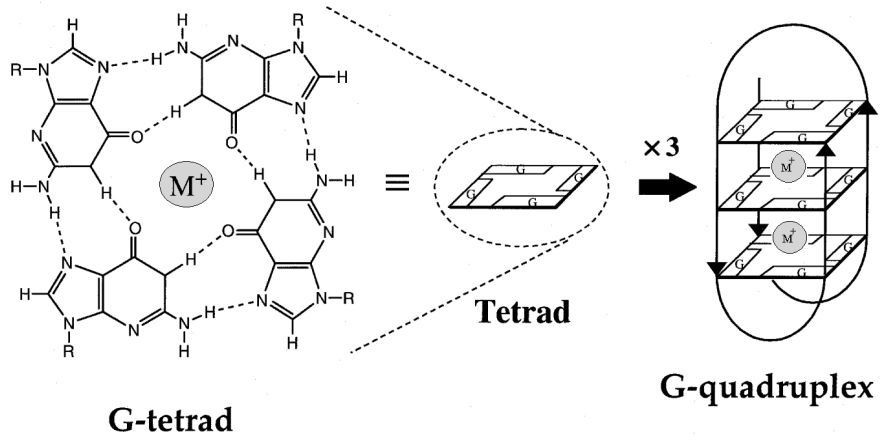


Fig.1 G-quadruplex formation.

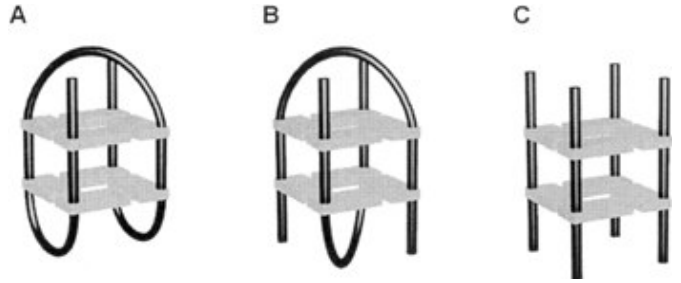


Fig.2 Various Strand Stoichiometries of G-Quadruplex Structures.
(A) A one-stranded structure yields a unimolecular G-quadruplex. (B) Two strands render a bimolecular G-quadruplex. (C) Four separate strands produce a quadrimolecular G-quadruplex (Simonsson **2001**).

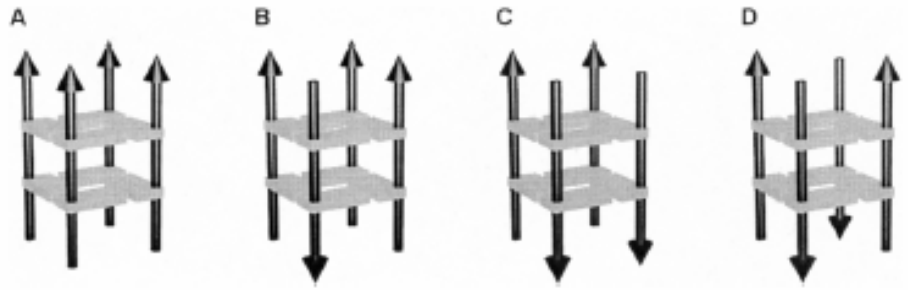


Fig.3 Different Strand Polarity Arrangements of G-Quadruplexes.
(A) All strands parallell. (B) Three parallell strands and one strand antiparallell. (C) Two pairs of adjacent parallell strands. (D) Alternating antiparallell strands. Arrows indicate 5'→3' polarity (Simonsson **2001**).

3. *conformation of guanosine glycosidic torsion angles (syn or anti)*:

   At a molecular level, the different directionality of the strands relates to the conformational state of the glycosidic bond between the guanine base and the sugar. This may be either syn or anti (fig.4). When all four strands are parallel, all the bases are in the *anti* conformation and the grooves between the backbones are all of equal size: the system is entirely C4 symmetric. When any of the strands are antiparallel, the bases must be in the *syn* form in order for the hydrogen bonds to be formed correctly. This then affects the orientation of the backbone relative to the G-quartets, and hence results in grooves of different sizes (fig.5). When successive guanines (starting with the guanine contributing N1 and N2) are both *anti* or both *syn*, the groove is medium in size (fig.5); if the first is *anti* and the second *syn*, the groove is wider, and if the first is *syn* and the second *anti*, then the groove is narrower (fig. 5).

4. *orientation of loops*:

   Depending on which strands are connected, loops may cross diagonally across the top of the structure, joining diagonally opposed antiparallel strands; go across a side, linking adjacent antiparallel strands; or may loop around the side of the structure linking parallel strands and forming a double-strand reversal loop.

These structural polymorphism of G-quadruplexes depends on a series of factors as:

- *oligonucleotides length:*

  The short Tetrahymena telomere oligonucleotide with sequence d(TG_4T) forms a tetramolecular parallel-stranded G-quadruplex structure, with the guanine glycosidic torsion angles in the *anti* conformation (Aboul-ela *et al.* **1992**, Phillips *et al.* **1997**) (fig. 6). Instead, the longer oligonucleotide sequence, like Oxytricha nova telomere d($G_4T_4G_4$), forms in solution a bimolecular G-quadruplex structure with parallel/antiparallel strand orientations and diagonal loops formed from opposite strands (Haider *et al.* **2002**) (fig. 6).

- *oligonucleotides sequence*:

  The two oligonucleotides d($G_4T_3G_4$) and d($G_4UT_2G_4$) (fig.6) with the same length but with a different sequence forms two different bimolecular quadruplexes. In fact, the sequence crystal structures contain two distinct quadruplexes with differing topologies, one with the

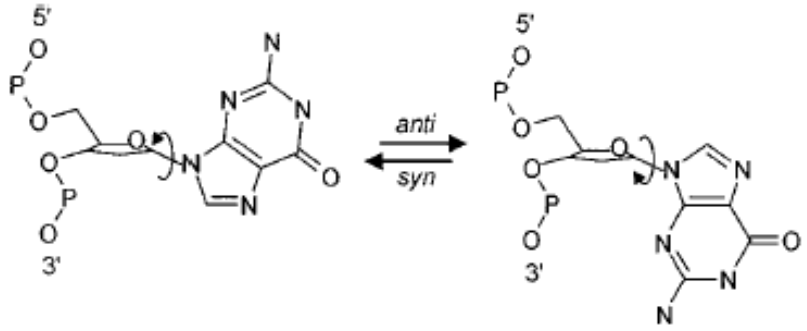


Fig.4 *Syn* and *anti* conformations of the glycosidic bond connecting the guanine base to the sugar-phosphate backbone.

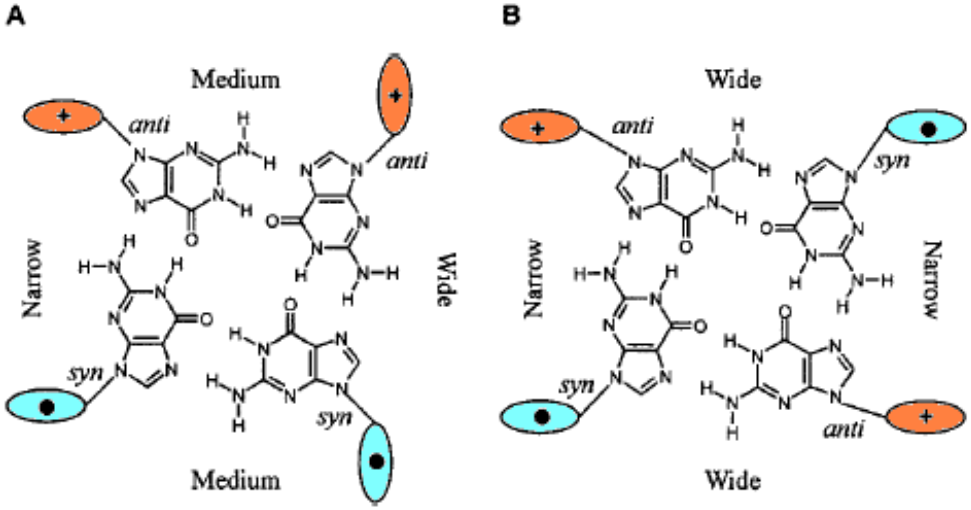


Fig.5 The interplay between glycosidic torsion angles and strand polarities, indicated by (+) or (•), Gives rise to grooves of different widths.

A. Guanine tetrads with two pairs of adjacent parallel strands that have identical glycosidic torsion angles generate two medium grooves, one wide groove, and one narrow groove.

B. Guanine tetrads with exclusively alternating antiparallel strands have two wide and two narrow grooves (**Simonsson**, 2001).

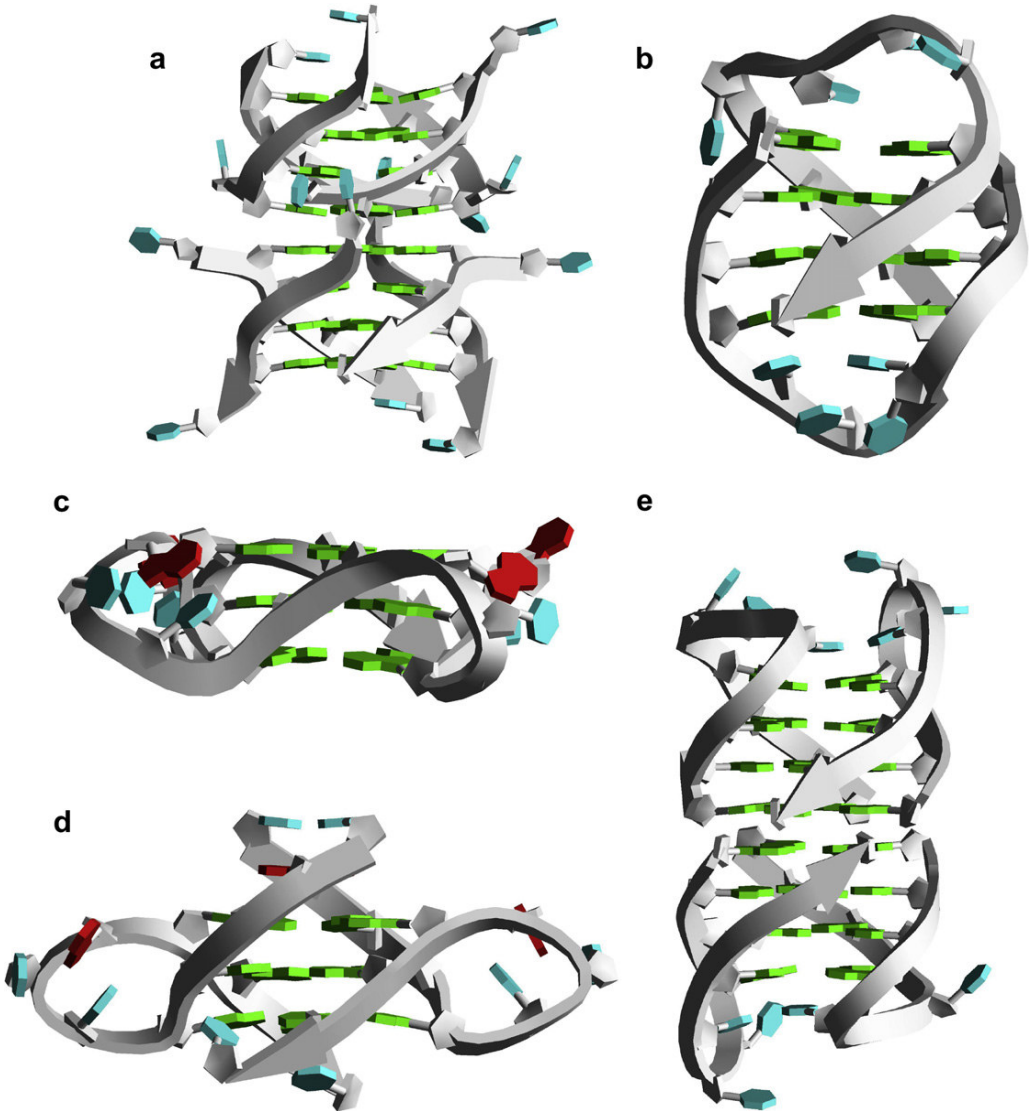


Fig.6 Views of various native quadruplex crystal structures.
The backbones are represented as ribbons, and bases are drawn in cartoon form. a) $d(TG_4T)$, b) $d(G_4T_4G_4)$, c) $d[AG_3(T_2AG_3)_3]$, d) $d(UAG_3UTAG_3T)$, e) lateral loop quadruplex in $d(G_4UT_2G_4)$ (Neidle and Parkinson **2008**).

diagonal loop as found in the Oxytricha nova structure ($d(G_4T_4G_4)$) , and the other having asymmetric lateral loops (Neidle and Parkinson, **2008**).

- *Type of cations stabilizing G-quadruplex:*

  The human telomeric sequence $d[AGGG(TTAGGG)_3]$ has been found to form different types of G-quadruplex structures (fig. 7). According to different cations-guanine coordination, in 1993 NMR studies (Wang & Patel **1993)** revealed that in Na^+ solution this 22 nucleotide (nt) sequence exhibits a monomeric antiparallel structure, with two lateral and one diagonal d(TTA) loop at the G-quartet ends (basket structure, fig. 7). Subsequently, structural information on the human telomeric G-quadruplex, formed under physiologically relevant conditions, was necessary for structure-based rational drug design. Therefore the K^+ structure has been the subject of intense investigation. In 2002 crystallographic studies in the presence of K^+ at a concentration approximately physiological ionic conditions showed a completely different monomeric parallel structure (Parkinson *et al.* **2002**) with three external d(TTA) loops (fig.8).

  Later, sedimentation and fluorescence studies in K^+ solution (Li *et al.* **2005**) suggest that the crystal structure (parallel structure) of the 22 nt sequence cannot be the predominant structure in K^+ solution but co-exists with an antiparallel form. The precise nature of any non parallel species formed by $d[AG_3(T_2AG_3)_3]$ in relatively dilute K^+ solution therefore remains to be determined, with evidence that a number of forms co-exist under these conditions.

  Recent NMR studies on the 26-nt human telomeric sequence $d[(TTAGGG)_4TT]$ (a four-G-tract native human telomeric sequence) in K^+ solution showed the presence of two conformations: the hybrid-1 and hybrid-2 human telomeric G-quadruplexes (Ambrus *et al.* **2006**) (fig. 19). The hybrid-2 type structure appears to be the major conformation on the extended human telomeric sequences (Dai *et al.* **2007**). Both hybrid-type structures contain three G-tetrads linked with mixed parallel/antiparallel G-strands, but they differ in loop arrangements, strand orientations, G-tetrad arrangements and capping structures (Fig. 19). On these bases, it has been proposed a model of DNA secondary structure composed of compact-stacking multimers of hybrid-type G-quadruplexes in human telomeres, with an equilibrium between hybrid-1 and hybrid-2 forms in K^+ solution (Dai *et al.* **2008**) (fig. 19).

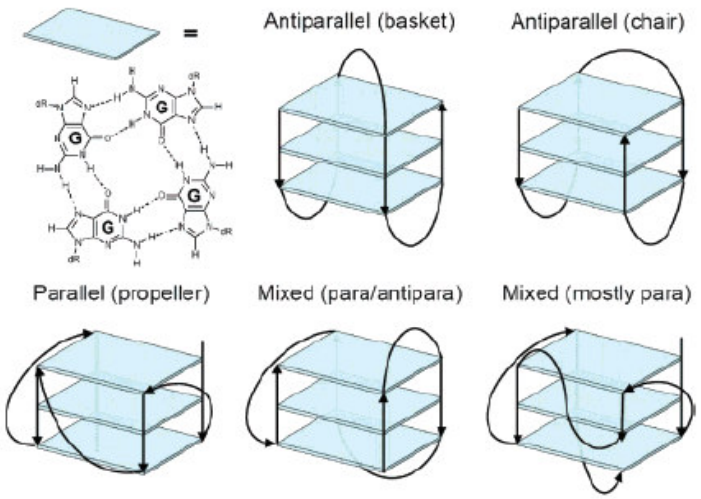


Fig.7 Several monomeric structures (Gabelica *et al.* **2006**).

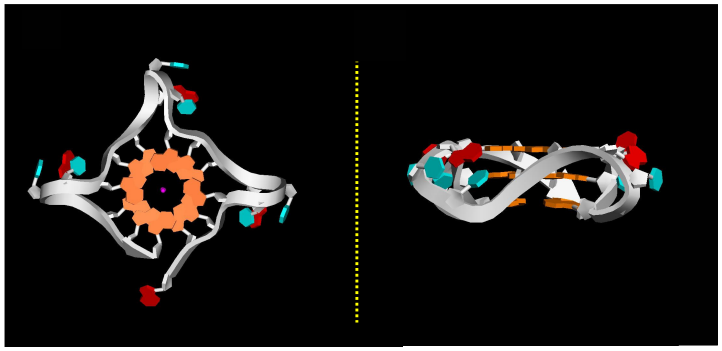


Fig. 8 Crystal monomeric structure from 22-nucleotide "F21T"
Sequence: d[AGGG(TTAGGG)$_3$] (Parkinson *et al.* **2002**).

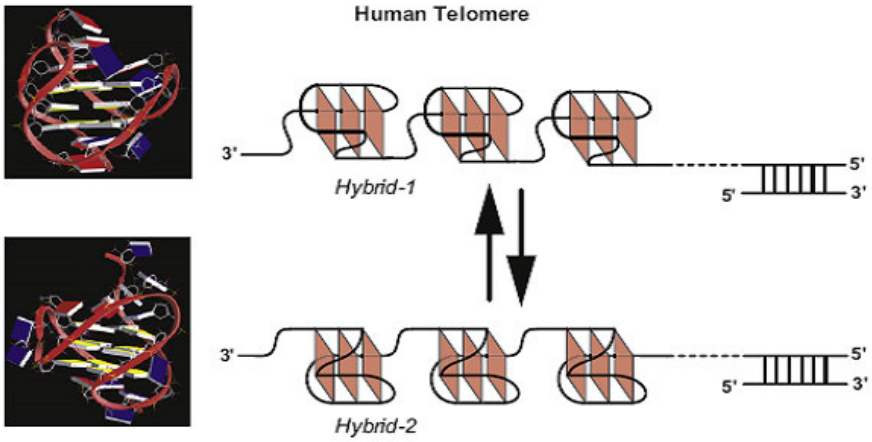


Fig. 9 A model of DNA secondary structure composed of compact-stacking multimers of hybrid-type G-quadruplexes in human telomeres, with an equilibrium between hybrid-1 and hybrid-2 forms in K^+ solution (Dai *et al.* **2008**).

1.2. G-quadruplexes: main applications

Guanine-rich segments able to form G-quadruplexes are found in biologically important regions such as telomeres (Blackburn, **1991**), oncogene promoter regions (Simonsson *et al.* **1998**) and immunoglobulin switch regions (Sen and Gilbert, **1988**). Not only because their biological importance but also because the potential applications in supramolecular chemistry and nanotechnology, G-quadruplex attracted much attention in the field of medicinal chemistry, pharmaceutical biology, and material sciences (Franceschin, **2009**).

Telomeres:

Telomeric G-rich single-strand DNA formed as 3' overhang of chromosomes is the substrate of telomerase, an enzyme necessary for telomeric replication, which is over-expressed in most cancer cells and participates in tumors genesis (Blackburn, **1991**). The formation of a telomeric G-quadruplex blocks telomerase activity and offers an original strategy for new anti-cancer agents. In the last years, several families of compounds have been identified which specifically bind to the quadruplex structures. These derivatives, called "G-quadruplex DNA ligands", are able to block telomeric replication in cancer cells and to cause replicative senescence and/or apoptosis after a few cell cycles (Chen *et al.* **1996**) (see below).

Transcription regulation :

An area of increasing interest is the role of G-quadruplexes in transcription regulation. A number of individual sequences in gene promoters have been shown to exist in vitro and affect transcriptional rates in vivo, including sequences in the chicken b-globin gene (Howell *et al.* **1996**), and the human c-myc gene (Siddiqui-Jain *et al.* 2002) , which has been particularly well characterized (Phan *et al.* **2004**; Yang *et al.* **2006**). This oncogene is an important transcription factor, which regulates a wide variety of genes, especially those involved with proliferation, differentiation and apoptosis. It has a number of nuclease hypersensitive regions (regions of DNA that are especially susceptible to enzymatic cleavage) in its promoter, one of which is responsible for a large majority of the transcriptional activity, and contains the sequence d(GGG GAG GGT GGG GAG GGT GGG GAA GG), which forms a family of stable G-quadruplex sequences. Stabilization of these structures using the porphyrin-derived G-quadruplex ligand TMPyP4 resulted in down regulation of the c-myc oncogene in vivo (Grand *et al.* **2002**). In the last years, G-quadruplex-forming motifs have been found in several promoter regions of a number of human genes (in regions immediately upstream of

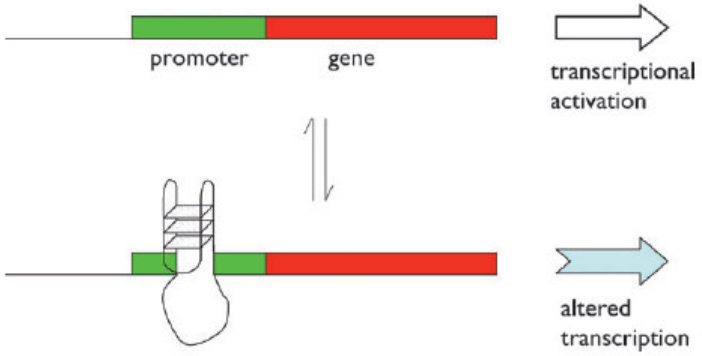


Fig. 10 The formation of a G-quadruplex in a promoter can affect the level and nature of transcription from that gene (Huppert **2008**).

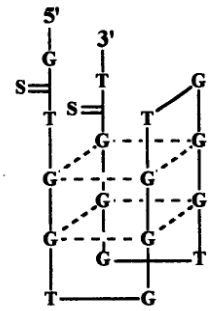


Fig. 11 Structure of Zintevir (Urata *et al*. **2001**).

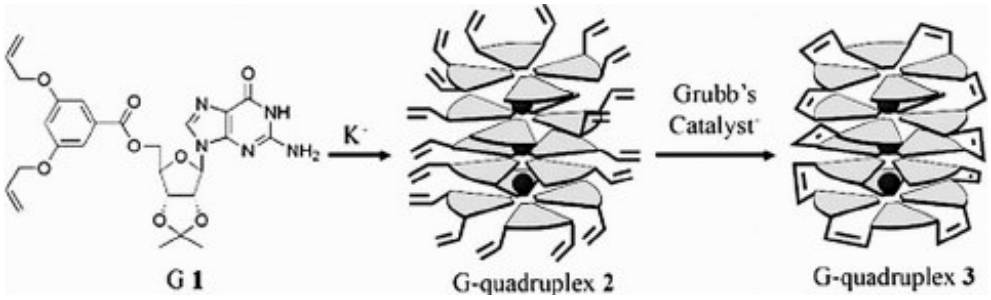


Fig. 12 Two-step synthesis of a lipophilic G-quadruplex by olefin metathesis (Kaucher et al. **2006**).

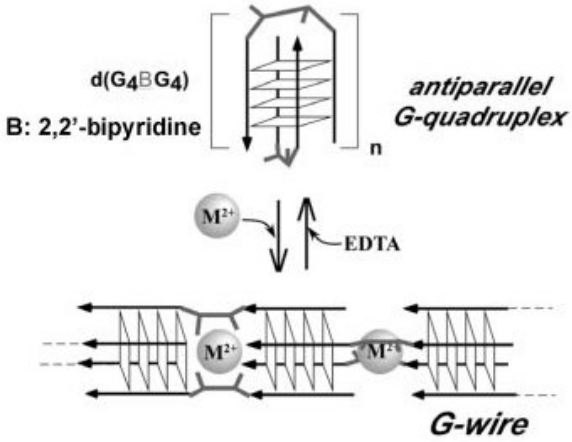


Fig. 13 G-wire formation controlled by divalent cations (Karimata *et al.* **2007**).

the transcription initiation site), including the important protooncogenes: VEGF (Sun *et al*. **2005**), HIF-1a (Armond *et al.* **2005**), BCL2 (Dexheimer *et al.* **2006**), PDGF-A (Qin *et al*. **2007**), c-kit (Rankin *et al.* **2005**) and c-Myb (Lee Palumbo *et al.* **2008**). This, suggests a significant role of G-quadruplex structures for the transcription regulation of a number of protooncogenes. One model for the effect of these sequences consists of an equilibrium between a folded and unfolded form of the Gquadruplex motif, resulting in transcriptional change (fig.10). One mechanism by which this could happen involves the G-quadruplex acting as a steric block to the transcription machinery. Quadruplex-interacting proteins may potentially mediate the equilibrium exchange (Shirude *et al*., **2007**). It should be noted that the formation of G-quadruplexes in promoters might result in either an increase or a decrease in expression (fig.10).

Studies on structural information of several genes G-quadruplex structures are still in progress in order to design selective ligands for new specific anticancer approaches.

RNA quadruplexes:

Although DNA quadruplexes have been studied in great detail over the past 15 years, only little is known about the stability and folding of corresponding G-rich RNA sequences. Whilst in DNA, G-quadruplex formation must overcome the challenge of strand separation and chromatin opening, these constraints do not apply in the case of single-stranded RNA sequences, where G-quadruplex formation is much more likely to occur. Additionally, it has been shown experimentally that RNA G-quadruplexes tend to be more stable than the equivalent DNA structures (Mergny *et al.*, **2005**).

Recent studies highlight the importance of RNA quadruplexes in cellular processes. For example, quadruplexes forming in mRNAs have, been found to interfere with gene expression in both pro- and eukaryotes (Aurora *et al.*, **2008**). Recently, bioinformatics studies have provided evidence that putative RNA quadruplexes are more frequently found at certain positions in untranslated regions of mRNAs (Huppert *et al.*, **2008**). In addition to mRNA-based quadruplexes, telomeres have been found to be actively transcribed by polymerase II, generating RNAs with the potential to form RNA G-quadruplexes that could be also involved in regulatory processes at the telomeres (Azzalin *et al.*, **2007**).

Aptamers:

Aptamers are short DNA- or RNA-based oligonucleotides able to efficiently recognize targets ranging from small molecules to proteins or nucleic acid structures. Like antibodies, they exhibit high specificity and affinity for target binding. As a result, they may display effective interference in biological processes, which renders them not only valuable diagnostic tools, but also promising

therapeutic agents. In particular, G-quadruplex-forming aptamers have shown promising biological and possibly pharmacological properties ranging from anticancer to anti-HIV activities.

A bioactive quadruplex-forming oligonucleotide, a 26-mer nucleotide named "AS1411" (Antisoma, London, UK), has been shown specifically to inhibit cancer cell proliferation: the activity of this aptamer has been related to its binding to certain nucleolin containing complexes (Teng *et al.* **2007**). Another family of guanine-rich oligonucleotides has been developed as potential anti-HIV therapeutic drugs (Held *et al.* **2006**). These compounds have demonstrated strong interaction with HIV-1 integrase in vitro and the ability to inhibit the integration of viral DNA into host DNA (Jing *et al.*, **2000**). Among several different G-quartet oligonucleotides proposed and studied for their ability to inhibit HIV, the most efficient was the 17-mer oligonucleotide 5'-GTGGTGGGTGGGTGGGT-3', referred to as Zintevir (Urata *et al*. **2002**) (fig. 11).

Recently, G-rich RNA aptamers against bovine prion protein (bPrP) were obtained in order to develop aptamers as therapeutic agents against prion diseases and Alzheimer's disease (Mashima *et al.* **2009**).

Supramolecular chemistry:

G-quadruplexes have attracted interest in the areas of supramolecular chemistry and nanotechnology, because of their potential applications for developing ion channels, wires, and molecular switches. Such self-organization of non-covalent assemblies of G-quartets, for instance, leads to liquid crystals in water: even without any covalent bridge, G-quartets stack to give columnar structures with hydrophilic surfaces and lipophilic cores. Depending on the concentration and other physico-chemical parameters, these aggregates generate arrangements in the cholesteric phase and in the hexagonal liquid-crystalline phase (Davis *et a.l* **2007**).

A remarkable ionophore was reported by J. T. Davis and colleagues, who covalently attached sixteen guanosine units together using olefin metathesis (fig. 12). This compound was synthesized using a templated synthesis in which multiples of four guanosines would self associate around K^+ cations to form "G-quadruplex" units. Under reaction conditions, four of these G-quadruplexes would stack, after which the polymerization would occur to give a templated hexadecamer of guanosine. Upon insertion into a lipid bilayer, the compound formed a G-quadruplex sodium transporter of similar length to gramicidin A (Kaucher *et al*. **2006**). Another intriguing potential application of quadruplex based nanodevices is represented by the so called G-wires, higher-order and parallel-stranded G-quadruplexes (fig.13). The G-wire is an especially useful structure as a nanomaterials because the G-wire has a well controlled and regulated alignment of numerous DNA strands with G-quartets (Karimata *et al.* **2007**).

1.3. Targeting telomeres and telomerase.

1.3.1 Telomeres

Telomeres are specialized DNA protein structures that cap the ends of linear chromosomes (fig.14). Telomeric DNA consists of tandem repeats of small G-rich sequences highly conserved during evolution. Human telomeres at birth contain 15–20-kilobase pairs of the repetitive sequence TTAGGG followed by a 3′ single-strand overhang on the G-rich strand (Blackburn, **2000**). Several electron microscopic studies have suggested that the structure of the ends of telomeres may be more complex than originally thought. Griffith and colleagues have found that telomeres do not end in a linear manner but form loop structures (Griffith *et al.* **1999**) with the 3'-G-rich strand, called T-loop (telomere loop) intercalating the duplex telomeric repeats of the 5'-end to form the D-loop (displacement loop). This conformation is maintained by the shelterin (de Lange **2005**), a complex of six proteins (TRF1, TRF2, TIN2, Rap1, TPP1, and POT1) whos bind to the single- and double-stranded telomeric DNA capping the chromosome ends (fig. 15). Recently, it has been proposed (Songyang & Liu **2006**) that the shelterin multi-complex and its associated partners provide the basis for constructing an interaction map of telomere regulatorsin mammalian cell, named "Telomere Interactome", which incorporates the various telomere signalling pathways and represents the molecular machinery that regulates mammalian telomeres. The telomeres has manifolds functions, reported as follows:

- **protects chromosomes from degradation and degenerative processes:** Usually, non-telomeric double-strand DNA breaks are rapidly repaired. The Shelterin complex, in "capped conformation" enables cells to distinguish their natural chromosome ends from DNA breaks, represses DNA repair reactions, prevents ligation with other chromosomes and regulates telomerase-based telomere maintenance (de Lange **2005**) (fig.15).

- **protects the genoma from loss of essential genes:**
  Dividing cells show a progressive loss of 25–200 DNA bp in each cell division. This loss of DNA is largely due to the "end replication problem" (Harley *et al.* **1990**) which refers to the inability of the DNA replication machinery to copy the final few base pairs of the lagging

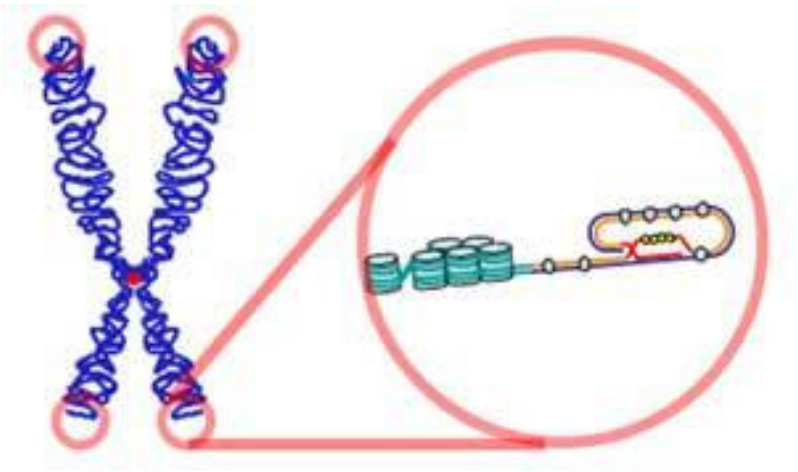


Fig.14
The telomere:
the end of
the chromosome

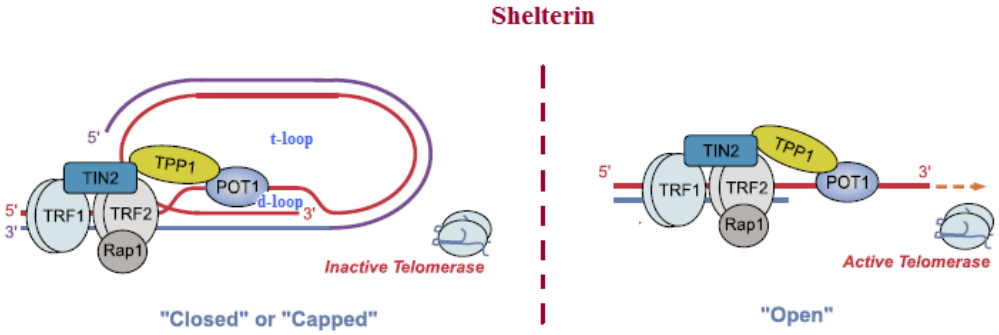


Fig. 15 The shelterin/telosome complex:
the capped state on the left and the open state on the right (de Cian *et al.* **2007**)

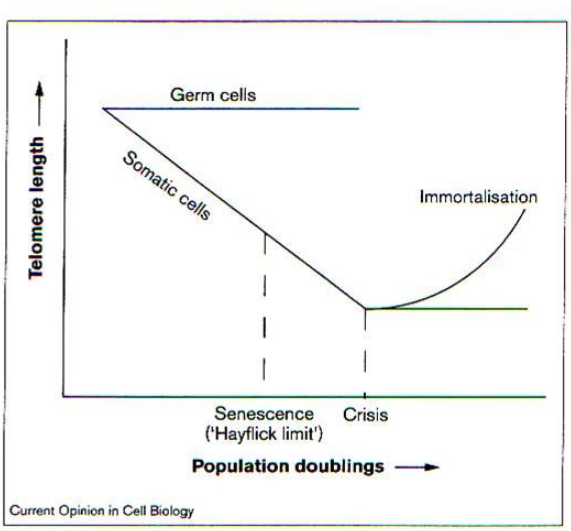


Fig.16 Telomere lenght vs.
population doublings:
the Hayflick limit and the replicative
senescence.

strand during DNA synthesis. Because the telomere consists of a repetitive no-encoding DNA, they protects the cell from the loss of more critical gene encoding sequences (Rezler *et al.* **2003**).

- **biological clocks for the cell proliferation:**

  As a result of the loss of telomeric DNA due to the "end replication problem", telomeres progressively shorten in all somatic cells with each cell division. When telomeres achieve a critic length (Hayflick limit) (fig. 16), a cellular signals induces cells usually undergo replicative senescence (mortality stage 1, M1) (Wright *et al.* **1989**). In addition to progressive telomere shortening (leading to replicative senescence), telomere dysfunction can be caused by a change of state (uncapping) that leads to a rapid induction of growth arrest (mortality stage , M2). However, in some rare cases, cells can escape this mortality stage leading to cell immortalization and cancer cell proliferation (Shay *et al.* **2006**) (fig. 17). These cells maintain telomere length because they express an enzyme that assures the addiction of telomeric repeats to chromosomal ends. This enzyme is telomerase.

1.3.2 Telomerase

Telomerase is a cellular reverse transcriptase responsible for telomere maintenance in most organisms (fig. 18). It does so by adding telomere repeats at the 3' extremity onto pre-existing ends using an integral RNA component as template (Lingner *et al*. **1998**).

Human telomerase is a ribonucleoprotein complex containing two major components (fig. 19) as well as many associated proteins that are critical for function. The first major component is a reverse transcriptase subunit (hTERT), a 120 kDa protein active as a dimer, which interacts with telomere DNA and telomere-binding protein, catalyzing telomeric DNA reverse transcription and telomere end capping (Blasco *et al.* **2005**). The second component (hTR) is an RNA subunit containing an 11 bp sequence which acts as template for the reverse transcription of TTAGGG repetitive sequence (Feng *et al.* **1995**). Recently, was resolved by X-ray crystallography the high-resolution structure of the Tribolium castaneum catalytic subunit of telomerase, TERT. The protein consists of three highly conserved domains, organized into a ring-like structure that shares common features with retroviral reverse transcriptases, viral RNA polymerases and B-family DNA polymerases. Domain organization places motifs implicated in substrate binding and catalysis in the

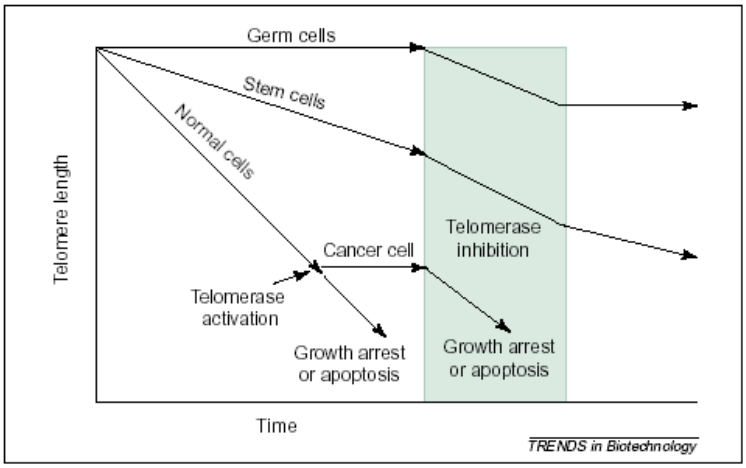


Fig. 17 Effects of telomerase inhibition on the telomere lenghts (White *et al.* **2001**).

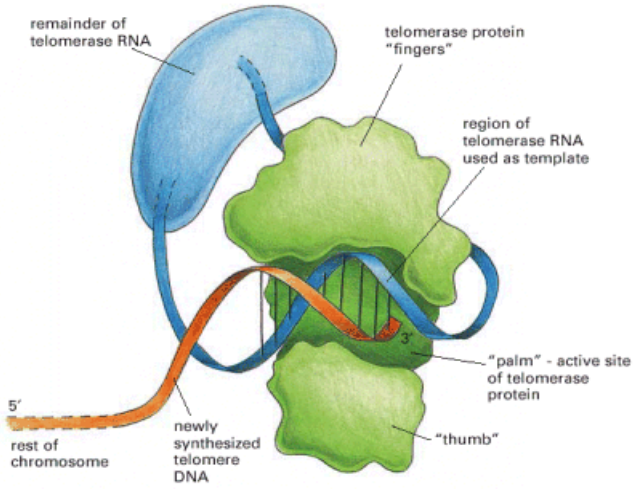
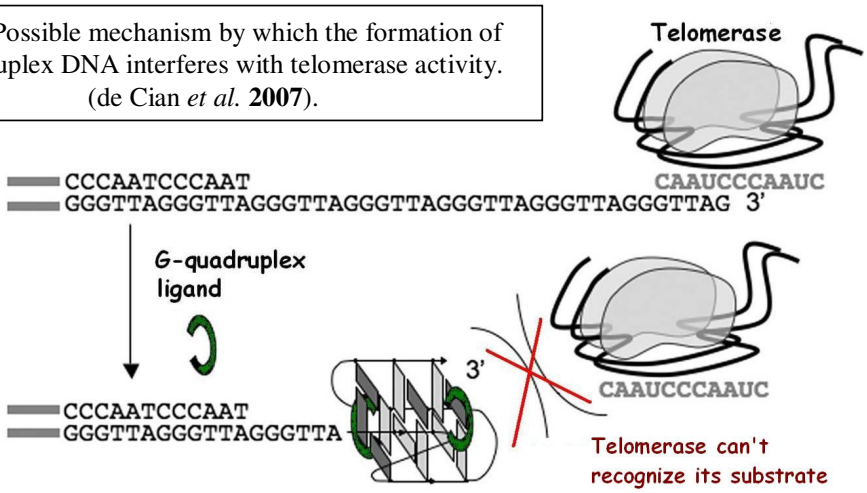


Fig. 18
The telomerase.

Fig. 19 Possible mechanism by which the formation of G-quadruplex DNA interferes with telomerase activity. (de Cian *et al.* **2007**).



interior of the ring, which can accommodate seven to eight bases of double-stranded nucleic acid (Gillis *et al*. **2008**).

Telomerase activity is physiologically required to regulate the structures of telomeres for continuous cell division during embryonic development and stem cell renewal and proliferation (Cesare & Griffith **2004**). There is no trace of telomerase activity in normal somatic cells, even in germline and stem cells this activity is low or absent (Shay & Wright **2006**, Hiyama & Hiyama **2007**) at contrary it has been found that in the majority of tumour cells (85-90%) this enzyme is overexpressed (Kim *et al.* 1**994**). In particular, introduction of hTERT into telomerase-silent cells is sufficient to reactivate telomerase, elongate or maintain telomeres, and to result in the bypass of both M1 and mortality stage 2 (M2). Moreover, some key-experiments have also been shown as inhibition of telomerase limits the growth of human cancer cells (Hahn *et al.* **1999**). These features make telomeres and telomerase an extremely attractive target for cancer therapy. Several approaches to target telomeres and telomerase activity have been extensively described (Hodes **2001**, White *et al.* **2001**, Shay & Wright **2001**, Saretzky **2003**, de Cian *et al.* **2007**, Pagano & Giancola **2007**).

One therapeutic strategy is based on immunotherapy, which uses vaccines to support the immune system in attacking cancer cells. The hTERT component of telomerase is the most attractive target for these vaccines, because it is expressed in cancer cell, but not in normal cells (Ulaner **2004**).

Another strategy is based on targeting the hTR component of telomerase with oligonucleotide inhibitors. These oligonucleotides should be considered as “template agonists” rather true antisense agents, as their targets are reverse transcribed rather than translated into peptides. Different strategies and chemical modifications have successfully been developed to target hTR, starting from traditional DNA-based antisense oligomers: peptide nucleic acids (PNAs) were the first oligomers tested that efficiently inhibit telomerase (Norton *et al.* **1996**).

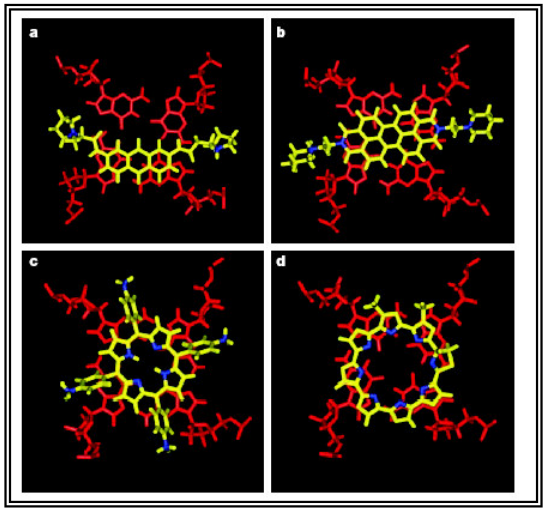
An alternative to targeting telomerase is to target telomeres themselves (fig.19). Optimal telomerase activity requires an unfolded single stranded substrat, therefore G-quadruplex structures formation by the telomere directly inhibits telomerase elongation in vitro (Zahler *et al.* **1991**). Infact, as described above, the 3′ overhang of telomeres are rich in guanine units and are able to form G-quadruplexe structures. Therefore, ligands that selectively bind to and stabilize G-quadruplex structures (see below) may interfere with telomere conformation and telomere elongation (de Cian *et al.* **2008**) (fig.16).

1.4. G-quadruplex ligands as telomerase inhibitors.

Studies *in vitro* on telomeric G-quadruplex structures in relevant physiologic condition has been fundamental in order to design and develop ligands especially to target the human telomeric repeat and hence block the action of telomerase. The number of known G-quadruplex ligands has grown rapidly over the past few years. These molecules recognize quadruplex DNA and are active in the telomere repeat amplification protocol (TRAP) assay (Gomez *et al*. **2002**).

Features shared by many of these ligands include a large flat aromatic surface, and, in most cases, basic side chains (positively charged under physiological conditions). NMR, Crystallographic and molecular modelling studies suggest, infact, a ligand/G-quadruplex general model in which ligands interact mainly through terminal π-π stacking at the end of the quadruplex (threading intercalation) (fig. 3), without internal intercalation, and by electrostatic interactions of the side chains in the grooves (Read and Neidle **2000**) (fig. 20, 21). Exceptions of this model are represented by some molecules, as distamycin A (Randazzo *et al.* **2002**, Pagano *et al.* **2007**), capable of stabilizing G-quadruplexes showing a binding mode based on loop and groove interactions (Müller *et al.* **2009**). Many molecules with the characteristics described above have been synthesized as potential telomerase inhibitors: fluorenones (Perry *et al.* **1999**), anthraquinones (Sun *et al.* **1997**) and 3,6-disubstituted acridines (Harrison *et al.* **1999**) are examples of the first generation of these inhibitors (fig. 22). Subsequently, trisubstituted acridines (Read *et al.* **2001**), porphyrins (Seenisamy *et al.* **2005**), and triazines (Gomez *et al.* **2003**) were synthesised to improve efficiency in telomerase inhibition (Figure 4). In fact passing from disubstituted compounds, as 3,6-disubstituted acridines, to trisubstituted compounds, as 3,6,9-trisubstituted acridine, the affinity and the selectivity of ligands toward G-quadruplex structures has been improved. This can be explained in term of G-quadruplex affinity as a consequence of a one more electrostatic interaction with a third groove of the G-quadruplex and in term of selectivity as consequence of major steric hindrance in the duplex binding (Read *et al.* **2001**).

One of the most promising compounds of this series is represented by BRACO19 (fig.21), a 3,6,9-trisubstituted acridine, synthesized and studied by Neidle's group (Gowan *et al.* **2002**). In fact, this compound has been shown to inhibit tumour growth in vitro and in vivo at concentrations under 1 μ. Another important class of G-quadruplex ligands are perylene derivatives as like PIPER (Fedoroff *et al.* **1998**), the first derivatives of this class reported and most intensely studied. Our research group has recently reported a series of new highly hydrosoluble perylene [such as DAPER4C(1,6)](Franceschin *et al.* **2007**) and coronene (CORON) (Franceschin *et al.* **2007**) derivatives with three or four side chains bearing various basic substituents. All of them exhibit



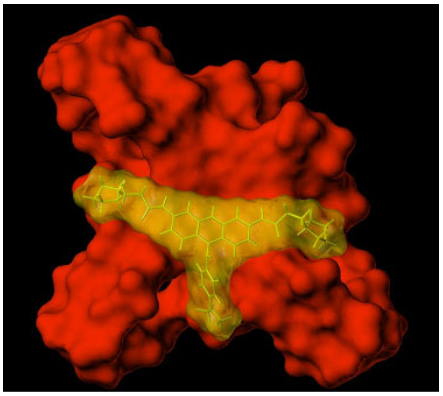


Fig.21 Simulated interaction of BRACO 19 with the G-quadruplex DNA (Gowan *et al.* **2002).**

Fig. 20 Simulation of the ligand/DNA G-quadruplex interaction by means of *molecular modelling* studies. Superimposition of the ligands on the terminal G-tetrad (Neidle *et al*. **2002**):
a. BSU 6039; **b.** PIPER; **c.** TMPyP4; **d.** Telomestatin.

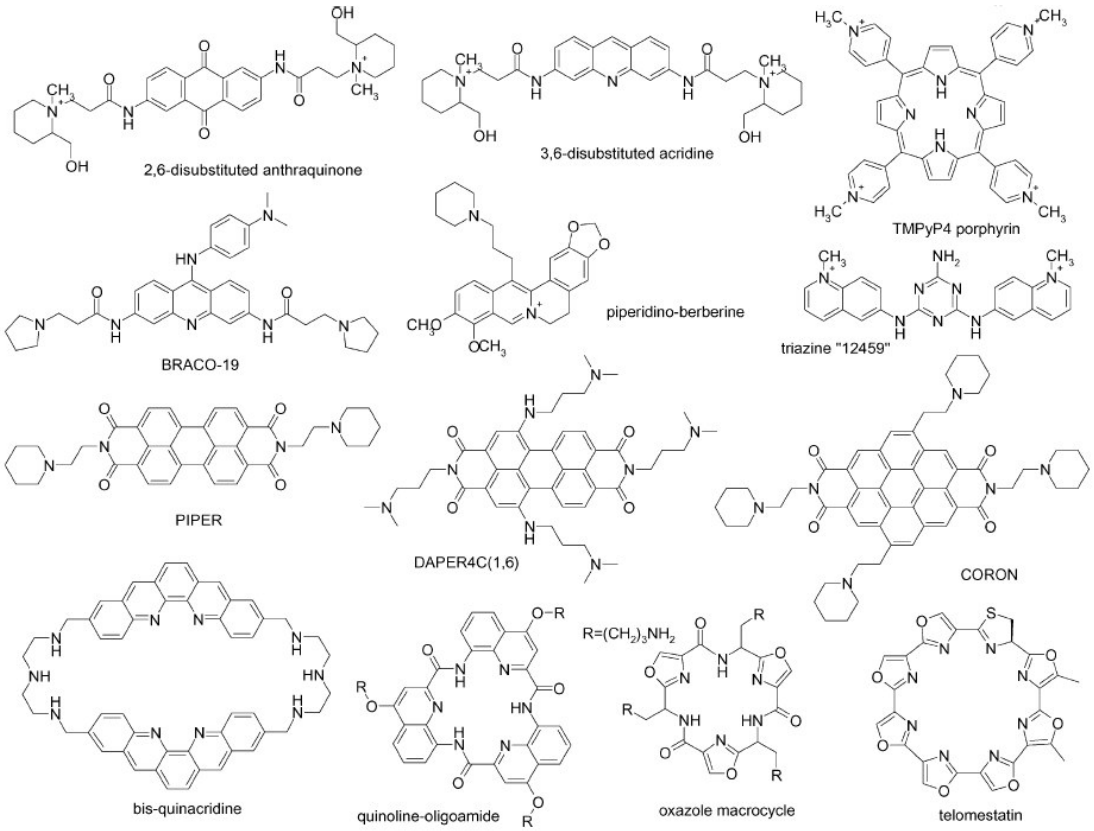


Fig.22 Structures of known G-quadruplex ligands (Franceschin **2009**) .

enhanced ability to interact with G-quadruplex structures and to inhibit human telomerase, with respect to the previously reported perylene derivatives.

Another promising class of G-quadruplex interactive compound is represented by porphyrin derivatives (Shi *et al.* **2001**). The most famous of them is TMPyP4 (fig. 20; fig. 22), that shows an affinity for G-quadruplex structures two times higher than for duplex DNA. Furthermore, this compound is able to promote the formation of the G-quadruplex in the promoter region of c-*myc*, so that it can inhibit telomerase also by the downregulation of the expression of hTERT (Siddiqui-Jain *et al.* **2002**). Another interesting feature of this porphirine is that this molecule is able to promote the formation of the i-motif DNA structure in the cytosine-rich telomeric strand, complementary to the guanine-rich sequence (Fedoroff *et al.* **2000**). For these reasons it has been supposed the intriguing possibility that TMPyP4 can trigger the formation of unusual DNA structures in both strands of the telomeres, which may in turn explain the biological effects of TMPyP4 in cancer cells (Grand *et al.* **2002**). More natural compounds such as berberine (Naasani *et al.* **2009**), quercetin (Sun *et al.* **2006**), ascididemin and meridine (Guittat *et al.* **2005**), cryptolepine and telomestatin (Figure 4) show a good affinity toward G-quadruplexes. The most famous and powerful compound discovered is telomestatin (Kim *et al.* **2002**) (fig. 22). The latter was isolated from *Streptomyces anulatus* 3533-SV4, after a wide range of specific screening for telomerase inhibitors. Indeed its name is due to its potent inhibitory activity against human telomerase at 5 nM concentration (Shin-Ya *et al.* **2001**) . After its structural characterization, showing the macrocyclic linkage of two methyloxazoles, five oxazoles and one thiazoline ring, the total synthesis of telomestatin was reported, and its absolute configuration was determined to be (*R*) (Doi *et al.* **2006**). Telomestatin has been shown to be able to reduce cell growth both in vitro and in vivo in a number of human tumours, such as multiple myeloma (Shammas *et al.* **2004**), neuroblastomas (Binz *et al.* **2005**) and leukaemia (Tauchi *et al.* **2006**). Following the promising results obtained with the naturally occurring telomestatin (Kim *et al.* **2002**), several synthetic macrocycles, based on quinacridine (Teulade-Fichou *et al.* **2003**), oxazole (Jantos *et al.* **2006**) and quinoline (Shirude *et al.* **2007**) moieties (Figure 4), were proposed and studied as G-quadruplex ligands. The last two classes of compounds are trimers characterized by amide bonds and obtained by classical peptide chemistry with protecting groups such as BOC. More recently, an analogous strategy has been used to synthesize macrocyclic hexaoxazoles (Tera *et al.* **2008**), which bear a closer structural resemblance to telomestatin than the previously reported macrocycles (see below).

2. The aim of the thesis

My PhD work deals with the design and synthesis of new G-quadruplex interactive compounds and the evaluation of their efficiency in inducing and stabilizing G-quadruplex structures. In particular, my efforts have been directed towards two synthetic targets represented by different indole-based oligomers: tetrameric macrocycles of indoles, where the monomers are linked by amidic bonds, and the triazatruxene core, a cyclotrimer of indoles were the three monomers are fused in a new aromatic core.

As a consequence of the high efficiency of telomestatin (fig. 24), a natural macrocycle able to strongly bind the macromolecular target even without lateral chains, to induce and stabilize the G-quadruplex DNA, in the last years several macrocyles were synthesized as G-quadruplex ligands. Particularly, oxazoles-based (Jantos *et al.* **2006**) and quinoline-based polyamido-macrocycles (Shirude *et al.* **2007**) are trimers characterized by amide bonds and obtained by classical peptide chemistry with protecting groups such as BOC (fig. 23). More recently, an analogous strategy has been used to synthesize a macrocyclic hexaoxazoles (Tera *et al.* **2008**) and a macrocyclic heptaoxazoles (Tera *et al.* **2009**) which bear a closer structural resemblance to telomestatin than the previously reported macrocycles (fig.24). These compounds combine a planar macrocyclic hexaoxazole pharmacophore with basic side chains: they showed good specificities for G-quadruplexes and potent telomerase inhibitory activities in both cell-free and cell-based assay systems.

In this context, our interest has been devoted to the design of new polyamido-macrocycles, considering the structural properties of the G-tetrad. In particular, I have thought to build a similar structure to optimize the stacking interactions between the ligand and the terminal G-tetrad. So, I have designed an indoles tetrad core, a macrocycle which can be considered as a simplified model of the G-tetrad, where guanines are replaced by indoles and the Hoogsteen H-bonds by four amidic bonds (fig. 25).

Molecular modelling studies have shown a good superimposing of the designed ligand with the terminal G-tetrad of the G-quadruplex (see below). In parallel, I have designed the indoles tetrad equipped of four hydrophilic chains to improve the interactions between the ligand and the G-quadruplex, according to the model of interaction previously described (par. 1.4.) (fig. 25). Therefore, I have designed and synthesized two suitable starting monomers **5** and **11** (fig. 26) for a synthetic strategy inspired by peptides synthesis in order to obtain the two desired macrocycles (fig.25). In both cases we have chosen monomers with the carboxylic and aminic functions suitable arranged in order to obtain the correct geometry of the final macrocycle (fig. 25): with the

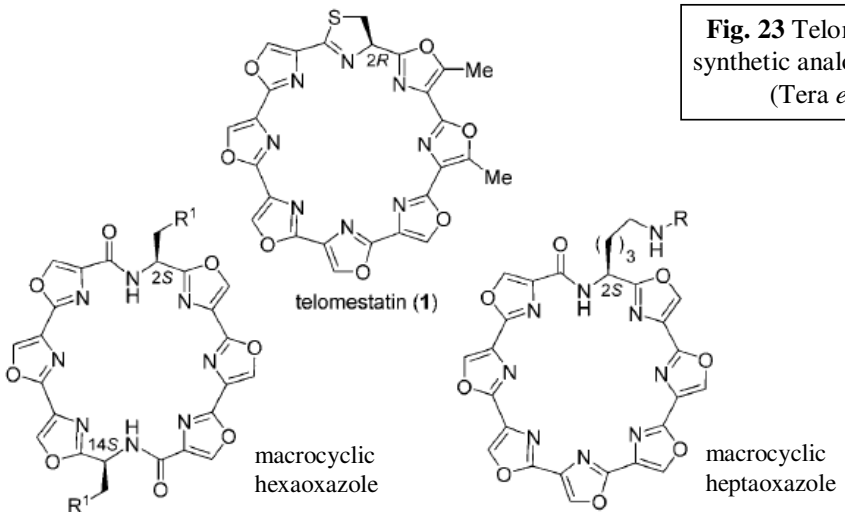


Fig. 23 Telomestatine and its synthetic analogue macrocycles. (Tera *et al.* **2009**)

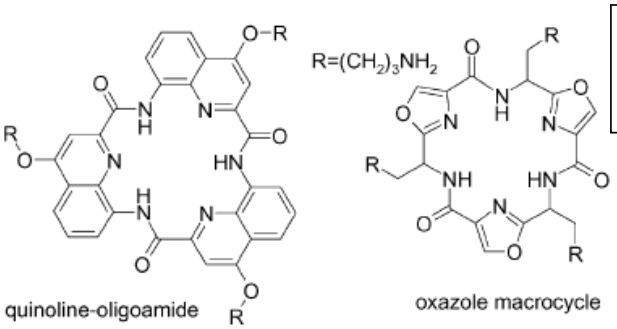


Fig.24 Polyamido-macrocyclic G-quadruplex ligands. (Jantos *et al.* **2006**; Shirude *et al.* **2007**)

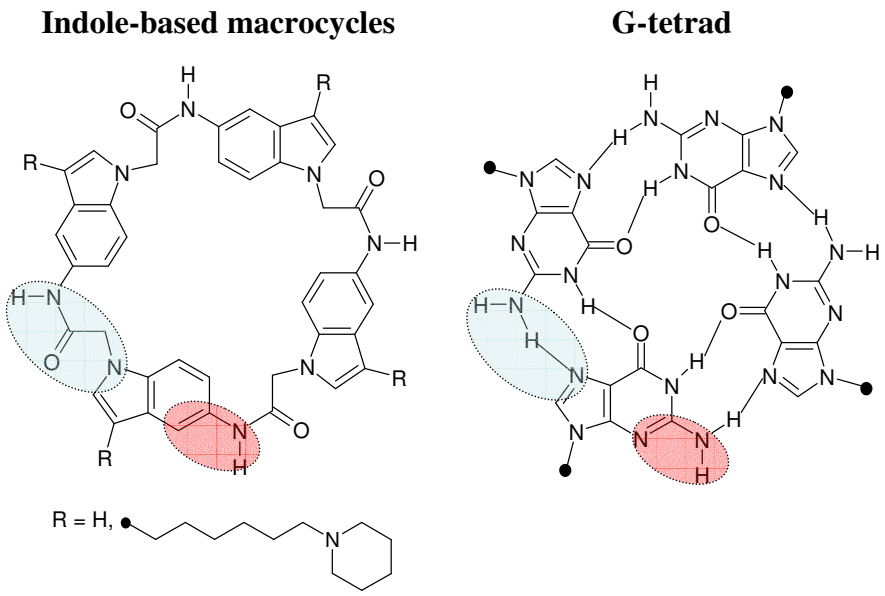


Fig.25 Indole-based macrocycles vs. G-tetrad structure.

methilenic bridge bearing the carboxylic group in 1-indole position and the amino group in 5-indole position. In both cases we detected the starting compound as the commercial product 5-nitroindole (**1**) (see below); while the difference of the two monomers is the insertion of the hydrophilic chain in 3-indole position in the case of the side-chained monomer (Fig. 26).

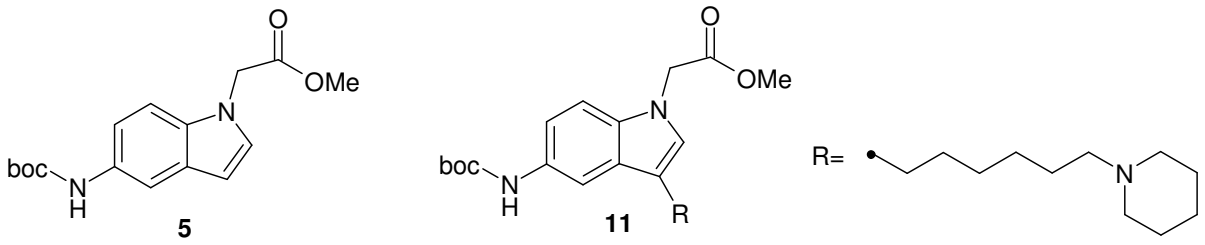


Fig. 26 Monomers.

On the other hand, in the context of fused-ring systems as G-quadruplex ligands, an important class other than macrocyclic compunds which includes perylene and acridine derivatives (par. 1.4.), indole looked promising as a base for new ligands. Considering the improved activity of the tri-substituted compounds with respect to the two-substituted and four-substituted compounds verified in different cases as BRACO-19 (par. 1.4., Read *et al.* **2001**), I worked on the synthesis and subsequently substitution of the triazatruxene core (fig. 27). In fact, the triazatruxene core consists of a C3 symmetric cyclotrimer of indoles, which presents a wide aromatic surface with three useful points for the attachment of side chains, namely the three indolic NH at 5, 10 and 15 positions (fig. 27) substituted triazatruxene derivatives. So, I decided to work on this kind of molecules in order to obtain tri-substituted compounds with optimal features as G-quadruplex ligands.

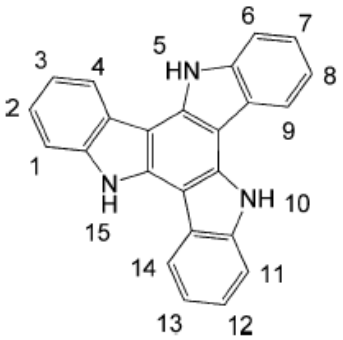


Fig. 27 Triazatruxene core.

Triazatruxene derivatives have also been of great interest in supramolecular chemistry and in particular in organic electronics (Lai *et al.* **2006**) but have not been reported for important pharmaceutical applications and we thought that this kind of compounds could represent a useful

basis to develop new interesting G-quadruplex ligands. In my work I considered different routes to obtain hydrosoluble triazatruxene derivatives, in order to rationalize the different approaches present in the literature and define the most efficient method for the preparation of several compounds of this series (Ginnari-Satriani *et al.* **2009**). So, I synthesized various hydrophilic compounds with different length of the chains in order to modulate the interaction of the ligand with the grooves of the G-quadruplex ligands, to optimize the binding conditions (Franceschin *et al.* **2009**).

Subsequently, in order to evaluate and to compare the activity of the new synthesized molecules we have studied their ability in inducing and/or stabilizing G-quadruplex DNA structures, as well as their selectivity towards G-quadruplex with respect to duplex DNA. To these aims, several experimental (ESI-Mass spectrometry, FRET, absorption spectroscopy) techniques have been used.

Upon the results of these biophysical studies, biological tests in vitro on the most promising compound have been performed to assay its potential anticancer activity.

At the same time, the spectroscopy behaviour (UV/Vis absorption spectroscopy and fluorescence emission spectroscopy) of synthesized compounds was studied.

In particular, since drug self-aggregation has been shown to be related to a lower telomerase inhibition and weaker interactions with the G-quadruplex with respect to unstacked molecules, the self-aggregation in aqueous solution of G-quadruplex ligands must be carefully considered in the model of their interaction with G-quadruplex DNA. For this reason, we have payed particular attention to the self association of the synthesized compounds as studied by UV-vis absorption spectroscopy and NMR spectroscopy.

3. Results and discussion

3.1. Indole-based macrocycles

In order to design new polyamido indole-based macrocycles as a simplified model of the G-tetrad, molecular modeling studies have been used to simulate the interaction of the ligand and the telomeric G-quadruplex structure. In particular *simulated annealing* experiments have been performed on the indole tetramer core on G-quadruplex monomeric structure (par. 1.1., Parkinson *et al.* **2002**). These studies have shown a good superimposing of the designed ligand with the terminal G-tetrad of the G-quadruplex (fig.28), with the four indolic unit superimposed on the four guanine of the G-quartet system. So, I decided to synthesize this macrocycle. In parallel, I thought to insert on this core four hydrophilic chains in order to improve the affinity for the quadruplex, due to an electrostatic interaction with DNA grooves. This led to the design of a second macrocycle with chains. Therefore, through a retro-synthetic analysis I found the two target monomers for a synthetic strategy inspired by peptides synthesis in order to obtain the two desired macrocycles with the same core (fig. 29).

The two resulting monomers, which differ for the insertion of the hydrophilic chain in 3-indole position, present the carboxylic and aminic functions suitable arranged in order to obtain the correct geometry of the final macrocycles (fig. 29). In particular they present the methilenic bridge bearing the carboxylic group, protected as ester, in 1-indole position and the amino group , BOC protected, in 5-indole position. In both cases we detected the starting compound as the commercially available 5-nitroindole (**1**).

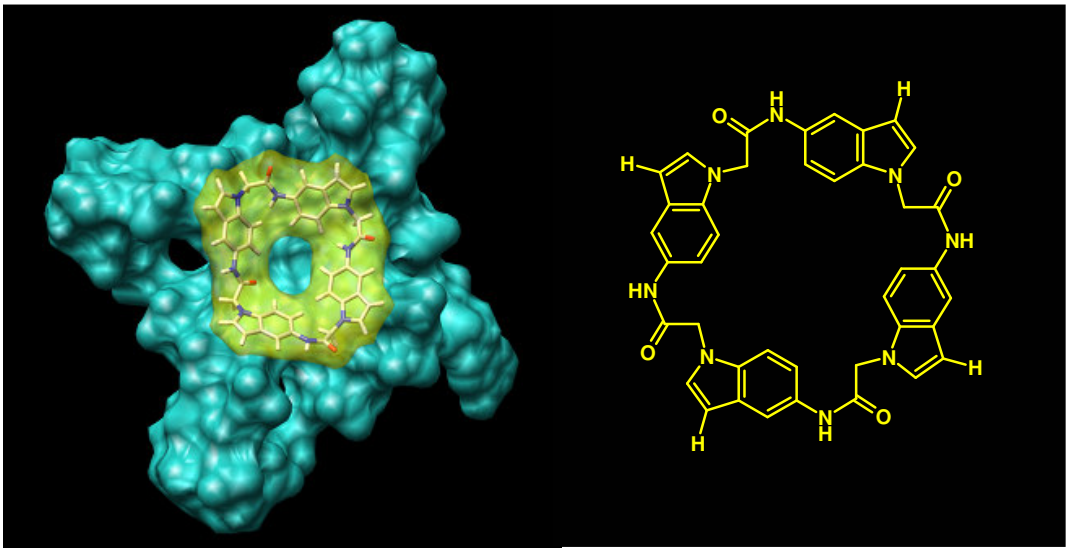


Fig.28 Simulated interaction of the indole-based macrocycle core with the G-quadruplex DNA.

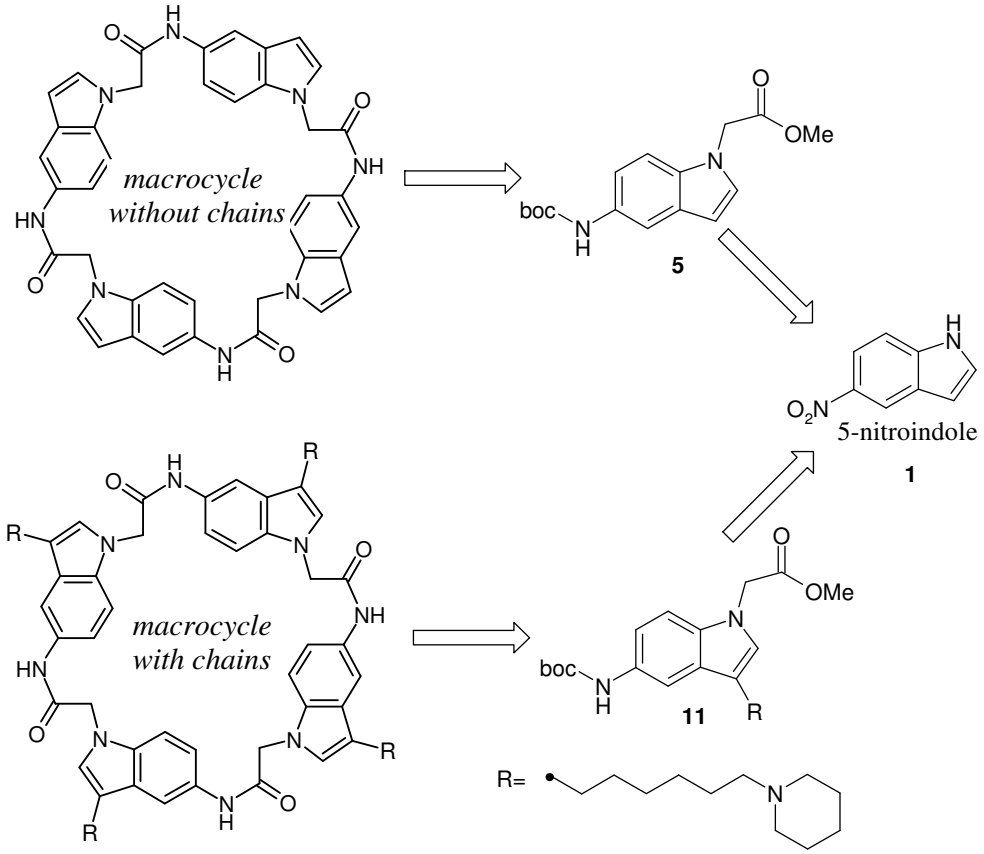


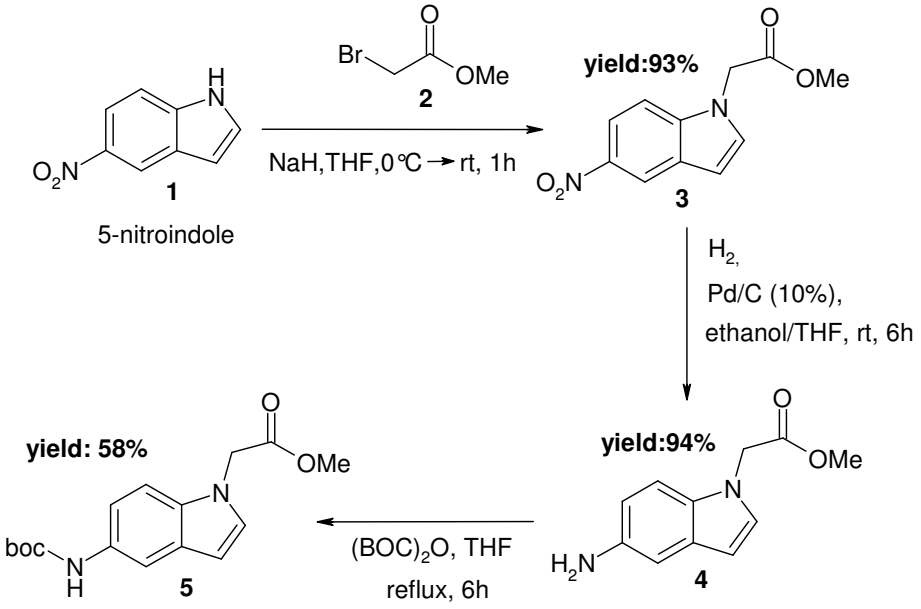
Fig. 29 Indole-based macrocycles, related monomers and the starting compound.

3.1.1. Preparation of the monomer without chains

To obtain the target monomer with the amino and carboxylic groups suitably directed and protected, we started from commercially available 5-nitro-indole (**1**) (Scheme 1).

In the first reaction we inserted the bridge bearing the carboxylic group protected as an ester, in basic conditions (NaH). This is a bimolecular nucleophilic substitution between the deprotonated indolic NH and methyl bromo acetate (**2**) to obtain the corresponding N-alkilated compound (**3**).

At this point the nitro group was reduced by a palladium-catalyzed hydrogenation to an amino group giving compound **4**, which was protected with BOC in a reaction with commercial di-tert-butyldicarbonate in THF to get the final compound **5** (Scheme 1).



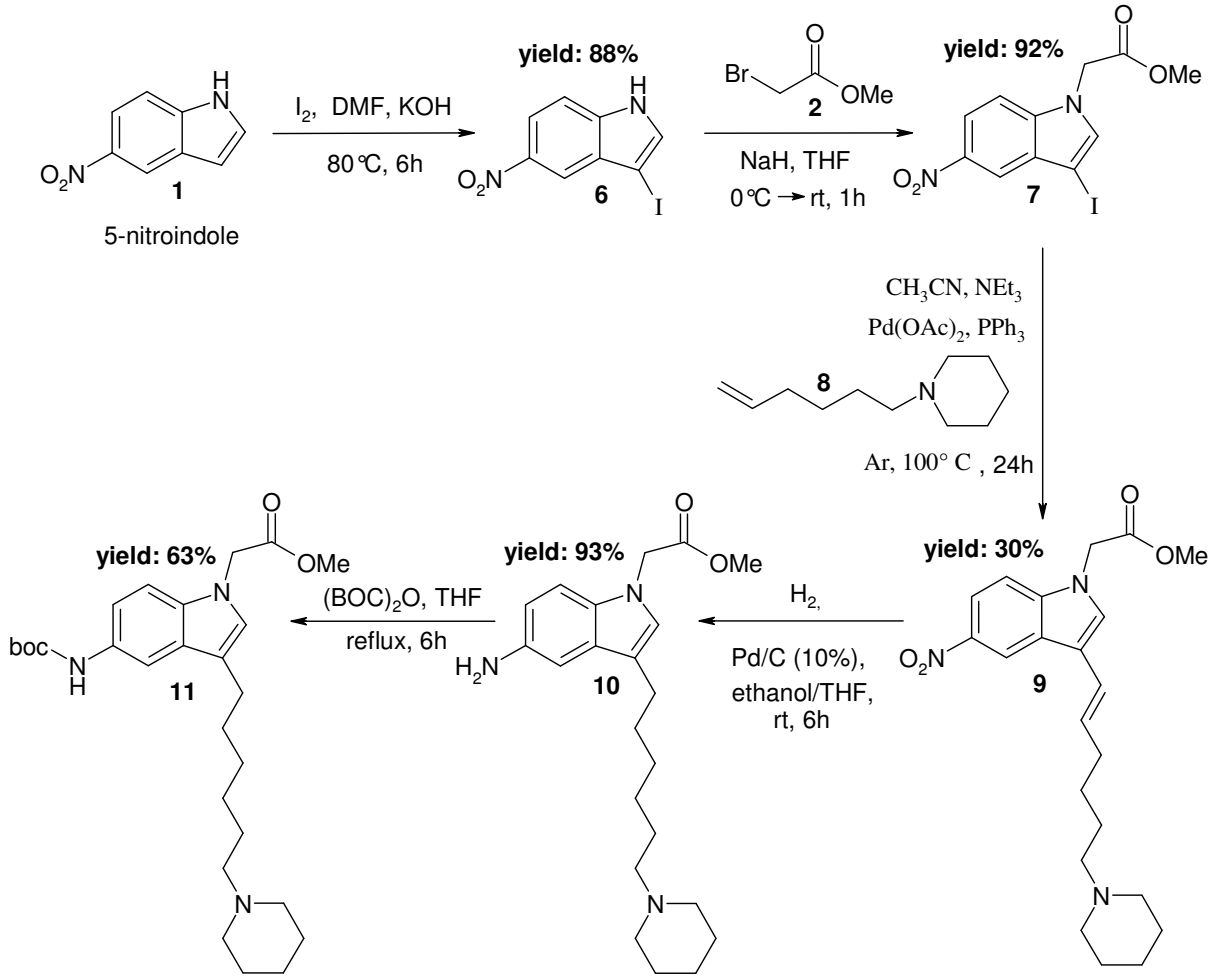
Scheme 1

3.1.2. Preparation of the monomer with the hydrosoluble basic chain

To obtain a target monomer similar to that previously described but equipped with a suitable hydrosoluble basic chain in order to get a final macrocycle with four chains, we decided to alkylate the monomer in 3-indole position.

As a first step, the commercial 5-nitroindole was iodinated in 3-indole position, in a reaction with I_2 in basic conditions to give the intermediated compound **6** (Scheme 2). Then, we inserted the bridge containing the protected carboxylic group, as described in the previous paragraph, to give the compound **7**. At this point we inserted the chain in the indole 3-position by a Heck reaction between

the compounds **7** and **8** to give the desired product (**9**). Subsequently, the latter was subjected to a palladium-catalyzed hydrogenation, reducing not only the nitro group but also the unsaturation on the side chain, to give the compound **10**. This compound was then N-BOC protected, as previously described, to give the final monomer (**11**) (Scheme 2).



Scheme 2

3.2. Coupling of two monomers in solution

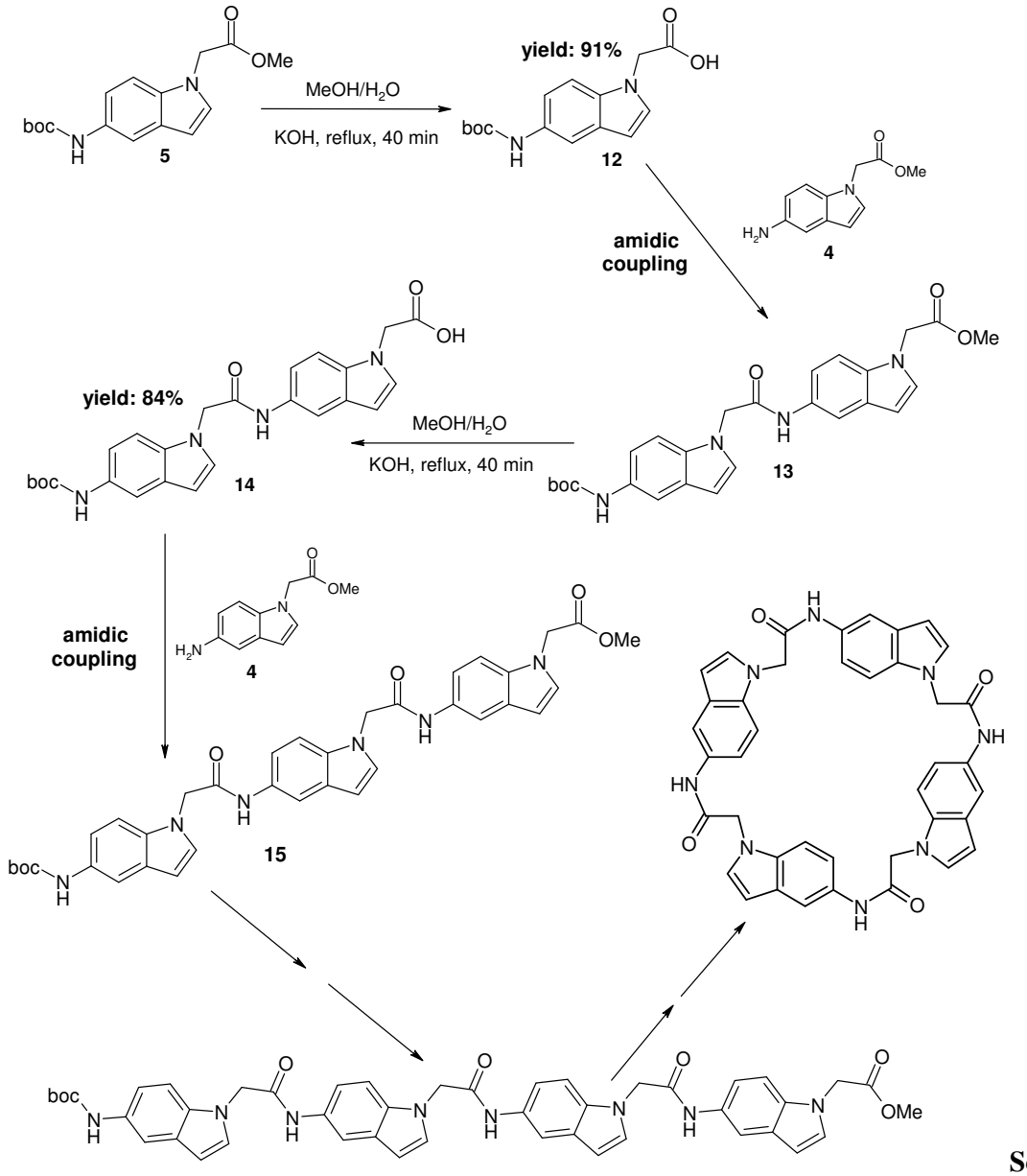
The next step in the synthetic pathway towards the designed indole-based macrocycles is the coupling of two monomers. We started from the simplest monomer in our hands: the monomer without chains (**5**). In particular, since we were inspired by peptides synthesis for this coupling, the precise steps are as follows:

- hydrolysis of the ester of monomer **5** in classic conditions to deprotect the carboxylic group;

- the amidic coupling of the obtained compound **12** with the precursor **4**, which has the free amino group, to give the dimer (**13**).

Further steps towards the final macrocycle include a ripetitive scheme (Scheme 3):

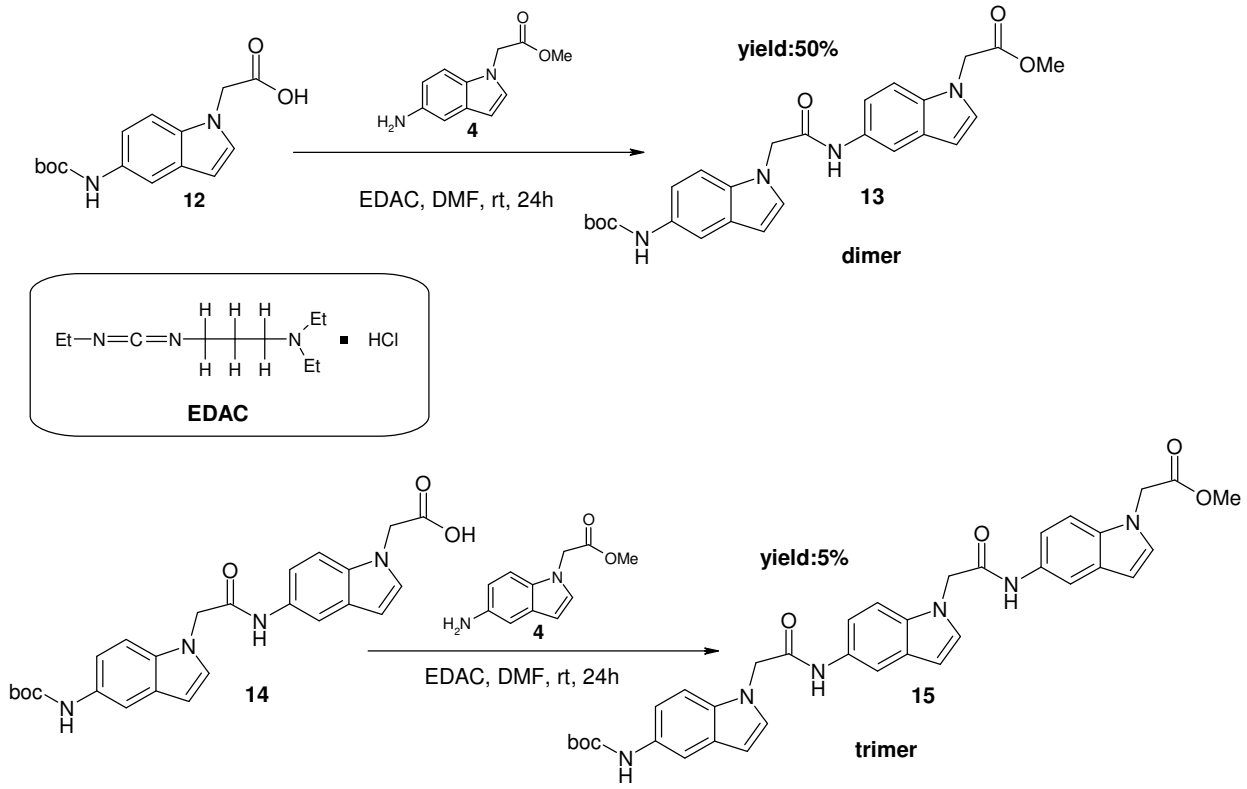
- deprotection of the carboxylic group of the dimer (**13**) to give the deprotected compound (**14**), which is coupled to the compound (**4**) with the same procedure described above to give the trimer (**15**);
- with the repetition of the same procedure on the trimer (**15**) is possible to synthesize the linear tetramer and subsequently to proceed towards the intramolecular coupling to give the desiderate macrocycle.



Scheme 3

3.2.1. Amidic coupling by EDAC

In this first approach for the coupling of two monomers in solution, we used a classic amidic coupling agent as EDAC (1-ethyl-3-(3-dimethylaminopropyl) carbodiimide. This is a useful hydrosoluble carbodimmide, easy to remove with water during the work up of the coupling reaction (Grayson **2000**). The first coupling in DMF between compound (**4**) and compound (**12**) led to the desiderate dimer (**13**) with a yield of 50%. Instead, the second coupling between the compound (**4**) and the dimer (**14**) got the desiderate trimer (**15**) with only 5% of yield (Scheme 3a).



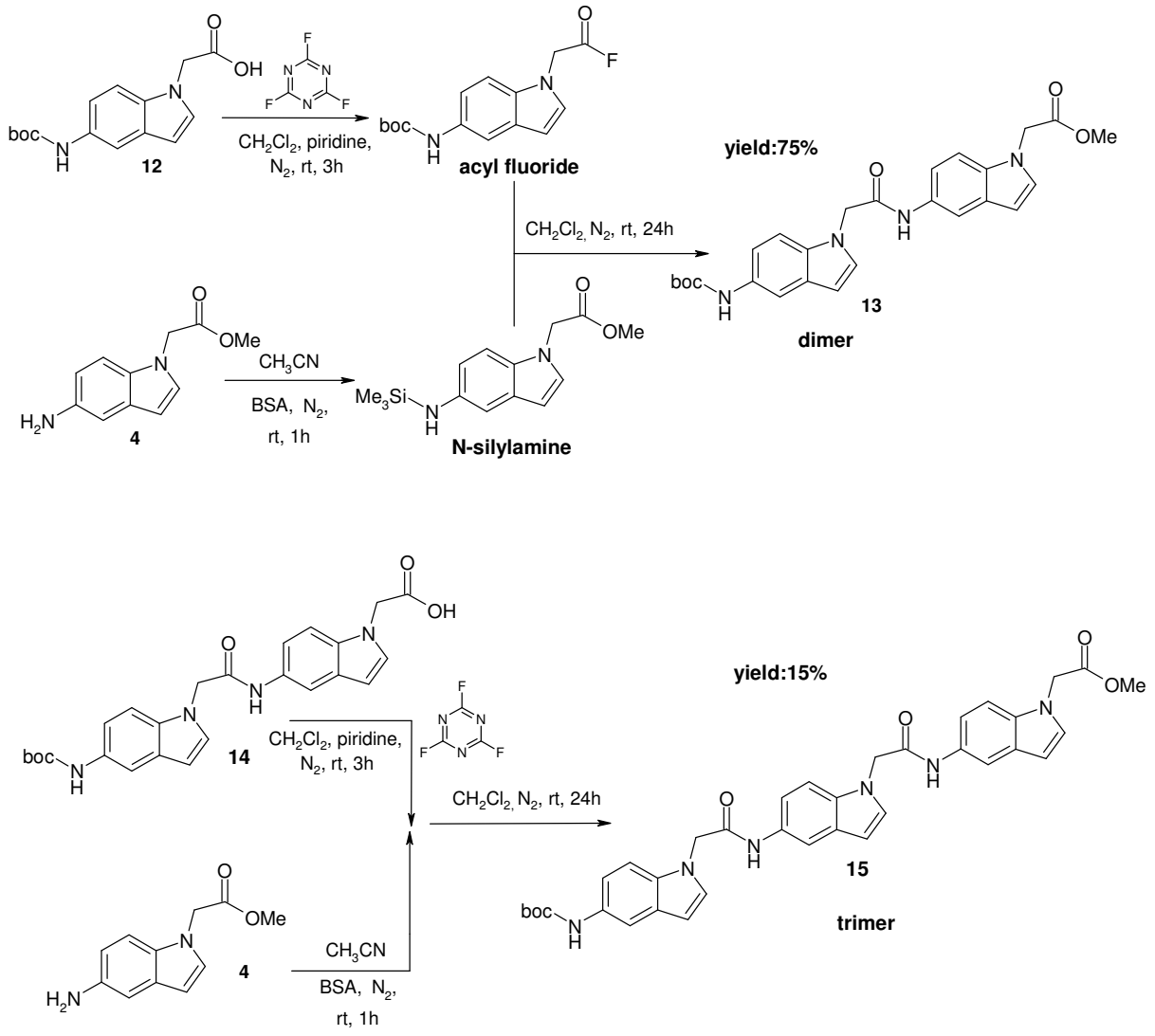
Scheme 3a

Probably, the drastic decrease of yield in the coupling step, in the case of the trimer with respect to the dimer, is due to the progressive increasing of the length of the polymer. In fact the increasing of the number of indole units advantages the possible intramolecular stacking between the different units causing the functional groups to be less reactive in the coupling reaction.

For this reason we tried to change the coupling agent to increase the reactivity of the ammine and carboxylic groups.

3.2.2. Amidic coupling between acyl fluorides and N-silylamines

Within a collaboration with Prof. Fernando Formaggio, I worked one month in his laboratory at the University of Padova to finalize new coupling methods for low reactive amminic and carboxylic groups. In this period, I improved the coupling conditions, bypassing the need of an external coupling agent. In fact, in this new strategy, the carboxylic group was activated as an acyl fluoride and the amine group as a silylamine (Rajeswari *et al.* **1987**): the two separately activated groups were then directly reacted together for 24h (Scheme 3b).



Scheme 3b

In particular, the carboxylic activation involved the reaction of the carboxylic acid with cyanuric fluoride in presence of pyridine for 3 hours (Carpino *et al.* **1990**), while the amine activation involved the reaction of the suitable aminic compound with BSA (N,O-

bis(trimethylsilyl)acetamide) in acetonitrile for 1 hour to afford the N-TMS derivative (Rajeswari *et al.* **1987**). After the subsequent work-up described in the experimental section, the two activated compounds were put to react together in dichloromethane under nitrogen atmosphere for 24 hours (Scheme 3b) .

In these conditions I obtained the desired compounds **13** and **15** with yields, respectively, of 75% and 15%: using this new method I improved the amidic coupling reaction with respect to the coupling reaction with EDAC (par. 3.2.1.). In particular, I increased the yield of the trimer (**15**) from 5% to 15%. Despite these results, the yield of compound (**15**) remained too low to continue the polymerization of the tetrameric compound in solution. So we thought to perform the synthesis of a linear tetramer, utilizing the approach of the solid phase synthesis of peptides. This step needs to be done in specialized research groups or industries. After this, I will utilize the obtained tetramer for the intramolecular cyclization in solution to obtain the desired macrocycle (Tera *et al.* **2008**; see Scheme 3). If successfully, the same procedure could be apllied to the monomer with the hydrophilic side chain (**11**) in order to obtain the second correspondent macrocycle.

3.3. Triazatruxene derivatives

Another approach towards an indole-based polycyclic aromatic system which I dealt with during my PhD work was represented by the preparation of the triazatruxene core, a cyclotrimer of indoles were the three monomers are fused in a new aromatic core. In recent years, triazatruxene (Fig. 2) derivatives have been of great interest in supramolecular chemistry and in particular in organic electronics (Gomez-Lor *et al.* **2009**), with regard to their potential applications as light-emitting diodes (Lai *et al.* **2008**), battery and capacitors (Talarico *et al.* **2008**). For their intrinsic photophysical and redox properties and their π-stacking capability (Garcia-Frutos and B. Gomez-Lor **2008**), these molecules can also behave as electroactive discotic liquid-crystals (Gomez-Lor *et al.* **2006**). More recently, a triaza-fullerene (fig.30) was synthesized, using the triazatruxene core as a precursor (Otero *et al.* **2008**). Nevertheless, so far, triazatruxene derivatives have not been reported for important pharmaceutical applications and we thought that this kind of compound could represent a useful basis to develop new G-quadruplex ligands, provided suitable hydrophilic side chains were added to the triazatruxene moiety.

In fact, the first hydrophilic triazatruxene derivative that I designed, AZATRUX (**24a**) (fig.31), showed optimal molecular features for the interaction with the G-quadruplex, as confirmed by molecular modelling simulations: the triazatruxene core stacks upon the terminal G-tetrad (fig.32a), while the hydrophilic side chains fit the DNA grooves (fig.32b).

In this section several synthetic approaches used to synthesize AZATRUX (**24a**) (fig.31)are described.

This compound represents the first of a series, in fact the defined versatile pathway has been then applied in our lab to synthesize other derivatives, with side chains of different length and basicity. The physico-chemical properties of the synthesized compounds have been studied, in order to evaluate their effects on the potential biological applications of these new hydrophilic compounds. Finally, several techniques have been used to study their affinity for the G-quadruplex structure and the selectivity with respect to duplex DNA.

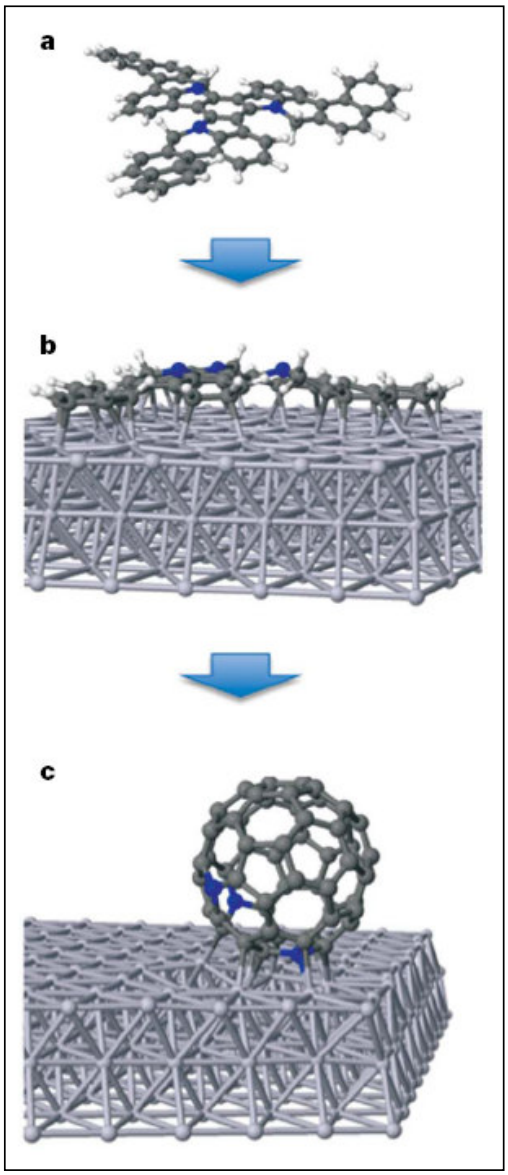


Fig.30

Optimized geometrical structure of the $C_{57}H_{33}N_3$ molecule at the different stages of the process (Otero *et al.* **2008**).

a) Representation of the optimized molecular structure of the $C_{57}H_{33}N_3$ molecule.

b) The same molecule adsorbed at room temperature (300 K) on the platinum surface.

c) The optimized structure for the $C_{57}N_3$ triazafullerene formed after the cyclodehydrogenation process at 750 K. Blue balls represent the nitrogen atoms in the structure.

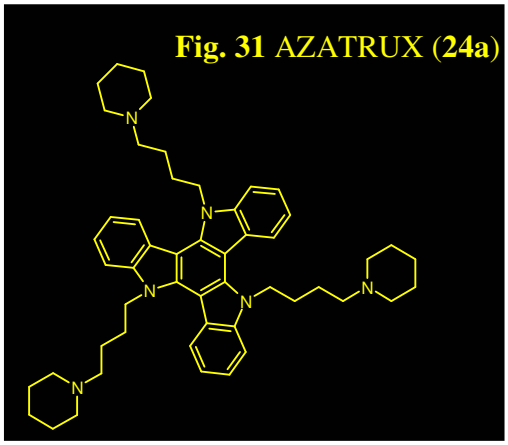
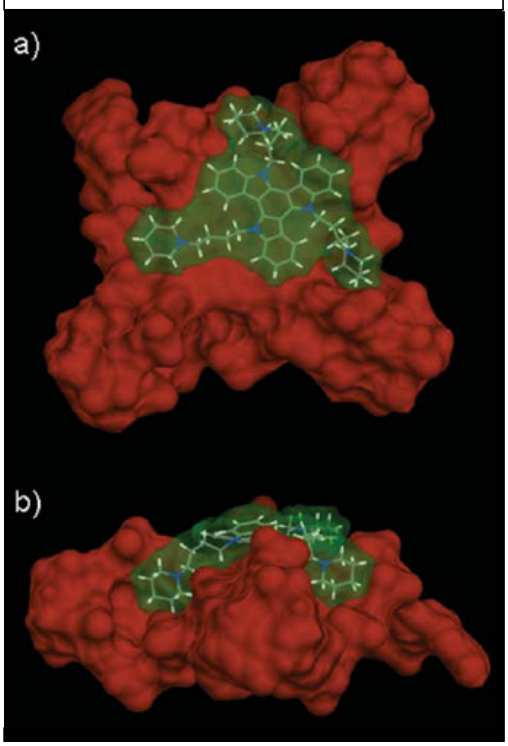


Fig. 31 AZATRUX (**24a**)

Fig.32 Complex of AZATRUX (stick model with yellow transparent surface) with the human monomeric G-quadruplex DNA**25** (red surface), obtained by simulated annealing. (a) top view, (b) lateral view.

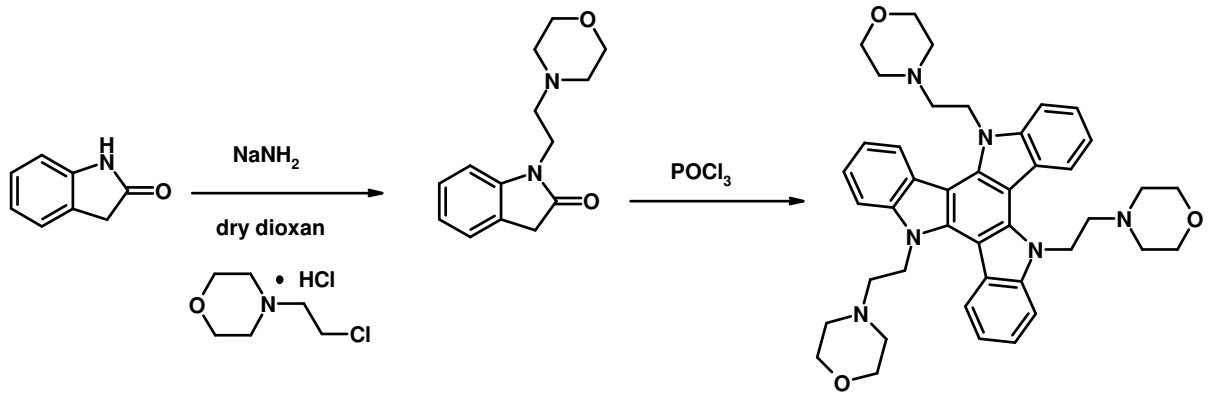


3.3.1. Synthesis of AZATRUX (24a)

Different routes to obtain AZATRUX (**24a**) (fig. 31) (Ginnari-Satriani *et al.* **2009**) have been considered, in order to rationalize the different approaches present in the literature and define the most efficient method for the preparation of several compounds of this series (Franceschin *et al.* **2009**).

3.3.1.1. Cyclotrimerization of N-substitued-2-indolones

As reported in the literature, a synthetic approach (Fig. 33) in order to obtain N-substituted hydrophilic triazatruxene derivatives is represented by the cyclotrimerization of N-substitued-2-indolones with $POCl_3$ (Hiyoshi *et al*. **2007**; Eissenstat *et al.* **1995**). This known procedure includes the preparation of the suitable N-substitued-2-indolone inserting directly the commercial chloro-amine side chain on the amidic NH of unsubstituted indolone in drastic basic conditions ($NaNH_2$) (Evans *et al.* **1965**).



Evans *et al.* **1965**

Hiyoshi *et al.* **2007**;
Eissenstat *et al.* **1995**

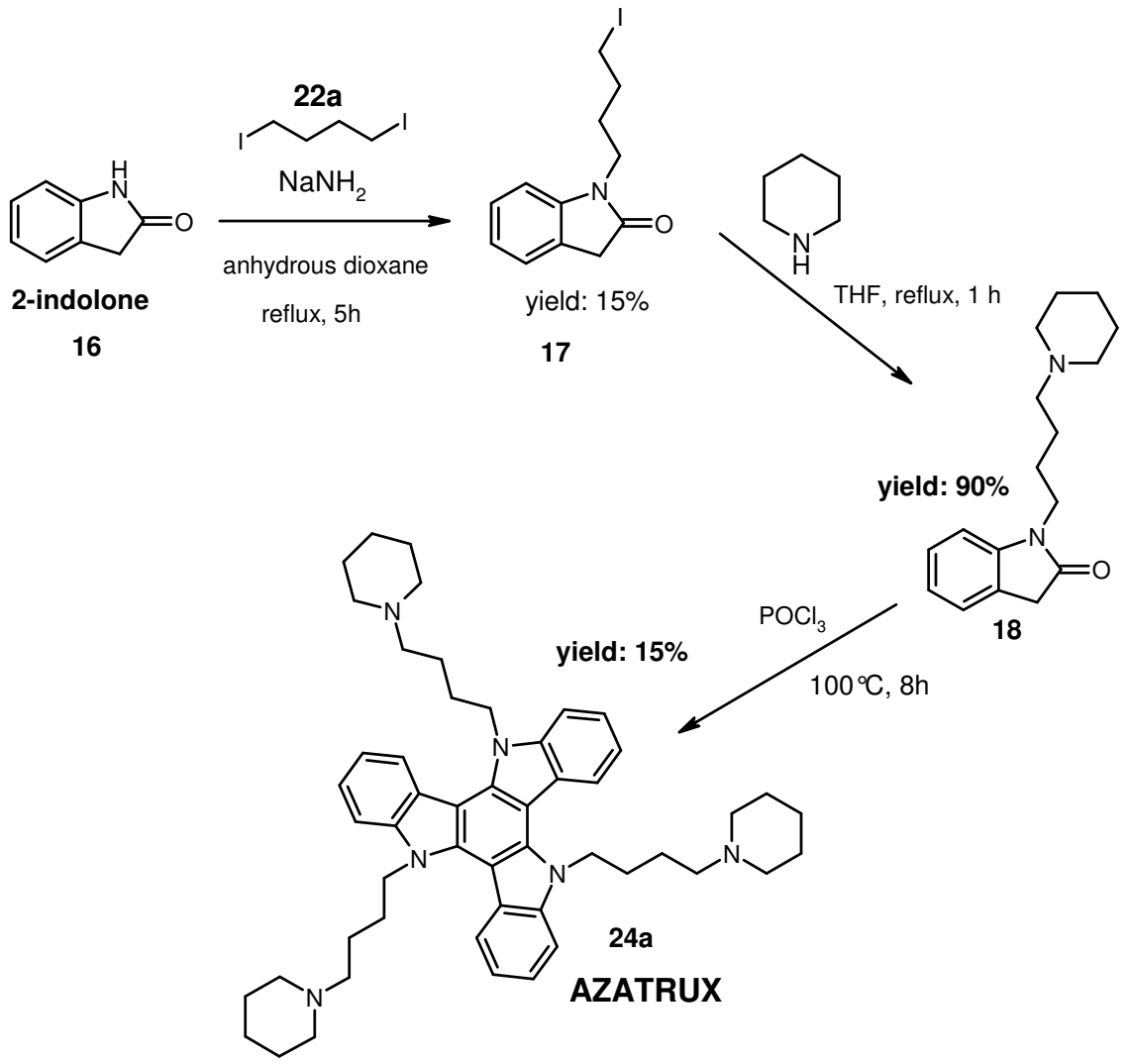
Fig. 33

Preparation of N-substitued-2-indolone (18)

Following this procedure, we decided to prepare the N-substitued-2-indolone (**18**) suitable for the preparation of AZATRUX (**24a**) (**Scheme 4**), the first synthesized triazatruxene derivative of our series. In our case the suitable halogen-amine wasn't commercially available, so we decided to insert the hydrophilic side chain in two steps in a similar way as previously reported (Franceschin *et al.* **2006**) (Scheme 4):

a) The N-alkylation in basic conditions ($NaNH_2$) with an excess of the commercial 1,4-diiodobutane (**22a**) to give the corresponding N-substituted indolone (**17**) (Evans *et al.* **1965**);

b) the following substitution of the iodine at the end of the chain with an excess of piperidine to get the final compound (**18**) (Franceschin *et al.* **2006**).



Scheme 4

The commercial unsubstituted indolone (**16**) was alkylated in basic conditions ($NaNH_2$) with an excess of the appropriate dihalide in dioxan under reflux. In this reaction it is very important to use an equimolar quantity of $NaNH_2$ and remove the NH_3 which is formed during the reaction (for example fluxing N_2 in the reaction mixture) after the adding of 1,4-diiodobutane (**22a**), in order to avoid a side reaction between NH_2^- or NH_3 and the iodine atoms present on the alkyl chain.

However, this reaction presents several problems, probably due to the scarce reactivity of the amidic NH, and the resulting yield was only 15%. The formation of (**17**) was confirmed by ^{1}H NMR and ^{13}C NMR. In particular, the ^{1}H NMR spectra in $CDCl_3$ shows the disappearing of the broad signal at δ 10.40 present in the spectra of the starting product relative to amidic NH, but is present a triplet at δ 3.73 relative to the CH_2 bounded at the indolic N, a triplet at δ 3.22 relative to the CH_2 bounded at iodine atom and the signal at δ 1.86 relatives to the two other CH_2 present on the side chain.

The low yield is surely also related to the necessity to use a dihalide instead of a monohalide.

The following substitution of the halogen at the end of the chain is easily carried out with an excess of piperidine under reflux to get the final compound **18** (Scheme 4) with a yield of 90%. The obtaining compound (18) is confirmed by ^{1}H NMR, ^{13}C NMR and by mass spectrometry which presents a peak at 295 m/z correspondent at $[(M+H)^{+}]$.

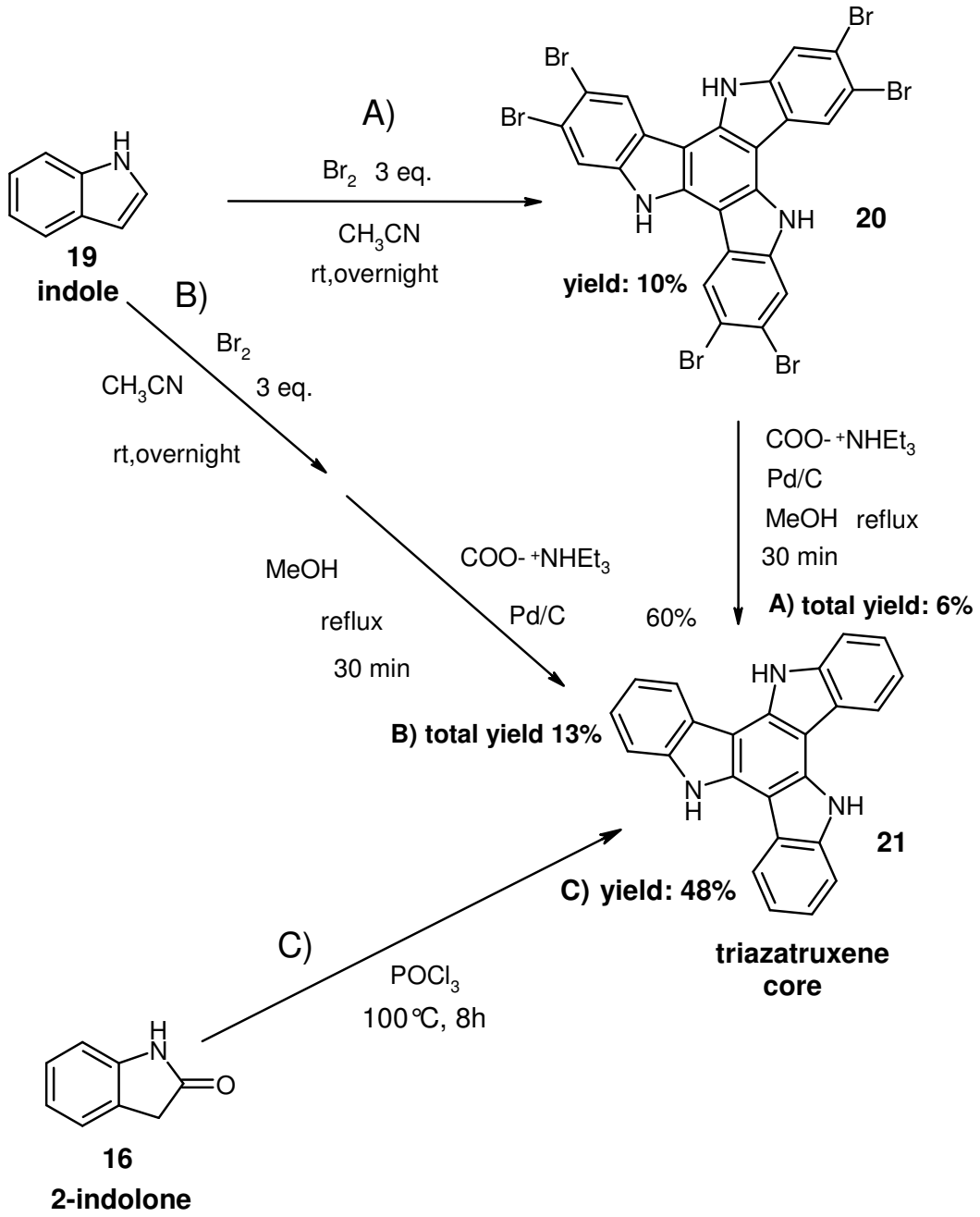
Cyclotrimerization step.

We have performed the symmetric cyclotrimerization of 2-indolone (**18**) in $POCl_3$ at 100°C, leading to AZATRUX (**24a**) in only 15% yield (Scheme 4). The resulting cyclotrimerization of **18** is confirmed by ^{1}H NMR, ^{13}C NMR , mass spectrometry and elemental analysis. In particular, ^{1}H NMR spectra shows the disappearing of the signal at δ 3.52 relative to the benzylic CH_2 bounded at the C=O of the starting product **18** suggesting the formation of a quaternary carbon. The shift of the triplet from δ 3.22 to δ 4.95 relative to the CH_2 bound to the nitrogen atom suggests the changing nature of this atom from amidic to indolic nitrogen. The number and the symmetry of signals are in according with the symmetric structure of the triazatruxene core. Moreover the increasing ^{13}H NMR spectra of the number of aromatic carbons from six to eight confirming the formation of indole skeleton and the symmetry of the indole condensation typical of the triazatruxene system.

Considering the scarce yield of the two key steps (chain insertion and cyclotrimerization), this synthetic approach did not result to be useful for the preparation of these triazatruxene derivatives. Consequently, in order to optimize the total yield, we tried to invert the steps order: first, the cyclotrimerization of the unsubstituted triazatruxene core and then its alkylation with the desired side chains. So, we began investigating the different preparations of the triazatruxene core.

3.3.1.2. Synthesis of the unsubstituted triazatruxene core

The most logical synthetic approach to obtain the unsubstituted triazatruxene core is the cyclocondensation of indole itself or unsubstituted 2-indolone via halogenation, similarly to what has been previously described for 2-indolone derivatives. In my work I compared these different synthetic routes for the preparation of this core, in order to optimize this cyclocondensation step (Scheme 5):



Scheme 5

Hexabrominated indole trimer: formation, isolation and subsequent dehalogenation (Scheme 5a)

The synthesis of the unsubstituted triazatruxene core in two steps has been reported in the literature:

1) cyclotrimerization of indole with 3 eq. of bromine (Robertson *et al*. **2000**),
2) isolation of the hexabrominated trimer and subsequent dehalogenation (Gomez-Lor and Echavarren **2004**).

Cyclotrimerization of indole with bromine

In the literature, different ways for the polymerization of indole units are known. In particular, indole can form different cyclotrimers, depending on the oxidative conditions. For example, the electrochemical polymerization of indole gives an asymmetric cyclotrimer (Mackintosh *et al.* **1995**) (fig. 34), formed by fusing the edges of indole molecules to give an extended π-system. Detailed NMR studies have confirmed the asymmetric structure of the electrochemically generated trimer.

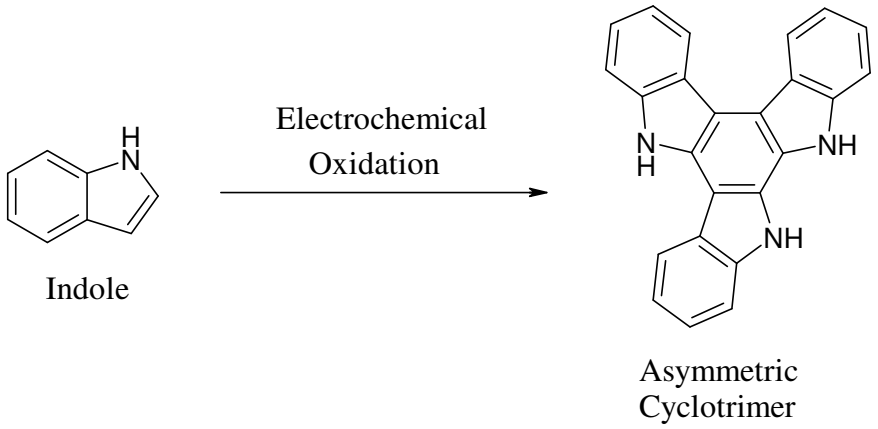


Fig. 34 Electrochemical cyclotrimerization of indole (Robertson *et al*. **2000**).

On the contrary, the symmetric cyclotrimerization of indole can be obtained by the bromination of indole in suitable conditions (Robertson *et al*. **2000**) in order to form the triazatruxene core.

Following this procedure, cyclotrimerization of indole with bromine was carried out reacting commercial indole (**19**) in acetronitrile with 3 eq. of bromine to give a mixture of brominated trimers of indoles. In this cyclotrimerization only the symmetric cycle-trimer can be obtained, since the reaction proceeds via a series of 2,3-couplings of monomeric units (fig. 35): probably, the reacting intermediate is the 3-bromo-indole. In this conditions, the formed triazatruxene core react with the excess of bromine in uncontrolled reactions to give a mixture of brominated compounds.

On the other hand, decreasing the quantity of bromine compromised the success of cyclotrimerization (Robertson *et al*. **2000**).

In fact, although in principle one equivalent of bromine would be sufficient to complete the cyclotrimerization, as a matter of fact in these conditions only the 2,3-dimer was isolated, while three equivalents were necessary to obtain brominated symmetric indole trimers (Robertson *et al.* **2000**) (Scheme 5a).

.

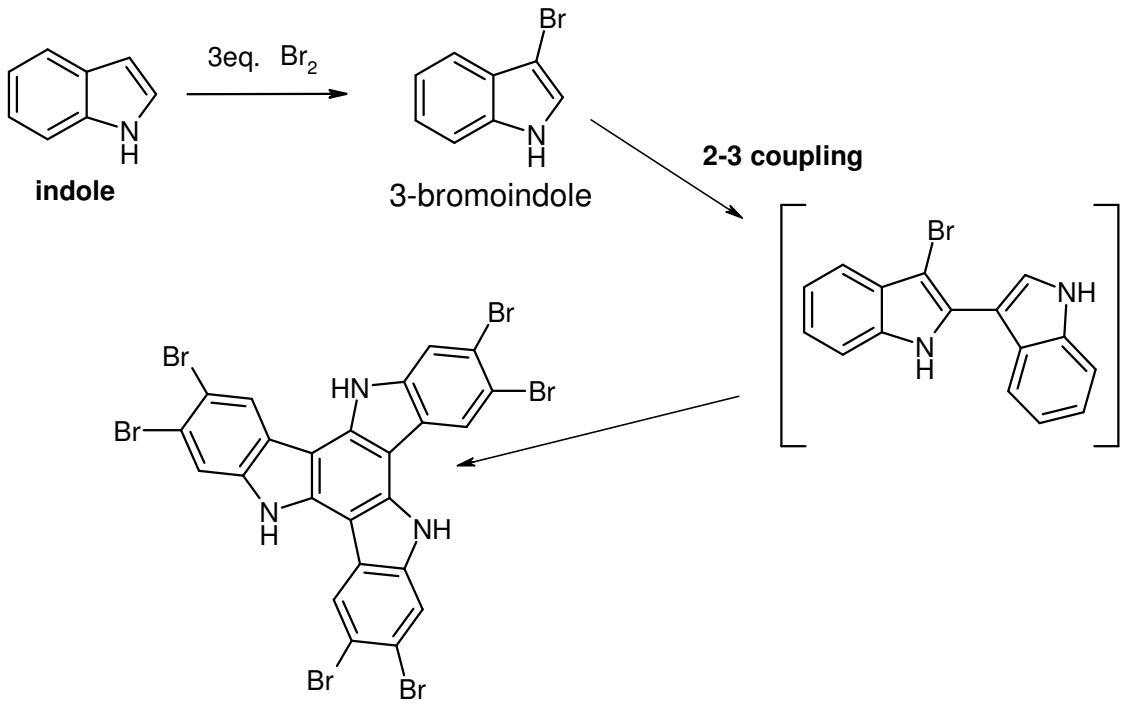


Fig. 35 Hypothesized mechanism of the symmetric cyclotrimerization of indole with bromine.

The hexabromination allows the trimer to be readily isolated due to its lower solubility, avoiding lengthy and difficultly chromatographic procedures. So, the hexabrominated cyclo trimer (**20**) was recrystallised with 10% of yield (Scheme 5a). The formation of **20** was confirmed by ^{1}H NMR and Mass Spectrometry. The ^{1}H NMR spectra in DMSO show three signals: one broad signal at δ 12.30 relative to indolic NH and two aromatic singlets at δ 8.59 and δ 7.89 related to two protons at 4- and 7- positions of indole. This spectra results consistent with the structure of a symmetric indole trimer brominated at the 5- and 6- positions of indole. The ESI-MS spectra confirmed the achievement of compound **20** showing a peak at 818 m/z corresponding of the ion $(M-H)^{-}$.

***Dehalogenation*:**

Following the reported procedure (Gomez-Lor and Echavarren **2004**), the dehalogenation of hexabrominated compound **20** was realized by reduction with formic acid in triethylamine, catalyzed by Pd/C, obtaining the desired compound **21** in 60% yield (Scheme 5a). The ^{1}H NMR spectra in acetone was consistent with the structure of a dehalogenated triaztruxene core (**21**): there are one broad signal at δ 11.08 relative to indolic NH, two doublets at δ 8.52 and δ 7.68 respectively related to two protons at 4- and 7- positions of indole, and two overlapped triplets in δ 7.4-7.2 related to two protons at 5- and 6- positions of indole.

The total yield for this schema (Schema 4, A) was 6%.

Cyclotrymerization of indole and subsequent dehalogenation without intermediate purification (Scheme 5b)

In this case the crude product of the cyclotrimerization of indole with bromine previously described was directly dehalogenated without further purification. In fact, as expected, the mass spectrum of the crude product of the first reaction step showed a mixture of brominated indole trimers (Fig. 36).

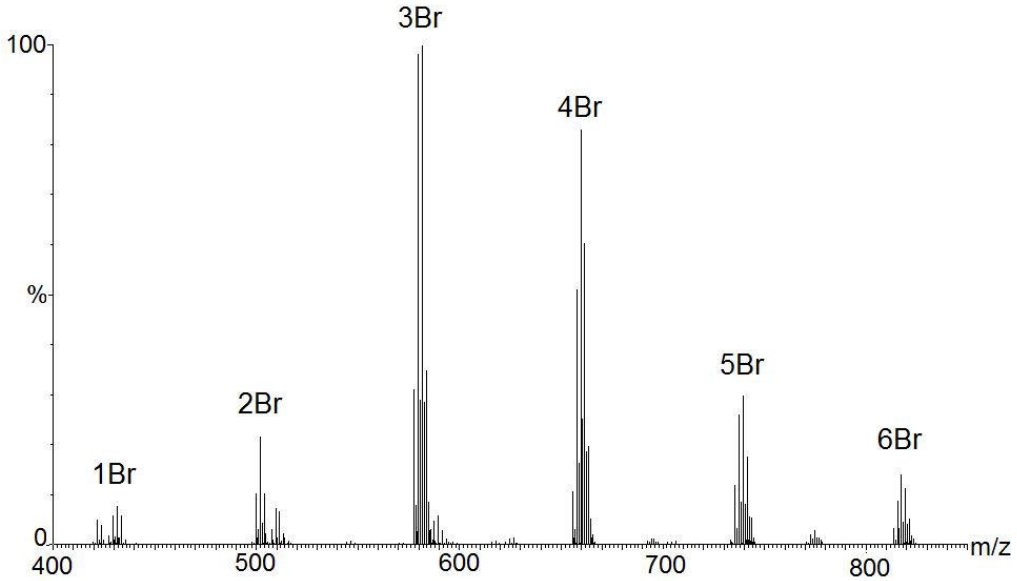


Fig. 36 Mass spectrum of the crude product of the first reaction step of Scheme 5A, showing a mixture of brominated indole trimers, whose number of bromine atoms is indicated by the labels.

Since, the main component of this mixture was not the hexabrominated trimer, we wondered if its isolation was really necessary. So, we poured this mixture, without further purification, in the reaction conditions for the Pd-catalyzed dehalogenation (Scheme 5b), obtaining the triazatruxene

core (**21**) in a total yield of 13% higher than before, but still comparable to that reported in the literature.

It is worth noting that the low yields of these two synthetic routes did not allow to consider promising the insertion of the side chains on the indole nitrogen atom before the cyclization step.

Cyclotrimerization of unsubstituted 2-indolone in $POCl_3$ (Scheme 5c)

As previously described, in the alternative synthetic approach, starting compounds are represented by N-substitued-2-indolones, which are reacted with phosphoryl chloride to give the corresponding tris-N-susbstituted-triazatruxenes (par. 3.2.1.1.). Following an analogous strategy, we have performed the symmetric cyclotrimerization of unsubstituted 2-indolone (**16**) in POCl3 at 100°C (Kim *et al*. **2005**), leading to the unsubstituted triazatruxene core (**21**) in 48% yield (Scheme 5c).

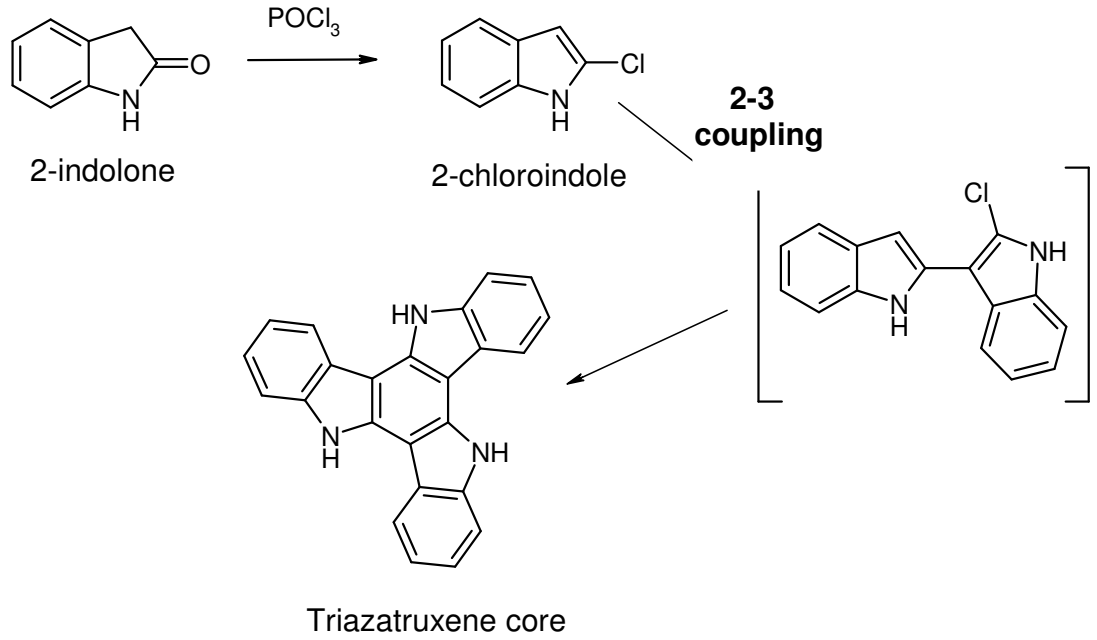


Fig. 37 Hypothesized mechanism of the symmetric cyclotrimerization of 2-indolone with $POCl_3$.

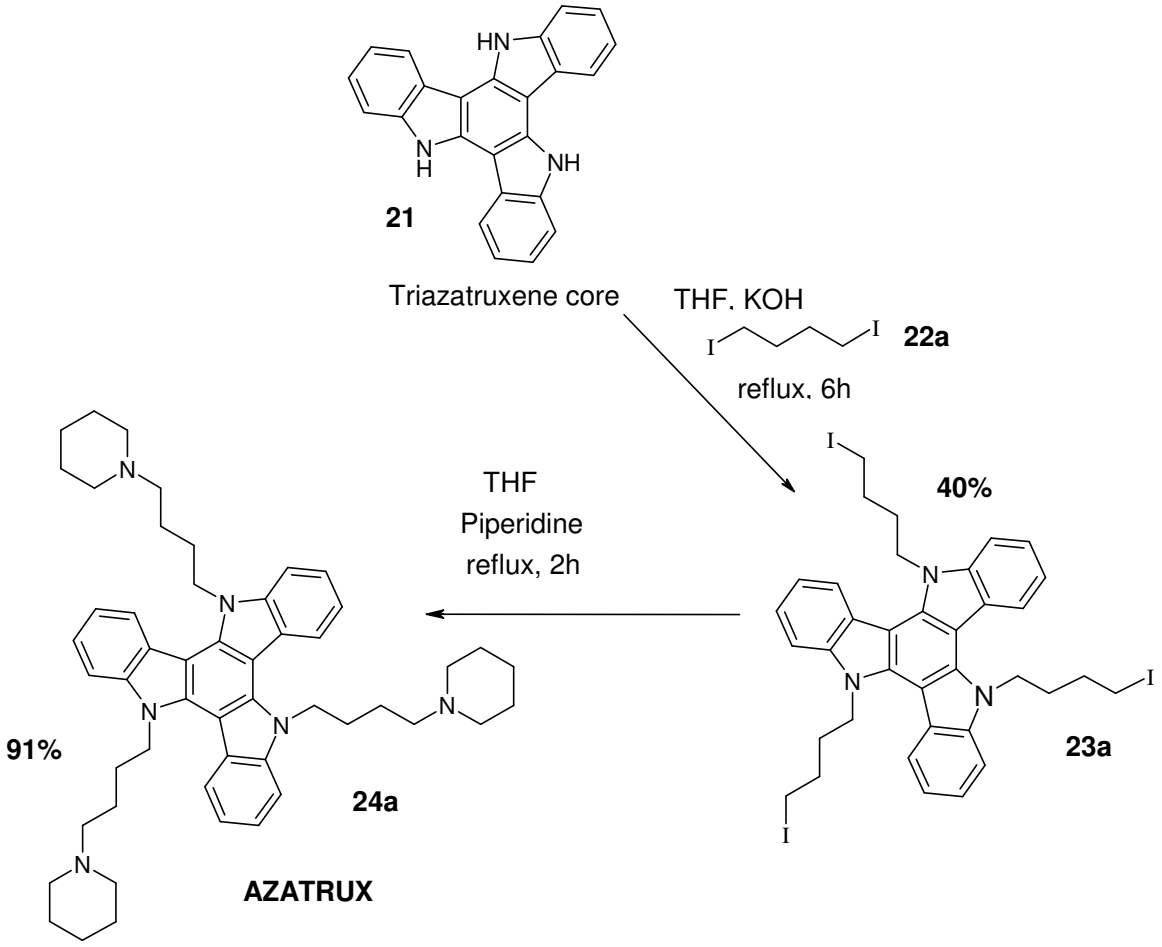
In this cyclotrimerization the starting product is 2-indolone instead of indole and the halogenating agent is $POCl_3$ instead of bromine. In this case the intermediate is probably 2-chloroindole and the reaction is supposed to proceed via a series of 2,3-couplings of monomeric units (fig. 37) similarly to the previously described cyclotrimerization of indole with bromine (fig. 35).

The main difference is the incapability of $POCl_3$ to substitute the indole 5- and 6- position before or after triazatruxene formation: so, the desiderate unsubstituted triazatruxene core can be obtained in

a one-pot reaction. To our knowledge, this way represents the best synthetic approach to obtain the unsubstituted triazatruxene core (**21**) in terms of time, number of steps and yield. In particular, with the cyclocondensation of the unsibstuted 2-indolone we improved the yield of this step with respect to the N-substituted 2-indolone previously reported (par. 3.3.1.1.) from 15% to 48%.

3.3.1.3. Insertion of the basic hydrosoluble chains in two steps

Once optimized the cyclocondensation step with the preparation of the unsubstituted triazatruxene core, we focused our attention on the insertion of the three hydrophilic side chains. As described above, in our case the suitable halogen-amines weren't commercially available, so we decided to insert the hydrophilic side chains in two steps.



Scheme. 6

In the first step, the triazatruxene core (**21**) was alkylated in basic conditions (KOH) with an excess of the appropriate dihalide in THF under reflux (Scheme 6). In these conditions each indolic NH is deprotonated, becoming a good nucleofilic agent able to substitute the iodine atom. The large excess of 1,4-diiodobutane (**22a**) is essential to limit the polymerization between different triazatruxene cores. It was possible to follow the partially substitution of the triazatruxene core from mono- to tris- substituted triaztruxene by TLC (hexane/ethyl acetate 80:20): after 6 hours only the tris-substituted compound (**23a**) was presents in 40% yield. ^{1}H NMR spectra of **23a** shows diagnostic signals in accord with the tris N-alkylation: the broad signal at δ 11.08 present in the spectra of the starting product relative to indolic NH is not present, but it is present a triplet at δ 5.03 relative to the CH_2 bounded at the indolic N, a triplet at δ 3.07 relative to the CH_2 bounded at iodine atom and the signals at δ 1.89 and δ 1.61 relative to the two other CH_2 present on the side chains. The number and the symmetry of the signals confirmed the tris N-substitution of the triazatruxene core. The formation of **23a** is also confirmed by ^{13}C NMR spectra and by mass spectrometry spectra which presents a peak at 892 m/z correspondent at $[(M+H)^+]$.

The better result of the substitution of the triazatruxene core (40% yield) with respect to 2-indolone (15% of yield) is probably related to the stronger reactivity of the NH group of indole (**19**) with respect to amidic NH of 2-indolone (**16**) in this kind of reaction, (par. 3.3.1.1., scheme 4).

The following substitution of the halogen at the end of the chain with an excess of piperidine was carried out to get the final compound **24a** (91% yield) (Scheme 6).

In conclusion, the best synthetic pathway to obtain AZATRUX (**24a**) is to perform first the cyclotrimerization in $POCl_3$ of the unsubstituted 2-indolone (**16**) (Scheme 5c) obtaining the triazatruxene core (**21**) and, then, the insertion of hydrophilic side chains in two steps (Scheme 6) (Ginnari-Satriani *et al.* **2009**). In this way we increased the total yield (cyclization and chain insertion, tab.) with respect to the first synthetic approach used, which was represented by the N-substitution of 2-indolone and the subsequent cyclotrimerization step in $POCl_3$ (scheme 4) (Franceschin *et al.* **2009**). The complete comparison of the two different synthetic routes, the partial yields of the two main steps and the total yields of the final compound AZATRUX (**24a**) are reported in Table 1.

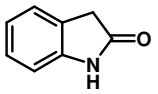
Starting compound	Schemes	Cyclization step (yield)	Chains insertion step (yield)	AZATRUX (24a) yield
 16	4	15%	13,6%	2%
	5C and 6	48%	36%	17%

Table.1 Comparison of the yields of the cyclization and chains insertion steps and the total yields of the final compound AZATRUX (24a) for the different synthetic routes.

For the reasons discussed above we decided to apply this useful method to prepare several compounds of this series with longer side chains with respect to AZATRUX (**24b-d**) (see below).

3.3.2. Studies of selective interactions of AZATRUX for G-quadruplex DNA with respect to duplex DNA by ESI-MS experiments.

The efficiency of AZATRUX (**24a**) in binding the G-quadruplex and its selectivity with respect to duplex DNA has been studied by electrospray ionization mass spectrometry (ESI-MS) (Ginnari-Satriani *et al.* **2009**).

Electrospray ionization mass spectrometry (ESI-MS) is a powerful tool for studying biomolecular structures and non-covalent interactions. In particular this technique allows the transfer of non covalently bound complexes into the gas phase without the disruption of the complex itself and therefore the determination of the stoichiometry and, in particularly favourable cases, modes and energies of interaction. In particular, in the field of the analysis of noncovalent complexes between nucleic acids and small molecules, De Pauw and coworkers (Gabelica *et al.* **2000**) used ESI-MS to study the interaction of double-stranded DNA with two classes of antitumor drugs: intercalators (ethidium bromide, amsacrine and ascididermin) and minor groove binders (distamycin A, Hoechst, netropsin, berenil and DAPI) (Gabelica *et al.* **1999**). Using the data obtained from mass spectra, they also evaluated the DNA affinities of the minor groove binders, by quantifying the equilibrium association constants of the observed complexes and, furthermore, they demonstrate the good consistence of these values with those obtained from other traditional techniques. In the past few years, ESI-MS has begun to be used in the study of nonconventional DNA structures, including DNA triplex and especially G-quadruplexes (Rosu *et al.* **2002**). In particular, this technique has been successsfully applied to the study of the binding of G-quadruplex ligands to their target sequences in order to determinate the stoichiometries and relative binding affinities of such complexes (Rosu *et al.* **2008**). In fact, quantitative analysis of binding affinities with quadruplex DNA structures is possible, because the association constants can be calculated directly from the relative intensities of the corresponding peaks found in the mass spectra, under the assumption that the relative intensities in the spectrum are proportional to the relative concentrations in the injected solution.

Evaluation of binding constants

For this study, we have chosen two oligonucleotides that can form different G-quadruplex structures: TG4T (5'-TGGGGT-3') which gives a tetrameric G-quadruplex and F21T (5'-GGGTTAGGGTTAGGGTTAGGG-3') which is composed by human telomeric repeats and is able to fold in a monomeric G-quadruplex structure, characterized by X-ray crystallography in K^+ solution (1.1.). The formation of stable complexes between AZATRUX and the G-quadruplex forming oligonucleotides is clearly demonstrated by the presence in the mass spectra (at 1:1 drug/quadruplex ratio) of intense peaks corresponding to drug-quadruplex complexes (fig.38). The evaluation of the binding constants obtained by the collected data (fig.39) demonstrates that the studied molecule is a good G-quadruplex ligand able to form both 1:1 and 2:1 drug–DNA complexes. This agrees with the general model for the G-quadruplex–ligand interactions characterized by the presence of two binding sites on the external tetrad surfaces of the G-quadruplex structure, even though other explanations are possible for the 2:1 stoichiometry, since the external tetrads of the quadruplex are not identical. The values of K_1 are above 10^5 for F21T and above 10^6 for $(TG_4T)_4$, while the order of magnitude for K_2 is 10^5 with both the oligonucleotides. Comparing K_1 and K_2 we can assume that no cooperativity is involved in the binding mechanism since K_2 values are always lower than those for K_1. In order to evaluate the selectivity of AZATRUX for quadruplex over duplex DNA we have studied preliminarily its affinity for a self complementary dodecamer: DK66 (5'-CGCGAATTCGCG-3'), one of the simplest duplex models widely used in the literature (see above). In this case, the spectra acquired at 1:1 ratios do not show any trace of the 1:1 or 2:1 complex peaks, suggesting a weaker interaction with respect to that revealed for G-quadruplexes. In order to detect an appreciable peak of the 1:1 drug/DNA complex, this molecule must be present in the sample at 3:1 and 4:1 ratios, but even at this concentration, there is no evidence of the peak relative to the 2:1 complex (fig.38). The order of magnitude for K_1 with DK66 is 10^4, about 25 times lower than that of K_1 with F21T and more than 150 times lower than K_1 with TG_4T (Table 2). These values suggest a good selectivity for quadruplex structures with respect to duplex DNA (Ginnari-Satriani *et al.* **2009**). Nevertheless, we have previously shown that a short duplex oligonucleotide is a really simplified model and physiologically non relevant interactions are possible with this kind of model.

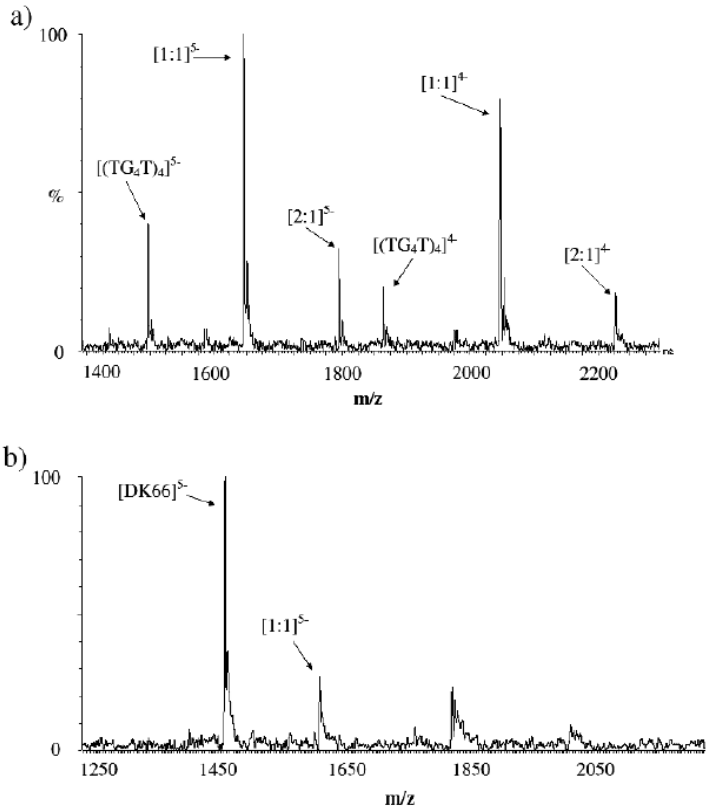


Fig. 38 ESI mass spectra of $(TG_4T)_4$ in the presence of AZATRUX at 1:1 molar ratio (a) and of DK66 in the presence of AZATRUX at 1:4 molar ratio (b). Arrows indicate the peaks corresponding to the oligos alone and to the relative 1:1 or 2:1 drug/DNA complexes, in different charge states.

DNA / Drug	F21T		$[TG_4T]_4$		DK66	
	log K1	log k_2	log K_1	log k_2	log K_1	log k_2
AZATRUX (**24a**)	5.5±0.1	5.0±0.1	6.3±0.1	5.3±0.1	4.1±0.1	n.d.

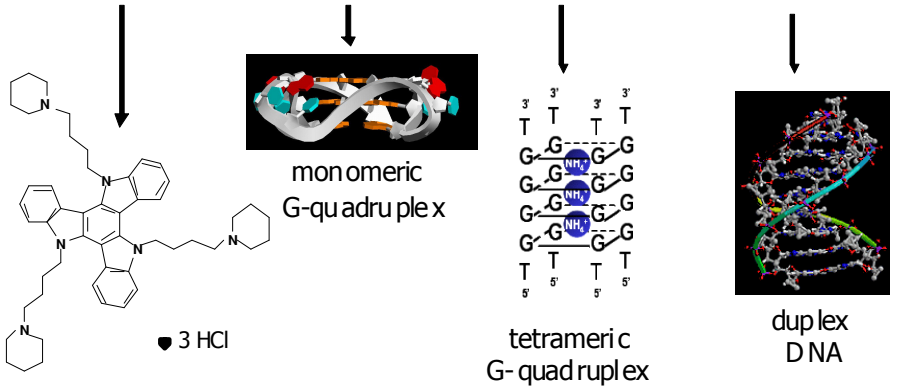


Table 2 K_1 and K_2 values on a logarithmic scale for the complexes between AZATRUX and the indicated oligonucleotides as derived by ESI-MS experiments.

Selective studies: competitive experiments

In order to evaluate selectivity between quadruplex and duplex DNA it is necessary to choose an appropriate model for duplex DNA: Dickerson-like dodecamers are among the simplest models and have been widely used in the literature (Gornall *et al.* **2007**).

The possibility of performing experiments in the simultaneous presence of G-quadruplex structures and a double-stranded genomic DNA is of particular interest for the biological relevance of this system, so after exploring this topic by a classical approach based on the very simple duplex model of an autocomplementary dodecamer, in my group we extended our analysis, reporting for the first time a competition ESI-MS experiment in the presence of genomic DNA fragments. Whereas those ligands showing high levels of selectivity between quadruplex and duplex oligonucleotides in terms of binding constants confirmed their selectivities in the competition experiment, the contrary was not always true: some ligands showing poor selectivity with respect to the autocomplementary dodecamer were selective in the presence of genomic DNA fragments (Casagrande *et al.* **2009**). This result suggests that physiologically non relevant interactions are possible with short duplex oligonucleotides. This is the case with the coronene derivative CORON, which would have been said not to be selective according to the data obtained with the short duplex oligomer, whereas its binding to quadruplex DNA is poorly affected by the presence of calf thymus DNA (Casagrande *et al.* **2009**).

So, in order to achieve a more reliable study of selectivity, we have performed competition experiments on AZATRUX, carried out in the simultaneous presence of G-quadruplex forming oligonucleotides and fragments of a double stranded genomic DNA in 1:1 and 1:5 ratios, calculated on the basis of the phosphate group concentrations (Casagrande *et al.* **2009**). Provided that calf thymus DNA (CT) cannot be detected in the used experimental conditions due to its high molecular weight, it is possible to report the percentage of bound quadruplex DNA in the presence of different duplex concentrations (Fig. 39): any decrease of this percentage must be caused by the drug binding to genomic duplex DNA. The analysis of the measures performed on both oligonucleotides shows that the percentage value of quadruplex bound at a 1:1 duplex/quadruplex ratio has a poor decrease: from 75 to 64 with TG_4T and from 46 to 39 with F21T, which is more than 80% left with respect to the value in the absence of CT (Tables 3 and 4).This result confirms the strong preference of this molecule for the quadruplex structure of DNA. At a 5:1 duplex/quadruplex ratio, AZATRUX is still able to bind 44% of the quadruplex DNA formed in the sample by TG_4T, which is about 2/3 of the quadruplex bound in the absence of duplex DNA, while in the case of F21T the percentage of quadruplex bound is remarkably reduced, probably due to a weaker affinity for this oligonucleotide with respect to TG_4T, even alone (Fig. 39).

	[CT] = 0	[Gq]:[CT] 1:1	[Gq]:[CT] 1:5
F21T	46	39	14
$(TG_4T)_4$	75	64	44

Table 3: Values of the percentage of bound quadruplex DNA for samples containing a fixed amount of both drug and G-quadruplex DNA (5 μM, 1:1 ratio) and different amounts of calf thymus DNA (CT), at the indicated quadruplex/duplex ratios (in phosphate ions). All values should be considered with an error estimated on at least three independent experiments of about ±5%.

	[CT] = 0	[Gq]:[CT] 1:1	[Gq]:[CT] 1:5
F21T	1	0.85	0.30
$(TG_4T)_4$	1	0.85	0.59

Table 4: Values of normalized percentage bound quadruplex (N%) are elaborated starting from the percentages reported in Table S1 using the following equation:

$$N\% = \frac{\text{\% quadruplex bound in presence of CT}}{\text{\% quadruplex bound in absence of CT}}$$

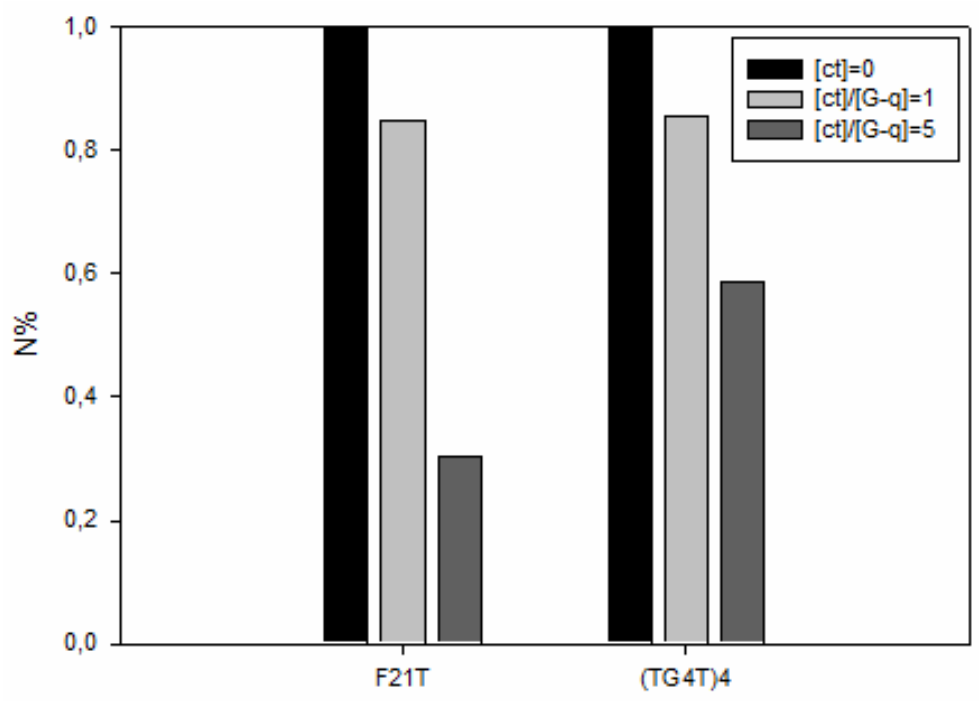


Figure 39: Competition experiments on F21T and $(TG_4T)_4$ oligos. Histogram diagrams reporting normalized percentage bound quadruplex (N%, Table 3).

3.3.3. Spectroscopic studies on AZATRUX

Because of the importance of the self-agreggation properties of the ligands on their interections with DNA, first of all we studied the behaviour of AZATRUX in aqueous solution by several spectroscopic techniques (NMR, UV and fluorescence spectroscopy). In order to study the affinity and selectivity of this compound towards G-quadruplex and duplex DNA, we performed experiments by some of the same techniques in presence of different DNA structures.

3.3.3.1. NMR spectroscopy

Since drug self-aggregation has been shown to be related to a lower telomerase inhibition and weaker interactions with the G-quadruplex with respect to unstacked molecules, the self-aggregation in aqueous solution of G-quadruplex ligands must be carefully considered in the model of their interaction with G-quadruplex DNA. For this reason, we have studied the self association of AZATRUX by NMR spectroscopy.

The high resolution spectra (fig. 43) and in particular the high resolution of the aromatic region of 1H NMR spectra in D_2O of AZATRUX (24a) at the high concentration of 10 mM at 23°C suggest the absence of aggregation. Moreover, increasing the temperature no improvement of spectra resolution was observed, as we expected for aggregation states (fig. 44). On the contrary we observed a normal worsening of the resolution due to increasing kinetic motions. These data suggest a poor tendency of AZATRUX to aggregate in water solution at high concentration (10 mM).

DOSY experiment:

NMR diffusion-ordered spectroscopy (DOSY) provide a way to separate the different compounds in a mixture based on the differing translation diffusion coefficients (and therefore differences in the size and shape of the molecule, as well as physical properties of the surrounding environment such as viscosity, temperature, etc) of each chemical species in solution. The measurement of diffusion is carried out by observing the attenuation of the NMR signals during a pulsed field gradient experiment. The degree of attenuation is a function of the magnetic gradient pulse amplitude (G) and occurs at a rate proportional to the diffusion coefficient (D) of the molecule.

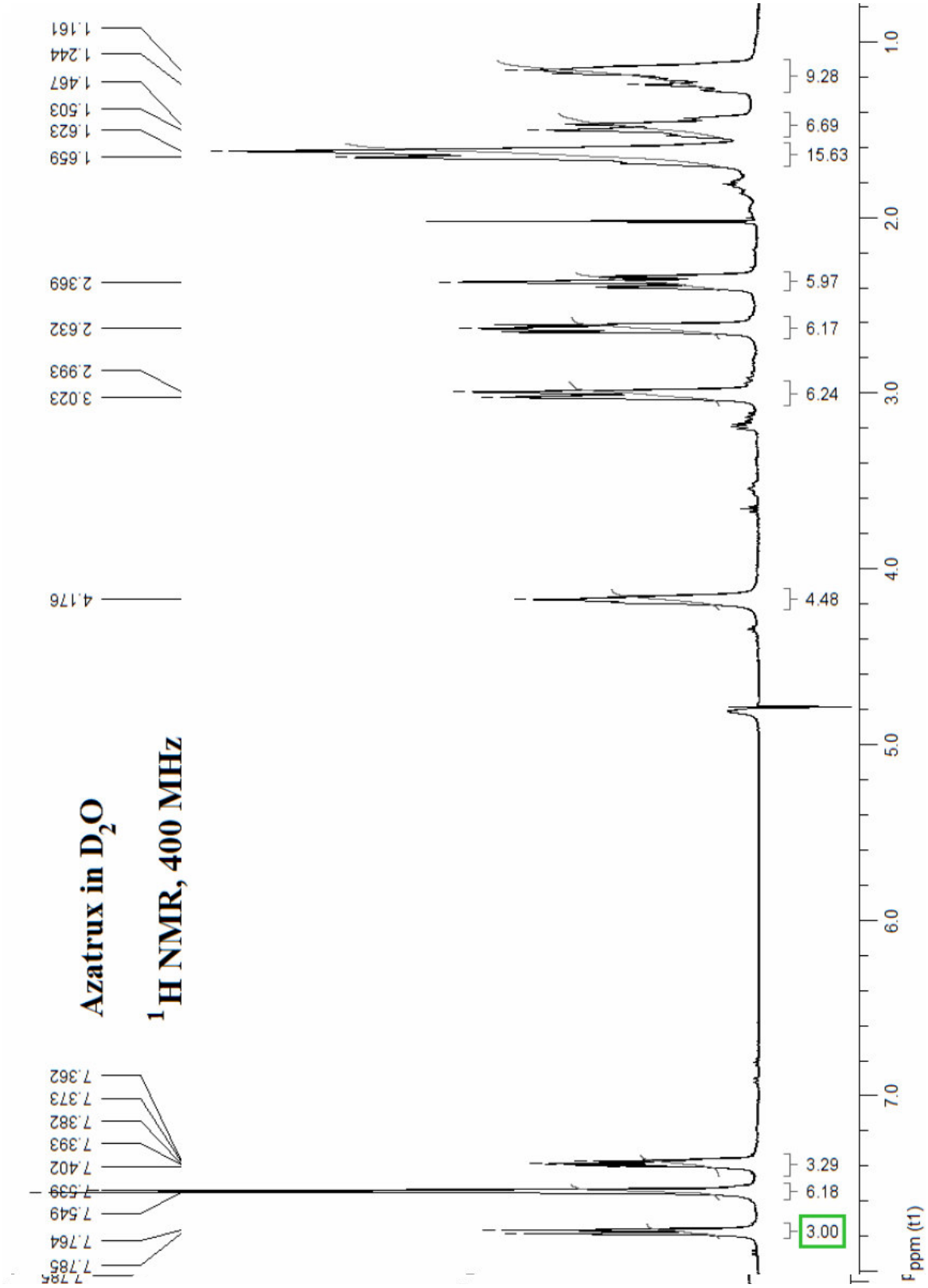


Figure 40: ^{1}H-NMR spectrum of AZATRUX (**24a**) as a hydrochloride in D_2O at 400 MHz.

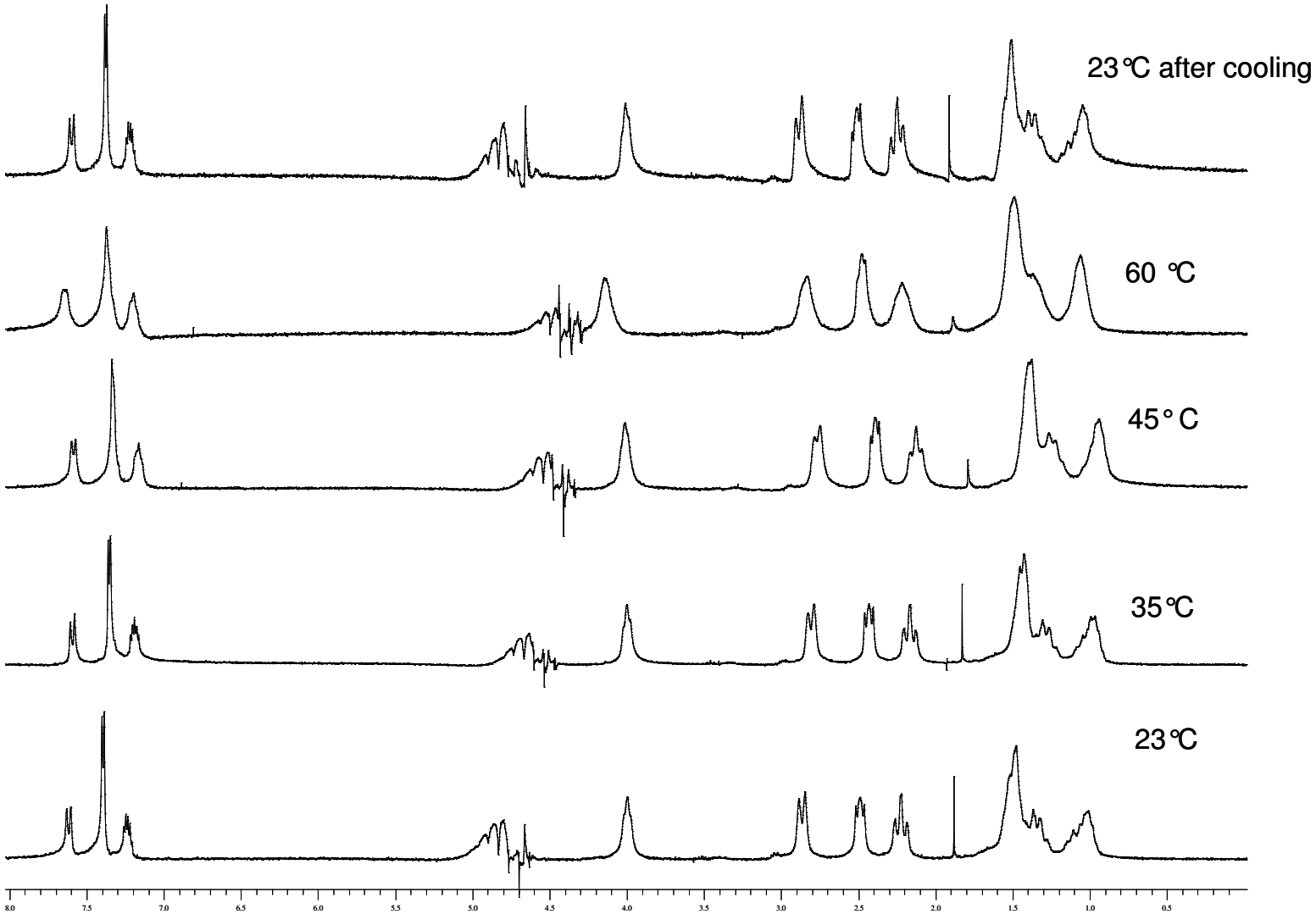


Figure 41: ^{1}H-NMR spectrum of AZATRUX (**24a**) in D_2O at different temperatures.

This no-invasive technique can be used to look the aggregation state of a sample (Agrawal *et al.* **2009**) monomers, dimers and polymers can be distinguish by the their different diffusion coefficients. In particular the experiment carried out on AZATRUX (**24a**) (Franceschin *et al.* **2009**) at 10 mM concentration in D_2O, at 25°C, clearly indicates that only one molecular species is present with a diffusion coefficient of D = 1, 86 X 10^{-10} m^2/s (log D= -9.73) D_2O (fig. 45). This data is calculated using as reference the HDO diffusion coefficient in D_2O at 298 K has a value of D = 1, 90 X 10^{-9} m^2/s (Longsworth, **1960**).

Due to the high resolution of the 1H-NMR spectrum in water (fig. 43), its behaviour at the increasing temperatures (fig. 44) and the diffusional properties noticed in DOSY experiment (fig.45) suggest that AZATRUX showed a poor tendency to aggregate in water solution (Franceschin *et al.* **2009**), despite the large aromatic core.

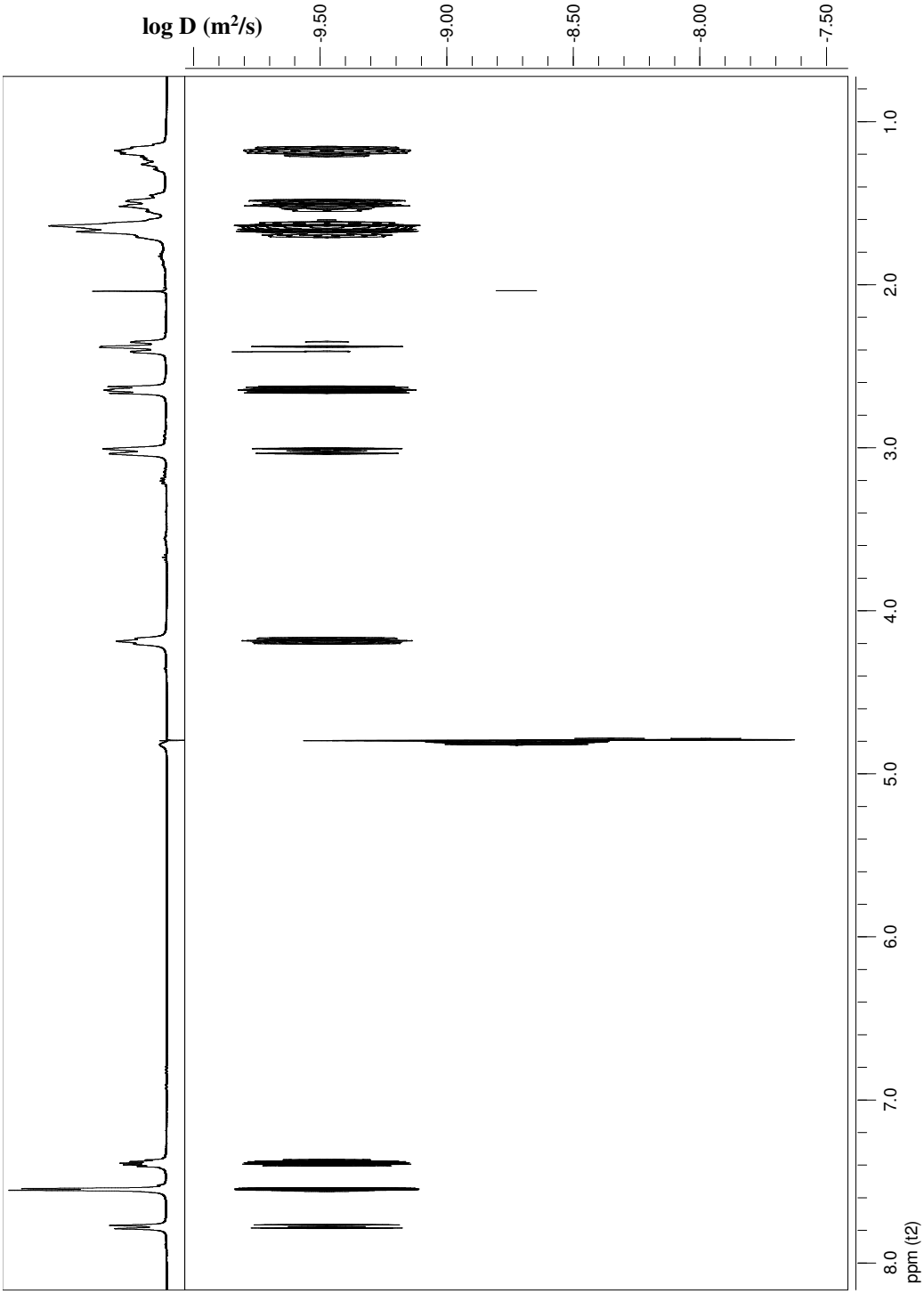


Figure 42: ^{1}H-NMR spectrum of AZATRUX (**24a**) in D_2O (left) and corresponding DOSY (Diffusion-ordered spectroscopy) experiment (right) (log D vs. chemical shift).

3.3.3.2. UV-vis absorption spectroscopy.

The poor variability of the UV spectrum appearance when increasing temperature (fig. 43) confirmed the scarce tendency of AZATRUX to aggregate in water solution (Ginnari-Satriani *et al.* **2009**) also at micromolar concentration. A further confirmation of this behaviour has been obtained by the observed linearity of absorbance versus concentration (Franceschin *et al.* **2009**) (fig. 44), both in MES/KCl buffer (pH 6.5) and in aqueous solution of HCl 0.1M. Since at this very low pH we can be reasonably sure to break any possible aggregate, we can conclude that no aggregates are present also at the physiological pH.

Since these results combined, also, with the previous NMR experiments it is reasonable to conclude that the unique species in solution, both at micromolar and at millimolar range concentration, is a monomeric form of AZATRUX (**24a**).

This is not a common behaviour, even though observed for instance for multiple-side-chained perylene derivatives (Alvino *et al.* **2007**; Franceschin *et al.* **2007**), since a large aromatic core is often related to a strong tendency of selfaggregation in water, because of hydrophobic interactions leading to stacking of several molecules one upon the other (Rossetti *et al.* **2005**; Franceschin *et al.* **2007**). It is worth noting that very recently self-association of perylene based G-quadruplex ligands has been reported to favour the binding to duplex DNA, compromising the specific recognition of Gquadruplex structures (Micheli *et al.* **2009**; Franceschin *et al.* **2008**). This could explain the good selectivity shown by AZATRUX for the G-quadruplex with respect to duplex DNA (Ginnari-Satriani *et al*. **2009**) showed in ESI-MS experiments.

UV titrations: study of selective interactions for G-quadruplex DNA with respect to duplex DNA .

As an independent confirmation of this quadruplex/duplex selectivity, I have performed titrations of the human telomeric F21T oligonucleotide and the biologically relevant CT DNA with AZATRUX by UV spectroscopy (fig. 47) in more relevant biological conditions with respect to ESI-MS experiment condition (in K^+ solutions at physiological concentrations) (Ginnari-Satriani *et al.* **2009**). In the UV titrations spectra showed in fig. and fig. a hypochromic effect can be clearly noticed in the maximum absorbiption of AZATRUX (**24a**), when increasing R (the molar DNA/drug ratio), at constant drug concentration (10μM). This behaviour is observed both with G-quadruplex (F21T) (fig. 45) and genomic duplex DNA (CT) (fig. 46), but to a different extent. This suggests a different affinity in binding the two DNA structures. Unfortunately a reliable quantitative analysis of the two titrations couldn't be performed because of the incresing importance of DNA absorbance with respect to the AZATRUX spectrum. This can be observed from R=2 at low

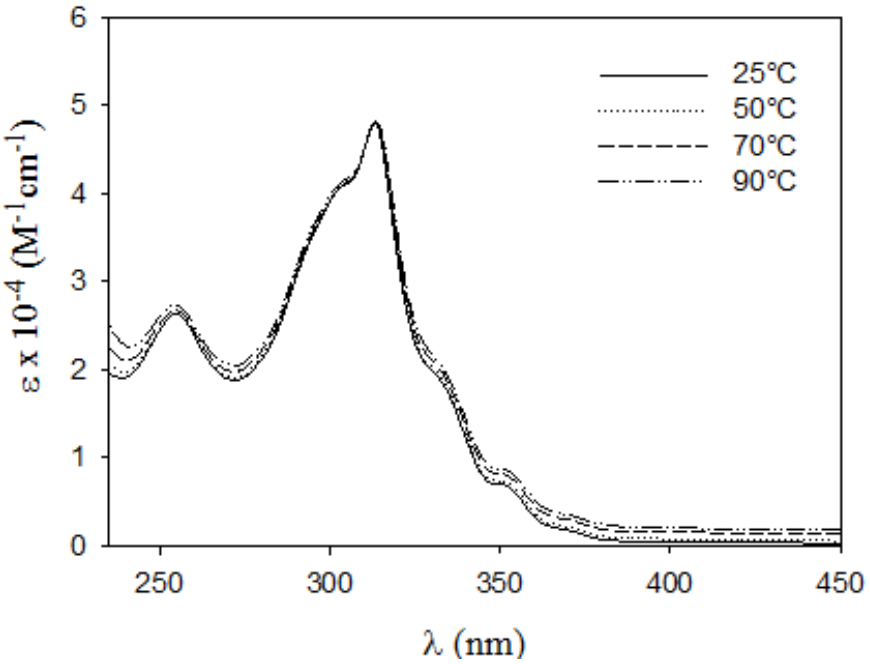


Fig. 43 UV-vis absorption spectra of AZATRUX in MES-KCl aqueous buffer at different temperatures (Ginnari-Satriani *et al.* **2009**).

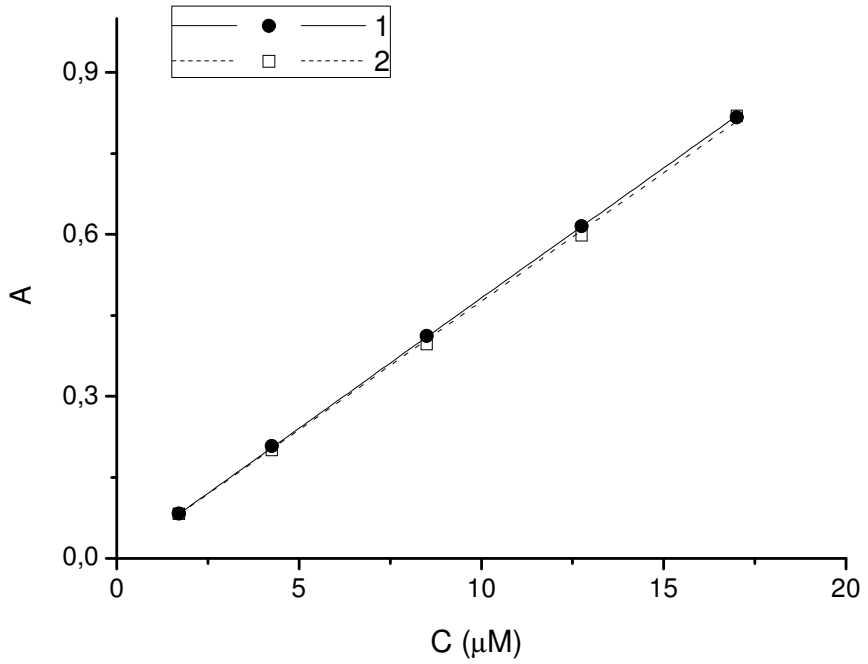


Figure 44. Absorbance versus concentration of AZATRUX (**7a**) in aq. HCl 0.1M (1) and MES/KCl buffer (2), showing linearity in the considered concentration range (Franceschin *et al.* **2009**).

wavelength, but this contribution becomes important at 313 nm only for R≥7 (F21T) (fig. 45) and R≥12 (ct) (fig. 46), anyway before saturation. In fact, at increasing DNA concentration and, consequently, its absorbiption contribution, a clear redshift of the absorbiption spectrum was observed, increasing also its intensity.

In order to evaluate in a qualitative manner the different appearence of the two titrations, I reported the normalized molar extinction coefficients (at 313 nm) vs. R (the molar DNA/drug ratio) (fig. 47), according to the equation: $\varepsilon N=\varepsilon_0-\varepsilon_R/\varepsilon_0-\varepsilon_f$ (where ε_0 is the ε at R=0, ε_R is the ε at a variable R and ε_f is the ε at final R used in the titration). So, I noticed (fig. 47) that the binding curve slope for quadruplex DNA is steeper than for duplex DNA and thus it tends to saturation at a lower ratio in the first case. This different behaviour of AZATRUX towards the two DNA structures confirms a stronger interaction for telomeric G-quadruplex with respect to genomic duplex DNA, also in these conditions (50 mM KCl solution) more similar to physiological conditions than those used in ESI-MS experiments.

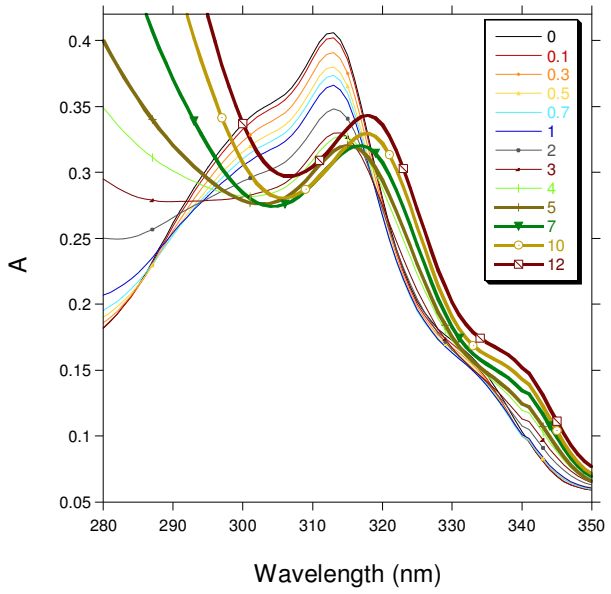


Fig. 45 UV-vis absorption spectra of AZATRUX in MES-KCl aqueous buffer in presence of F21T (G-quadruplex DNA) at different **R** (were **R** is the DNA/drug molar ratio).

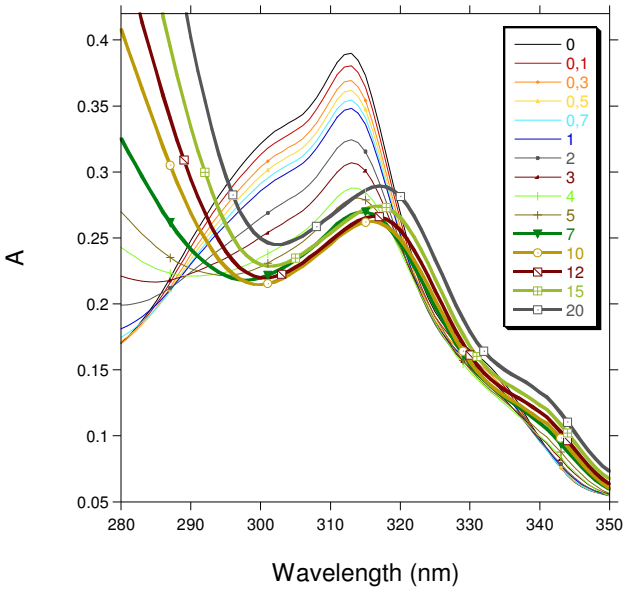


Fig. 46 UV-vis absorption spectra of AZATRUX in MES-KCl aqueous buffer in presence of genomic ct (duplex DNA) at different **R** (were **R** is the DNA/drug molar ratio).

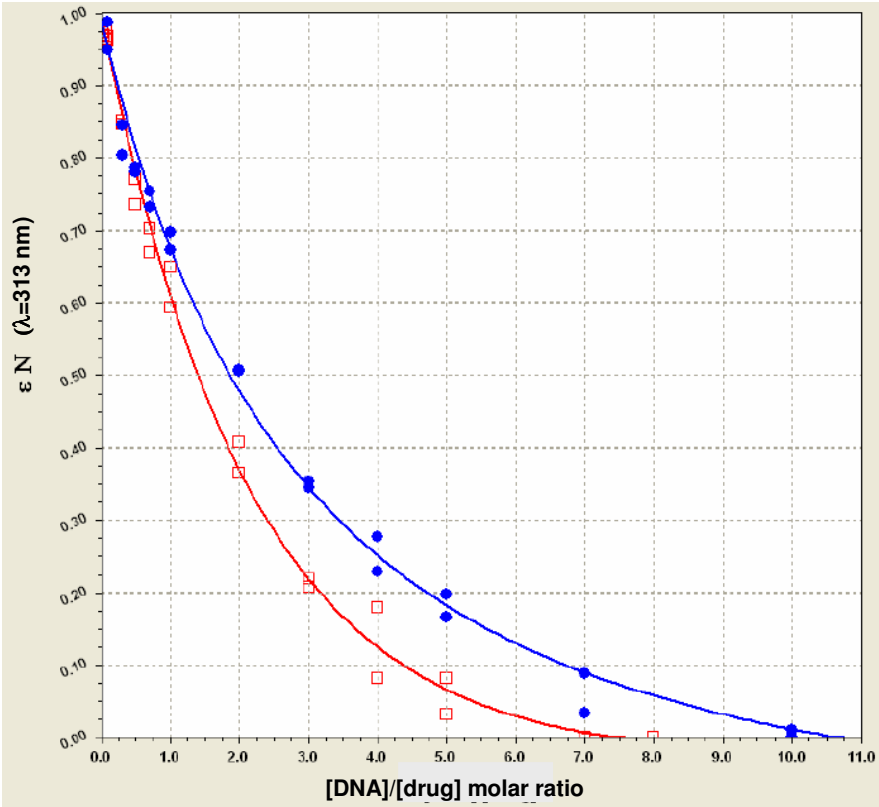


Figure 47
UV Titrations: Normalized molar extinction coefficient (εN) versus DNA/drug concentration ratio of AZATRUX (10µM) at 313nm with F21T oligo (□) and CT DNA (●).

3.3.3.3. Fluorescence emission spectroscopy.

Fluorescence properties of AZATRUX:

AZATRUX (**24a**) showed also very interesting fluorescence properties. In fact, the unsubstituted triazatruxene core (**21**) showed a very high fluorescence intensity in DMSO even at a low voltage (Fig. 48a), as reported also for several other lipophilic N-substituted derivatives in organic solvents (Franceschin *et al.* **2009**) . On the other side, no fluorescence was detected in water in the same conditions, while a high voltage (800V) was necessary to detect a weak fluorescence (fig. 48b). This is probably due to a very scarce solubility of this compound in water and/or a strong self-aggregation, as suggested by the strong hypochromic effect observed in the UV spectrum. On the contrary, the hydrophilic triazatruxene derivatives showed a good fluorescence both in DMSO and in water at a voltage of 500V (Fig. 49).

In fact, AZATRUX (**24a**) is fluorescent not only in organic solvent, as reported in the literature for other lipophilic derivatives, but also in water, making them particularly suitable for biological applications (Franceschin *et al.* **2009**).

Fluorescence titrations:study of selective interactions for G-quadruplex DNA with respect to duplex DNA .

Due to fluorecence properties of AZATRUX in aqueous solution, fluorescence spectroscopy looked also promising for a quantitive study of the interactions of AZATRUX with quadruplex and duplex DNA, since differently from UV-vis absorption spectroscopy (par.3.3.3.2) in this case there is no contribution of DNA to the emission spectrum.

Fluorescence titrations of AZATRUX with quadruplex and duplex DNA have been performed in collaboration with the group of prof. Giancola at the University of Naples “Federico II”. In the figure 50 the quenching of the emission signal of AZATRUX is reported upon interaction with wtTel24b oligo (5’-AGGGTTAGGGTTAGGGTTAGGGTT-3’) and an autocomplementary dodecamer, representing a model for quadruplex and duplex DNA respectively. The experimetal data (black sqaures) have been interpolated with an equation based on an equivalent binding sites model (red curve, fig. 50), obtaining the binding constants reported in Table 5.

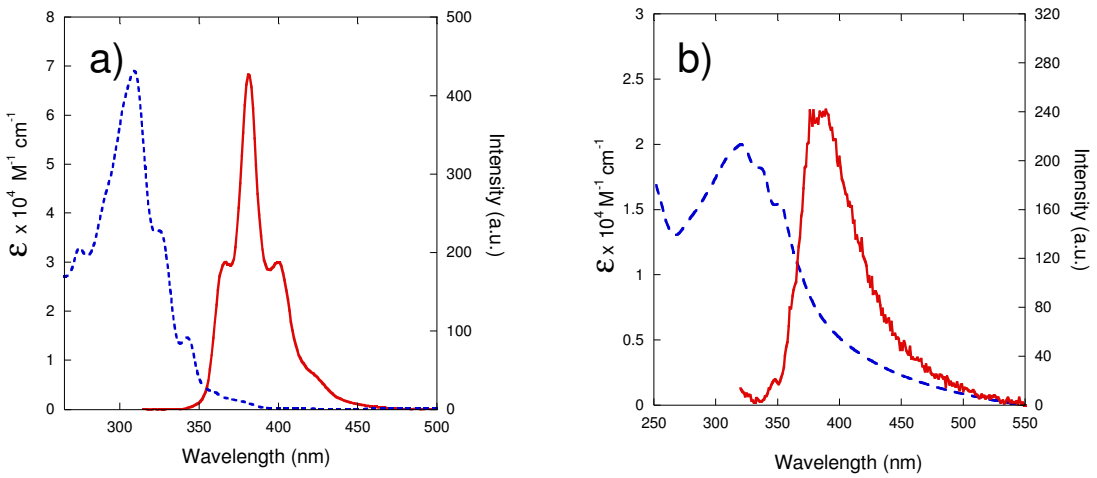


Fig.48 UV/vis absorption (dotted line) and emission (continuous line) spectra of triazatruxene core (**21**) in DMSO (a) and in queous MES/KCl buffer (b). For the fluorescence spectra the following conditions were used: compound concentration 10 µM, voltage 450V (a) and 800V (b), excitation wavelengths 309 nm (a) and 310 nm (b).

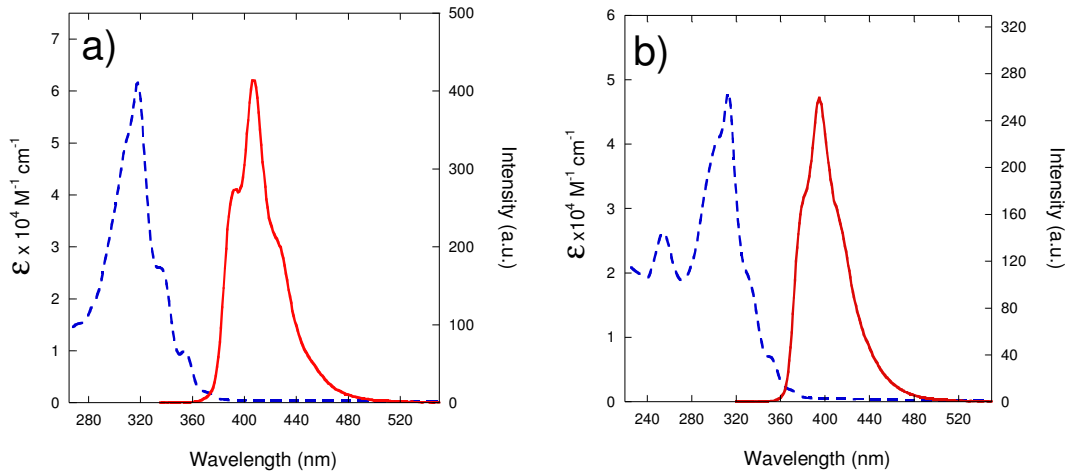


Fig. 49 UV/vis absorption (dotted line) and emission (continuous line) spectra of AZATRUX (**24a**) in DMSO (a) and in aqueous MES/KCl buffer (b). For the fluorescence spectra the following conditions were used: compound concentration 10µM, voltage 500V (b, c), excitation wavelengths 318 nm (a) and 310 nm (b).

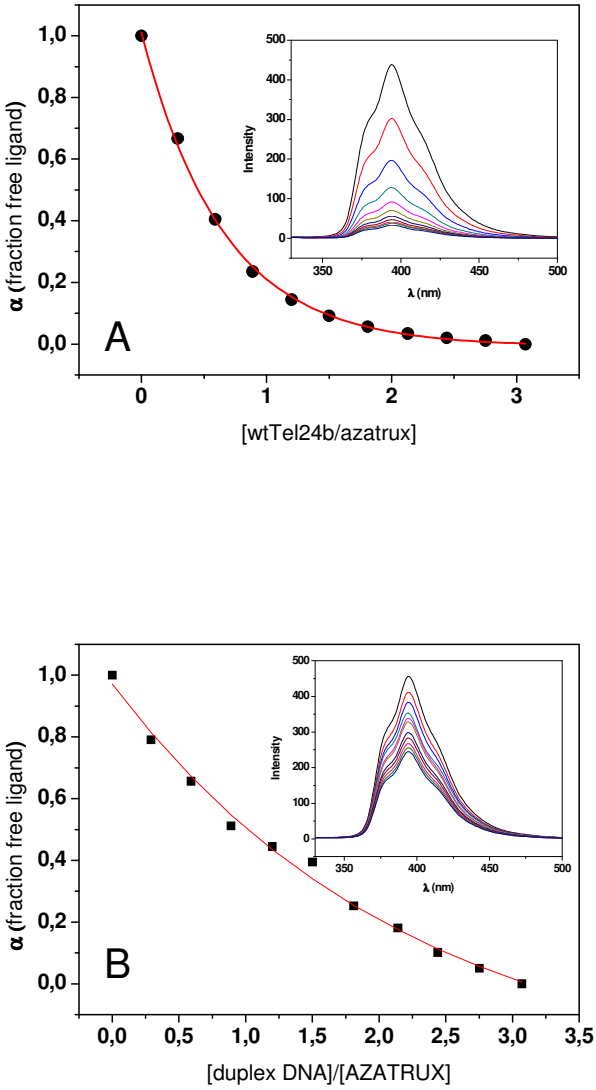


Fig. 50

• Fluorescence emission spectra of AZATRUX in aqueous solution in presence of wtTel24b (G-quadruplex DNA, **A**) and autocomplementary dodecamer (duplex DNA, **B**) at different **R** (were **R** is the molar DNA/drug ratio).

• Fluorescence titrations: fraction free ligand (α) versus molar DNA/drug ratio (**R**) with wtTel24b (●, **A**) and duplex DNA (■, **B**).

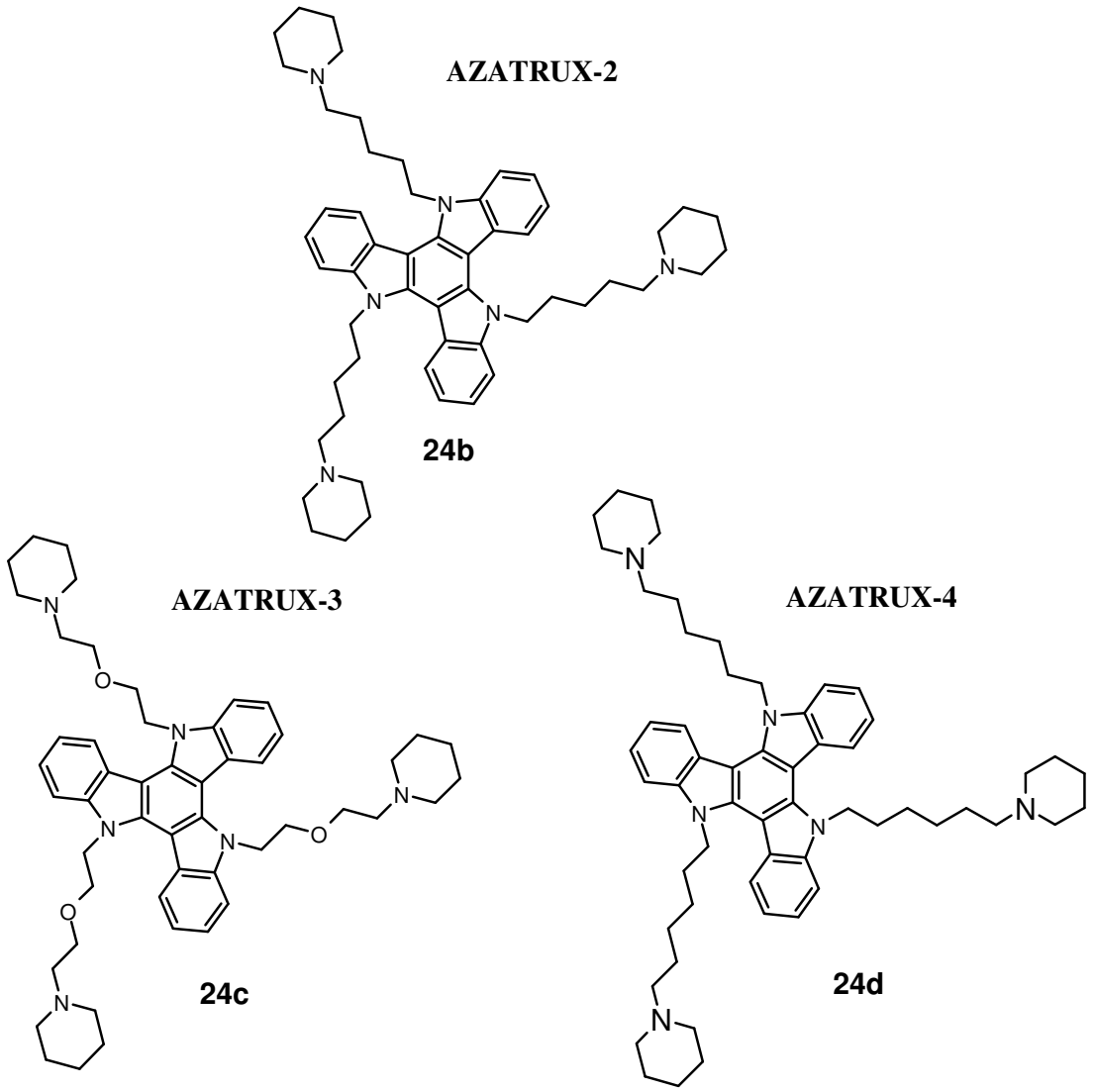
Oligonucleotides	K
wtTel24b (G4)	$5X10^6$
duplex	$5X10^4$

Tab.5 Binding costants for the complexes between AZATRUX and the indicated oligonucleotides as derived by elaboration of fluorescence titrations.

The difference of two orders of magnitude in K values confirms the selectivity of AZATRUX for the G-quadruplex with respect of duplex DNA. On the other hand, a more detailed comparison with binding costants derived from different techniques, such as ESI-MS, is not possible because of the different experimental conditions, in particular the different cations.

3.3.4. Synthesis of a series of tris-N-substituted triazatruxene derivatives

Bypassing the problem of the limited commercial availability of halogeno-amines, we decided to apply the versatile method described above for the preparation of AZATRUX (**24a**) (par. 3.3.1.2.; 3.3.1.3.) to synthesize several compounds of this series, with different chains in length and bascicity: AZATRUX2-4 (**24b-d**) (Scheme 7) (Franceschin *et al.* **2009**) .

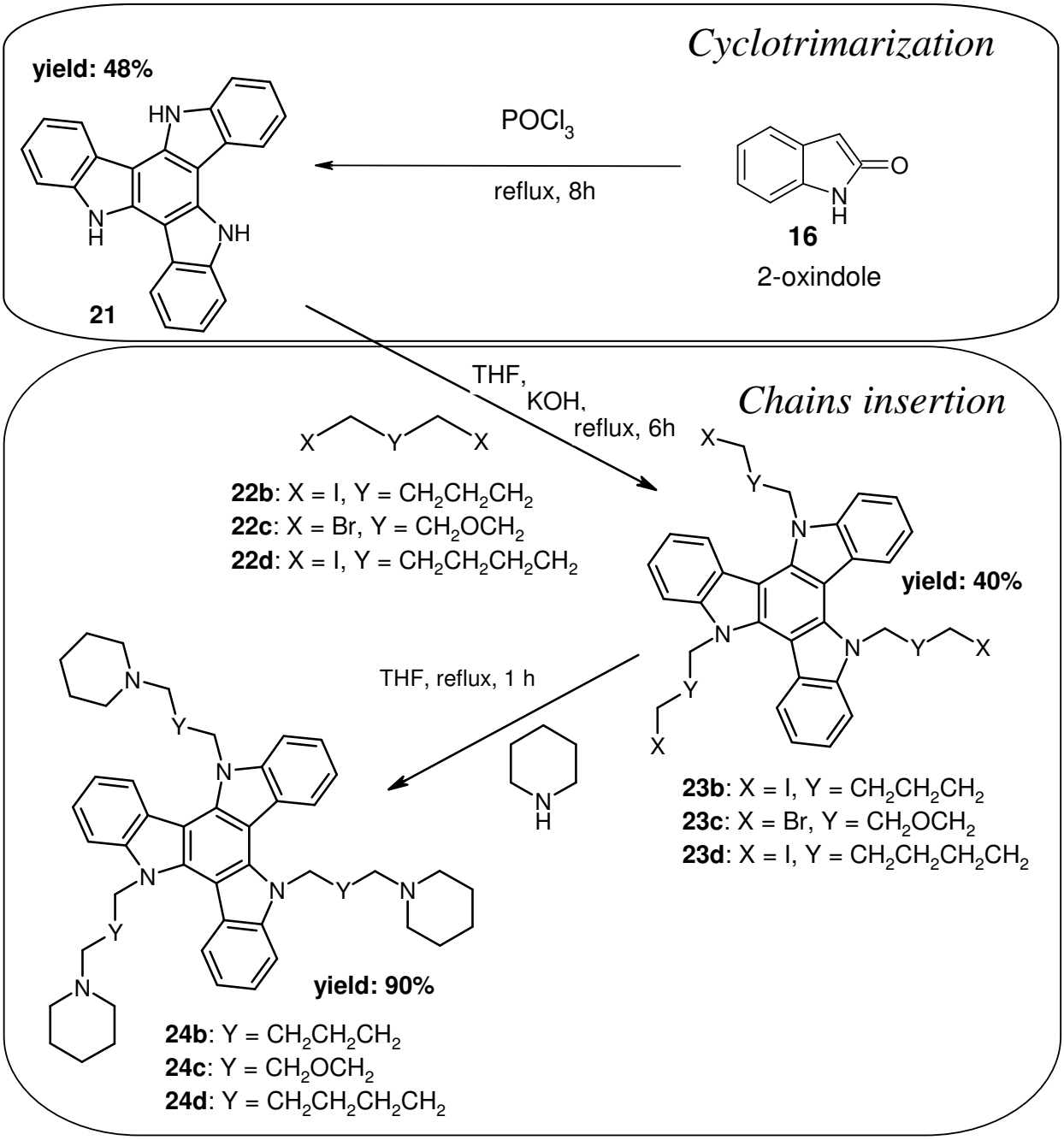


Scheme 7 Triazatruxene derivatives synthesized.

The synthetic pathway described above include the cyclotrimerization of 2-indolone (**16**) in $POCl_3$ and the insertion of side chains on the triazatruxene core (**21**) in two steps (Scheme 8):

a) The N-alkylation in basic conditions (KOH) with an excess of the appropriate dihalide (**22b-d**) to give the corresponding tris-N-substituted derivative (**23b-d**).

b) The following substitution of the halogen at the end of the chain with an excess of piperidine to get the final compounds (**24b-d**).



Scheme 8

The four synthesized compounds (**AZATRUX 1-4**, **24a-d**) represent just few examples of many hydrophilic triazatruxene derivatives, which can be easily prepared through this synthetic pathway. In fact, length and basicity of the side chains can be efficiently modulated choosing appropriate and, if desired, substituted dihalides (such as **24c**), as well as different amines for the final substitution of the remaining halogen (Franceschin *et al.* **2009**).

3.3.5. FRET assays of the triazatruxene derivatives on quadruplex and duplex DNA.

Fluorescence resonance energy transfer (FRET) is a distance-dependent interaction between the electronic excited states of two dye molecules in which excitation is transferred from a donor molecule to an acceptor molecule *without emission of a photon* (51a). The efficiency of FRET is dependent on the inverse sixth power of the intermolecular separation, making it useful over distances comparable with the dimensions of biological macromolecules. Thus, FRET is an important technique for investigating a variety of biological phenomena that produce changes in molecular proximity (Cantor *et al.* **1980**).

In particular, FRET can be used to study the stabilisation of G-quadruplex DNA was assessed using a modified fluorescence resonance energy transfer (FRET) between a donor (fluorescein) and an acceptor (tetramethylrhodamine) covalently attached to the 5' and 3' ends of the DNA, respectively, that form an acceptor/donor system. This is a method to measure the melting temperature of the intramolecular DNA G-quadruplex reported in the first time by Mergny and coworkers (Mergny *et al.* **2001**, Koeppel *et al.* **2001**). In short, the dual-labelled DNA probe is designed to fold into an intramolecular quadruplex at low temperatures, thus allowing resonance-energy transfer to occur. As the temperature is increased, the fluorescence donor–acceptor distance increases as the DNA undergoes a transition from an ordered quadruplex structure to a random-coil, thus emitting a detectable signal from the donor fluorophore and enabling the temperature for the transition to be measured (the melting temperature, ***T*m**) (fig.51b). The T**m** (*T*1/2) temperature is defined as the temperature for which the normalized emission is 0.5 (emission of Fluorescein is normalized between 0 and 1).

In a fret experiment are measured different melting temperatures keeping the DNA concentration constant and increasing the drug concentration, starting with a melting of DNA in drug absence in order to obtain the change melting temperature (Δ***T*m**). The change in melting temperature (Δ***T*m**) can be studied in function of the drug concentration (or R, where R represents the drug/DNA ratio defined as the ratio between the molar drugs concentration and the molar DNA concentration, expressed in oligonucleotides molecules) in order to compare in a qualitative manner the quadruplex-stabilising ability of the ligands in defined series (Guyen *et al.* **2004**).

Mreover, in order to study by FRET the interaction of the drugs with duplex DNA the same two fluorescent probes used in the quadruplex model, are linked in appropriate manner to the ends of an autocomplementary sequence, able to fold in a duplex structure. So, are obtained a quickly method

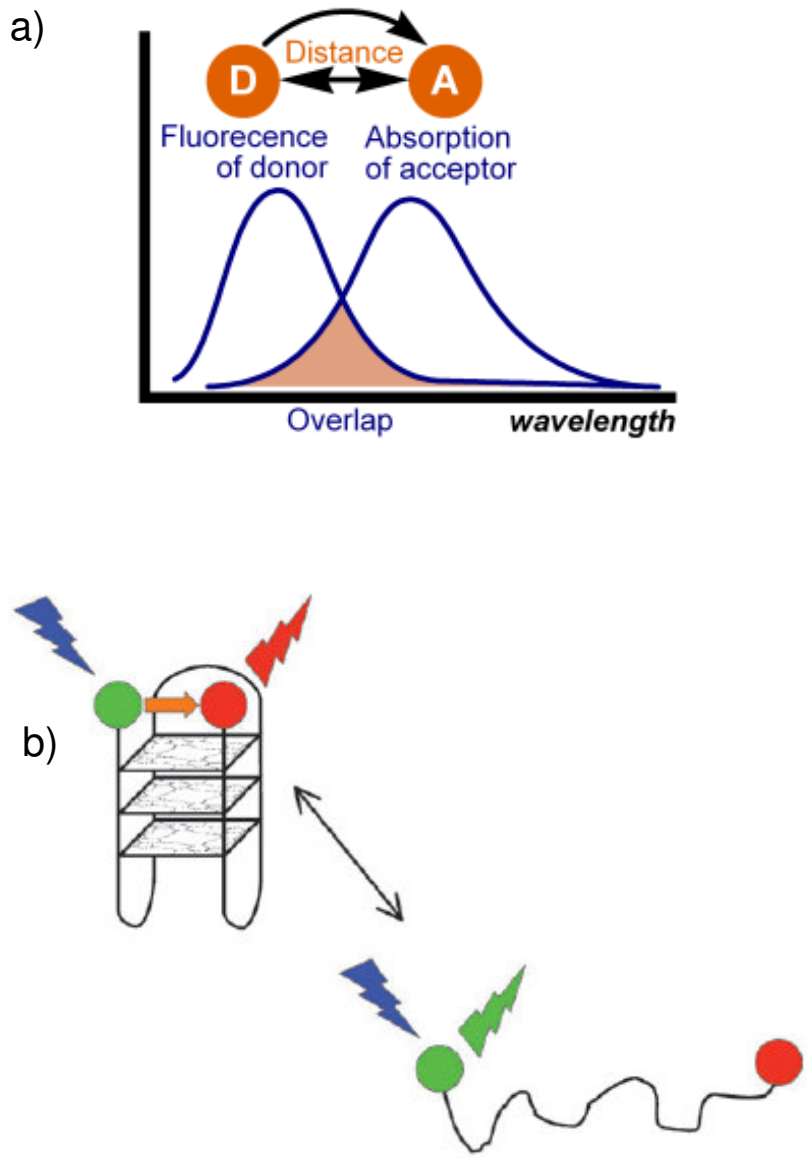


Fig. 51
a) Schematic representation of the superimposing of the absorption spectrum of the acceptor and the fluorescence emission spectrum of the donor;
b) Fluorescence resonance energy transfer assay. Two dyes are used: when the two dyes are close, energy may be transferred from the excited dye to the other dye, which then re-emits at a longer wavelength. The amount of direct emission versus FRET emission indicates how close the dyes are.

to evaluate the selectivity of the analysed drugs for G-quadruplex DNA with respect to duplex DNA (Guyen *et al.* **2004**).

In order to study and compare the affinity of the synthesized triazatruxene derivatives (**24a-d**) for the G-quadruplex structure and the selectivity with respect to duplex DNA, I carried out FRET assays on these compounds in the laboratories of the Professor Stephen Neidle (Cancer Research UK Biomolecular Structure Group, School Pharmacy of London).

As described in the experimental section, in these experiments I used two labelled oligonucleotides: F21T, which is a 21-mer human telomeric sequence (5'-FAM-dGGG(TTAGGG)$_3$-TAMRA-3', par.5.2.5., Mergny *et al.* **2001**) able to fold into an intramolecular G-quadruplex structure, and F10D, 5'-FAM-dTATAGCTATA-(CH_2-CH_2-O)$_6$-dTATAGCTATA-TAMRA-3', able to fold in a duplex DNA structure. The experimental conditions are identical to those used for the crystallization of the human telomeric sequence [AGGG(TTAGGG)$_3$], whose structure was resolved as a monomeric G-quadruplex structure in the presence of K^+ approximately physiological ionic conditions (Parkinson *et al.* **2002**).

The G-quadruplex melting curves for all the four compounds show a similar behaviour (fig. 52), as for the changing in shape with the increasing of drug concentration. In fact, when the drug concentration is ≤ 1 μM (R=5) the melting curves appear as regular sigmoids with one inflection point which represents the melting point. Instead, when the drug concentration is > 1 μM, melting curves change gently shape resulting a combination of two sigmoids with two distinct steps, characterized by two very different melting temperatures: ≈58 and ≈72°C. This suggests two distinct melting steps: the first step could be attributed to the threading intercalation mechanism proposed by Hurley and co-workers for perylene diimides (Fedoroff *et al.* **1998**), which assume that the ligand is initially stacked onto the ends of the G-quartet system. When the drug-DNA ratio R is ≤ 5 only this step is observed. Instead, increasing this ratio up to 25, a second melting step appears: this could be due to additional interactions of stacked drug with the hydrophobic grooves inside, where the drug could easily assume stacked forms. The stacking should significantly contribute to the stabilization of the DNA-drug complexes, so that the melting temperature is shifted to higher values.

In order to evaluate the selectivity for G-quadruplex DNA with respect to duplex DNA of the synthesized triazatruxene derivatives and compare the different behaviour of the four compounds,

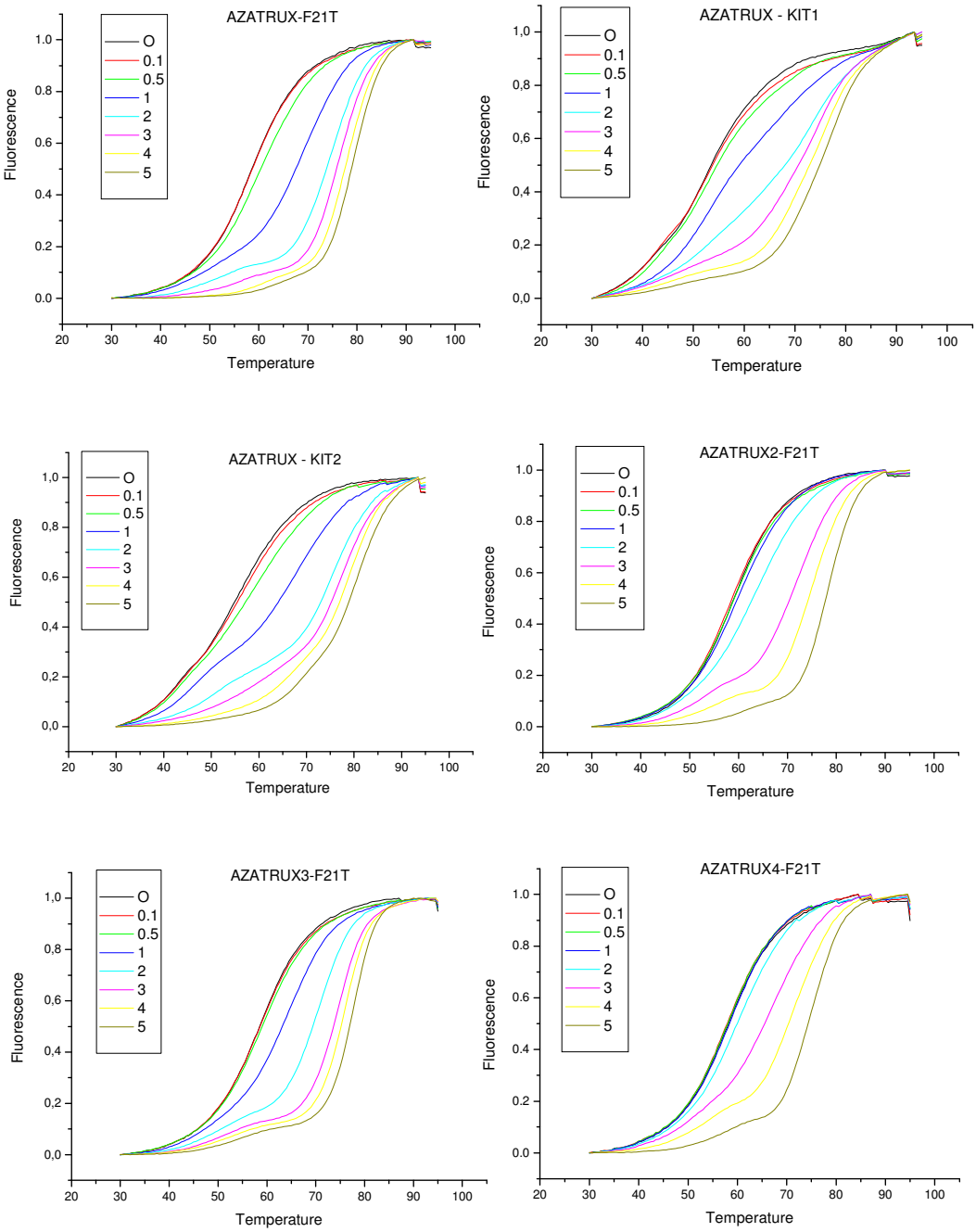


Fig. 52 Curves obtained by FRET experiments on F21T (intramolecular telomeric G-quadruplex) in presence of the triazatruxene derivatives AZATRUX (**24a**), AZATRUX-2 (**24b**), AZATRUX-3 (**24c**) and AZATRUX-4 (**24d**). Legends show the drug concentration (μM). Fluorescence is obtained after normalization.

are reported in fig. 53, 54 and 55 the graphics of Δ*T***m** vs drugs concentration for the two DNA structures (F21T and F10D). The four synthesized triazatruxene derivatives have been found to interact with the folded quadruplex DNA showing a good selectivity with respect to duplex structure. In particular, AZATRUX (**24a**) is resulted the best of the series: in fact, at the concentration of 1 μM (R=5) (tab. 6) it increases the G-quadruplex melting temperature of 11.3 °C, while at the same concentration (R=5) increases the duplex melting temperature of only 0.4 °C (fig. 53 and 56, tab. 6). AZATRUX-2 (**24b**) provokes a moderate effect on the quadruplex stabilization at drug concentration of 2 μM (R=10) with Δ*T***m**=5.6 °C and an important effect at drug concentration of 3 μM (R=15) with Δ*T***m**=13.8°C. At these concentrations a stabilization of the duplex structure is not noticed (fig. 54). AZATRUX-3 (**24c**) is able to induce a moderate G-quadruplex stabilization even at a concentration of 1 μM with Δ*T***m**=5 °C, while provokes an important effect at a concentration of 2 μM (R=10) with Δ*T***m**=11.6 °C. Also in this case, at the same concentrations no effect on duplex structure could be noticed (fig. 55). AZATRUX-4 (**24d**), instead, causes an important quadruplex stabilization only at a concentration of 4 μM (R=20) with Δ*T***m**=9.3 °C without interactions with the duplex structure (fig. 55).

From the reported data (tab. 6; Fig. 56) these new hydrophilic three side chained triazatruxene derivatives result in potent and selective G-quadruplex ligands and in particular AZATRUX (**24a**) is the best G-quadruplex ligand of this series showing a good selectivity with respect to the duplex DNA in according with the previous experiments in vitro (par. 3.3.2. and 3.3.3.).

Moreover, in order to study the possible pharmacological effects of AZATRUX (**24a**) not only at telomeres but also with respect to the regulation of oncogenes (par. 1.2), we carried out FRET melting assays on two other different G-quadruplexes (*c-kit1* and *c-kit2*).

c-kit1 and *c-kit2* are two sequences identified in the proto-oncogene *c-kit* promoter upstream of the transcription start site that are capable of forming quadruplex structures (Hsu *et al.* **2009**). Functionally significant mutations and overexpression of the *c-kit* gene play a role in oncogenic transformation of certain cell types and are found in several highly malignant human cancers, notably gastrointestinal stromal tumors (GIST) (Yang *et al.* **2008**).

The resulting data show that AZATRUX (**24a**) is able to stabilize the human telomeric (F21T), the *c-kit1* and *c-kit2* quadruplexes with very similar Δ*T*m at the same drug concentation, except at 1 μM (R=5) where it is notable a significantly lower Δ*T*m value for the *c-kit1* quadruplex with respect to the other two quadruplex structures (fig. 53 and 57, tab. 7). These results suggest possible effects by AZATRUX (**24a**) both at the telomeres and on the regulation of oncogenes.

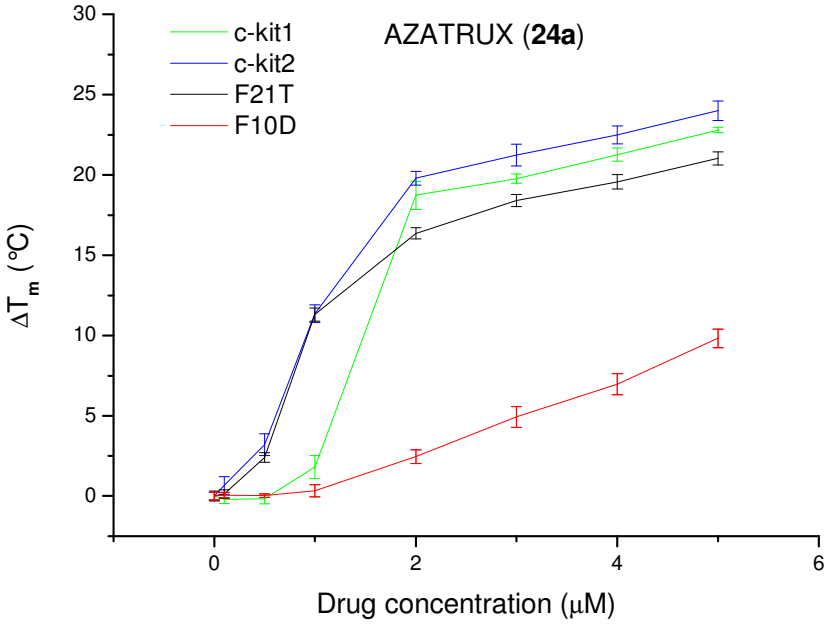


Fig. 53 Concentration dependency of the ΔTm. The stabilization (ΔTm) induced by AZATRUX (**24a**) is presented for F21T (telomeric G-quadruplex) with a black line, for c-kit1 and c-kit2 (*c-kit* G-quadruplexes) with green and blue lines, and for F10D (double-stranded DNA) with a red line.

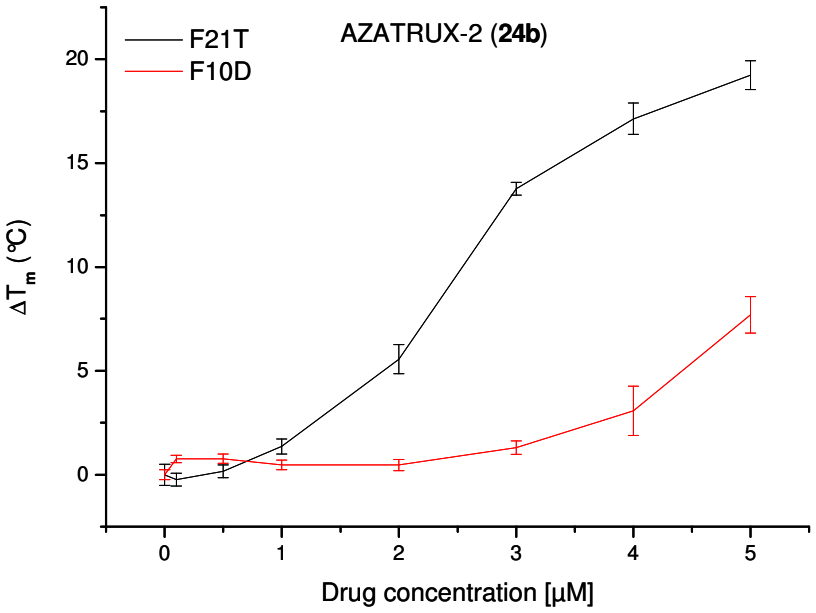


Fig. 54 Concentration dependency of the ΔTm. The stabilization (ΔTm) induced by AZATRUX 2 (**24b**) (B) is presented for F21T (intramolecular G-quadruplex) with a black line and for F10D (double-stranded DNA) with a red line.

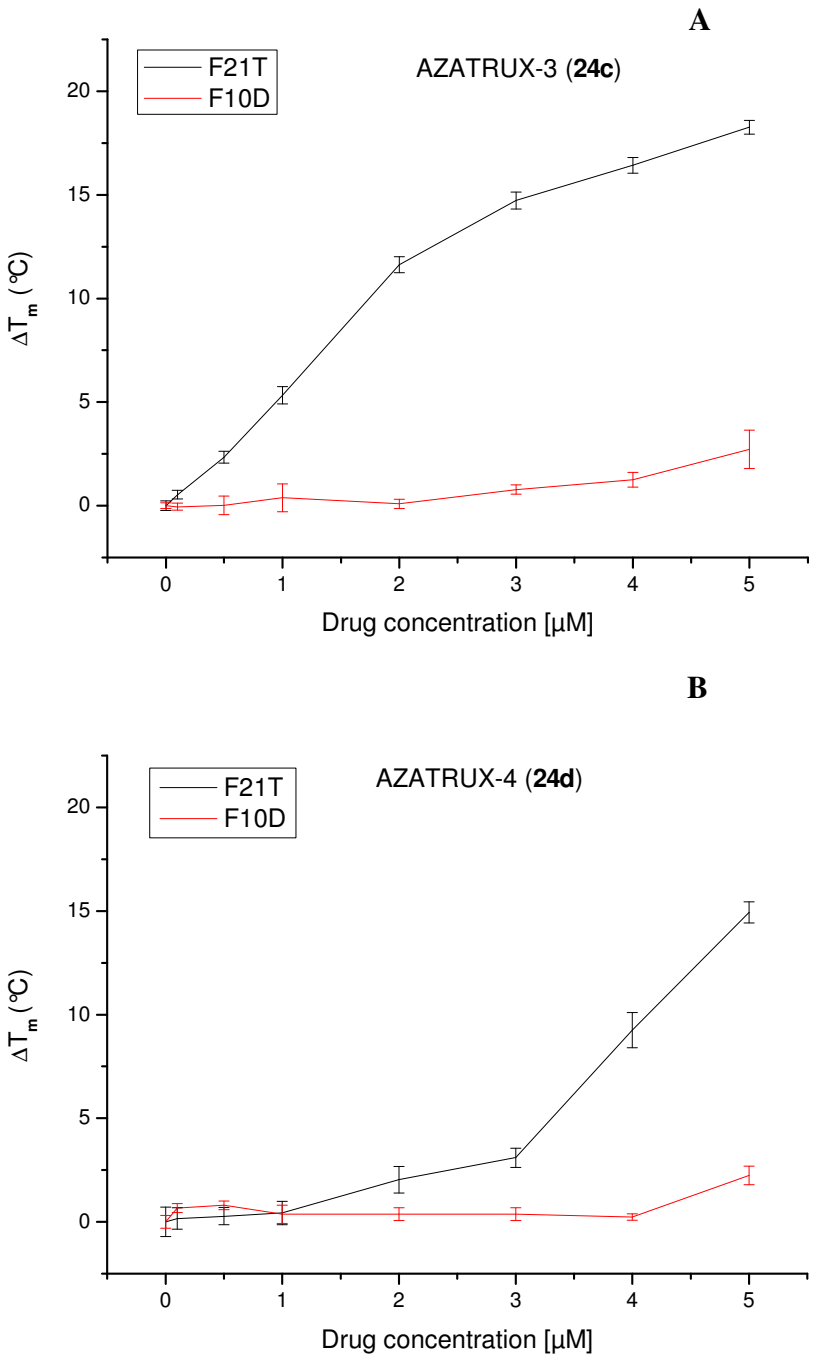


Fig. 55 Concentration dependency of the ΔTm. The stabilization (ΔTm) induced by AZATRUX 3 (**24c**) (A) or AZATRUX 4 (**24d**) (B) is presented for F21T (intramolecular G-quadruplex) with a black line and for F10D (double-stranded DNA) with a red line.

Drug conc.	1μM		2μM	
	F21T	F10D	F21T	F10D
AZATRUX	11.3±0.4	0.3±0.4	16.4±0.4	2.5±0.4
AZATRUX-2	1.4±0.4	0.5±0.2	5.6±0.7	0.5±0.3
AZATRUX-3	5.3±0.5	0.4±0.4	11.6±0.4	0.1±0.2
AZATRUX-4	0.4±0.6	0.4±0.4	2.0±0.6	0.4±0.3

Tab 6. Comparison of the DNA stabilization by triazatruxene derivatives AZATRUX (**24a**), AZATRUX-2 (**24b**), AZATRUX-3 (**24c**) and AZATRUX-4 (**24d**): ΔTm (°C) of different DNA (F21T and F10D) at different drug concentrations.

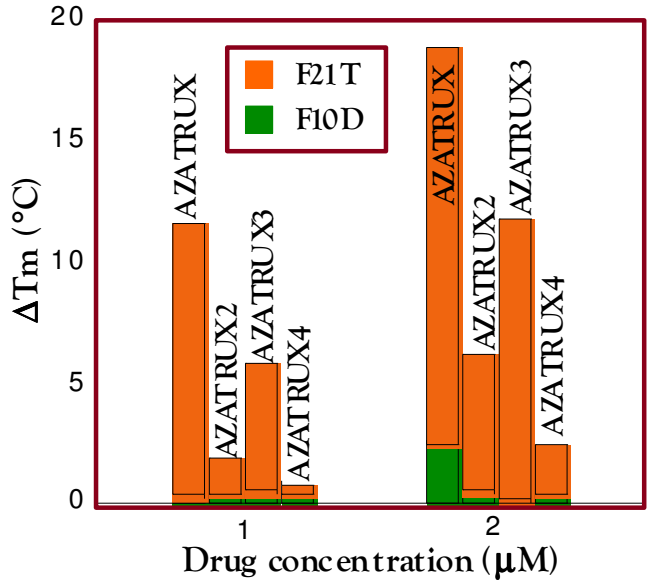


Fig. 56 Histogram of the ΔTm values of telomeric G-quadruplex (F21T) and duplex DNA (F10D) obtained from the FRET melting studies at different drug concentration (Table 5).

Drug conc. \ oligo	F21T	c-kit 1	c-kit 2
1 μM	11.3±0.4	1.8±0.7	11.7±0.6
2 μM	16.4±0.4	18.7±0.9	19.8±0.4
3 μM	18.4±0.4	19.8±0.3	21.2±0.7
4 μM	19.6±0.5	21.3±0.4	22.5±0.6
5 μM	21.0±0.4	22.8±0.2	24.0±0.6

Tab 7. G-quadruplex stabilization by AZATRUX (**24a**) on different G-rich oligonucleotides: ΔTm (°C) of different G-quadruplexes (G4-TEL "*F21T*", G4-kit "c-*kit1*" and "*c-kit2*") at different drug concentrations.

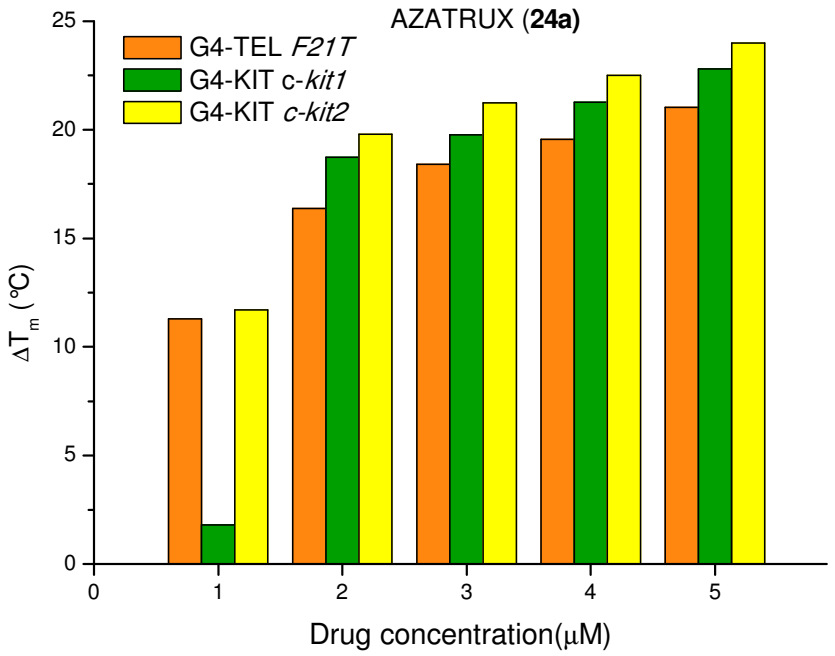


Fig. 57 Comparison of G-quadruplex stabilization by AZATRUX (**24a**): histogram of the ΔTm values of various G-quadruplexes (F21T, c-kit1 and c-kit2 sequences) obtained from the FRET melting studies at different drug concentration (Table 6).

3.3.6. Comparing studies for the triazatruxene derivatives towards telomeric G-quadruplex and duplex DNA by ESI-MS.

In order to evaluate the binding affinity of the complete series of triazatruxene derivatives (**24a-d**) and to compare these results with the FRET qualitative data we decided to perform ESI-MS experiments. We then calculated the binding costants values for both telomeric G-quadruplex and duplex DNA.

We couldn't choose the oligo F21T as a model of telomeric G-quadruplex, as previously done in the ESI-MS experiments (par. 3.3.2.) for AZATRUX. In fact the ESI-MS spectrum of this oligo in presence of AZATRUX-4 (**24d**) presents the peak relative to the $[2:1]^{5-}$ ion overlapped to the peak relative to the unbound $[F21T - 4H^+]^{4-}$ species (par. 5.2.2). This problem don't allow a reliable evaluation of the relative binding costants.

For this reason, we decided to use the oligo F21TTT (5'-GGGTTAGGGTTAGGGTTAGGGTT-3') as a telomeric G-quadruplex model. This oligo presents the same sequence of F21T and has two further nucleotides (TT) at the 3' end so that this DNA strand has a heavier mass with respect to F21T, thus allowing us to evaluate the binding costants for all the four triazatruxene derivatives (**24a-d**). The evaluation of the binding constants obtained by the collected data (Tab. 8) demonstrates that all the studied molecules are good telomeric G-quadruplex ligands, able to form both 1:1 and 2:1 drug–DNA complexes. AZATRUX-2, AZATRUX-3 and AZATRUX-4 show similar K_1 values, with log K_1 = 4.6, 4.6, and 4.7 respectively while AZATRUX shows the highest value of K_1 with log K_1 = 5.1 (Tab. 8).

In the examination of K_2 values AZATRUX, AZATRUX-2, AZATRUX-3 AZATRUX-4 show similar values of K_2 with log K_2 = 4.8, 4.5, 4.3 and 4.8, respectively (Tab. 8)

As previously reported for AZATRUX, in order to establish the selectivity of the other three triazatruxene derivatives synthesized, we have tested AZATRUX-2,4 (**24b-d**) with the self-complementary double-stranded DNA DK66 (Tab. 8). As previously described for AZATRUX (par. 3.3.2.), all of the three triazatruxene derivatives show spectra acquired at 1:1 ratios with no traces of 1:1 or 2:1 complex peaks, suggesting almost no interaction with double-stranded DNA. Also in these cases, in order to detect an appreciable peak of the 1:1 drug/DNA complex, these molecule must be present in the sample at 3:1 and 4:1 ratios, but even at this concentration, there is no evidence of the peak relative to the 2:1 complex. In particular AZATRUX, AZATRUX-3 and AZATRUX-4 show similar K1 values with log K_1= 4.1, 3.9 and 4.1 respectively, while AZATRUX-2 show a lower value of 3.4 (Tab. 8).

Comparing these data with the obtained data for the telomeric G-quadruplex, we can observe in all cases, except from AZATRUX-3, that K_1 values with DK66 are one order of magnitude lower than K_1 values with F21TTT, suggesting a good selectivity for these compounds in binding the telomeric G-quadruplex structure with respect to duplex DNA.

	F21TTT		**DK66**	
	Log K1	log K2	log K1	log K2
Azatrux (24a)	5,1±0,1	4,8±0,2	4,1±0,1	----------
Azatrux2 (24b)	4,6±0,1	4,5±0,2	3,4±0,2	
Azatrux3 (24c)	4,6±0,2	4,3±0,1	3,9±0,1	----------
Azatrux4 (24d)	4,7±0,1	4,8±0,1	4,1±0,1	----------

Tab. 8 K1 and K2 values on a logarithmic scale for the complexes between the synthesized triazatruxene (**24a-d**) derivatives and the indicated oligonucleotides as derived by ESI-MS experiments.

In conclusion, in according with the previously reported FRET results (par. 3.3.5.; Tab. 6), the ESI-MS data show that the four synthesized triazatruxene derivatives result in potent and selective G-quadruplex ligands. In particular AZATRUX result, again, the best G-quadruplex ligand of the series (Tab. 8).

For this reason we decided to test the anticancer activity of AZATRUX (**24a**) in cell biology studies in vitro (see below).

3.3.7. Cell biology studies on AZATRUX.

The good anticancer activity of AZATRUX (**24a**) in vitro was confirmed by extensive antitumor in vitro tests performed at the National Cancer Institute (NCI).

The NCI drug screening procedure (Monks *et al.* **1991**) utilize a panel of 60 human tumor cell lines organized into subpanels representing leukemia, melanoma, and cancers of the lung, colon, oviarian, renal, breast, prostate, ovary, and central nervous system (fig. 58, 59). Briefly, each cell line is inoculated onto 96-well microtiter plates, then preincubated for 24–28 hours. Subsequently, test agents are added in five 10-fold dilutions and the culture is incubated for an additional 48 hours. For each test agent, a dose-response profile is generated. End-point determinations of the cell viability or cell growth are performed by in situ fixation of cells, followed by staining with a protein-binding dye, sulforhodamine B (SRB assay) (Rubinstein *et al.* **1990**). The sulforhodamine B, that binds to the basic residues of proteins, is used to quantify the amount of living cells after short-term exposure to the test compound. The amount of dye per well, which can be quantified spectroscopically as function of the absorbance at 540 nm, is related to the total protein content and consequently to the number of cells.

The vitro antitumor activities is expressed as log GI50 (fig.60) where GI50, measures the growth inhibitory power of the test agent.

AZATRUX (**24a**), in particular, showed a good growth inhibitory activity for all 60 cancer cell lines coming from different tissues having the log GI50 in the range -5.98 to -6.29 with an exception of -8.12 for a SR leukemia cell line (fig. 60).

These biological results on the growth of several different tumour cell lines suggest an interesting anticancer activity of AZATRUX (**24a**) in vitro, although obviously at the moment it is not possible to correlate this finding to quadruplex binding. In the future it will be interesting to verify the correlation of anticancer activity with effects on telomeres and on specific oncogenes and to study more in detail the biology activity in terms of selectivity for the cancer cells with respect to normal cells.

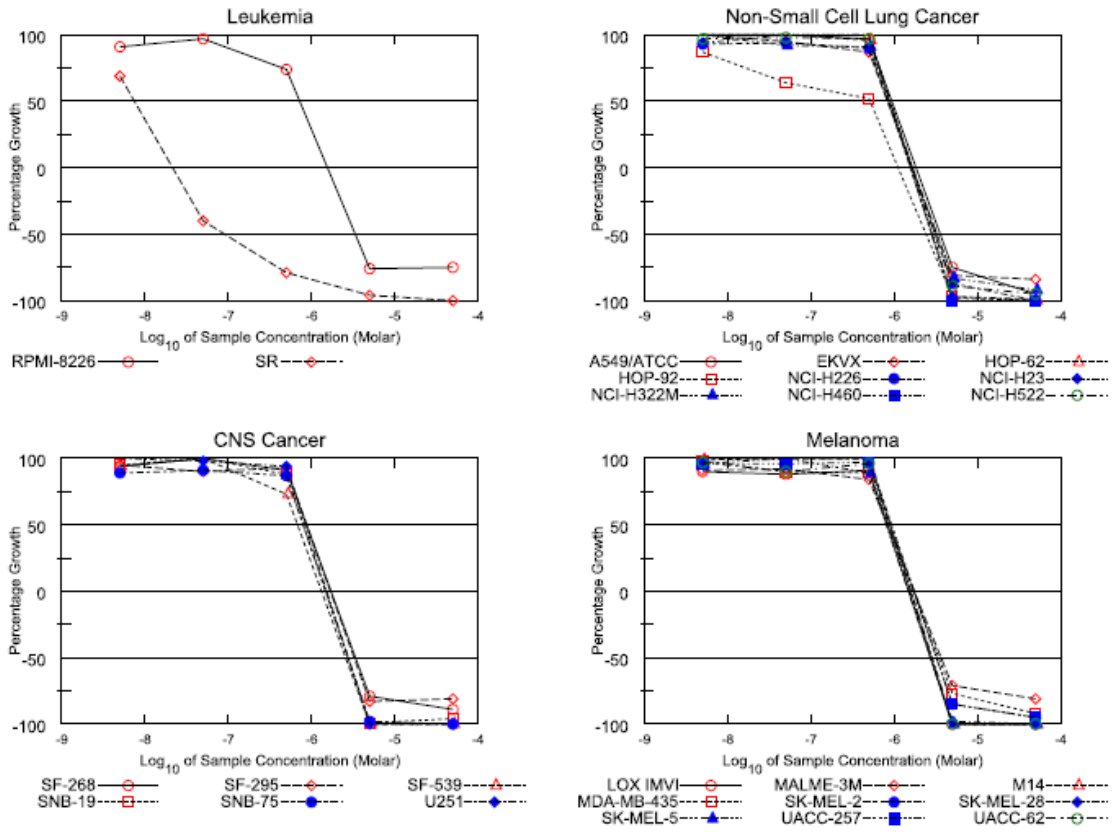


Fig. 58 Dose response curves of AZATRUX (**24a**) on cancer cell lines: leukemia, melanoma, non-small lung cell and central nervous system cancer cells.

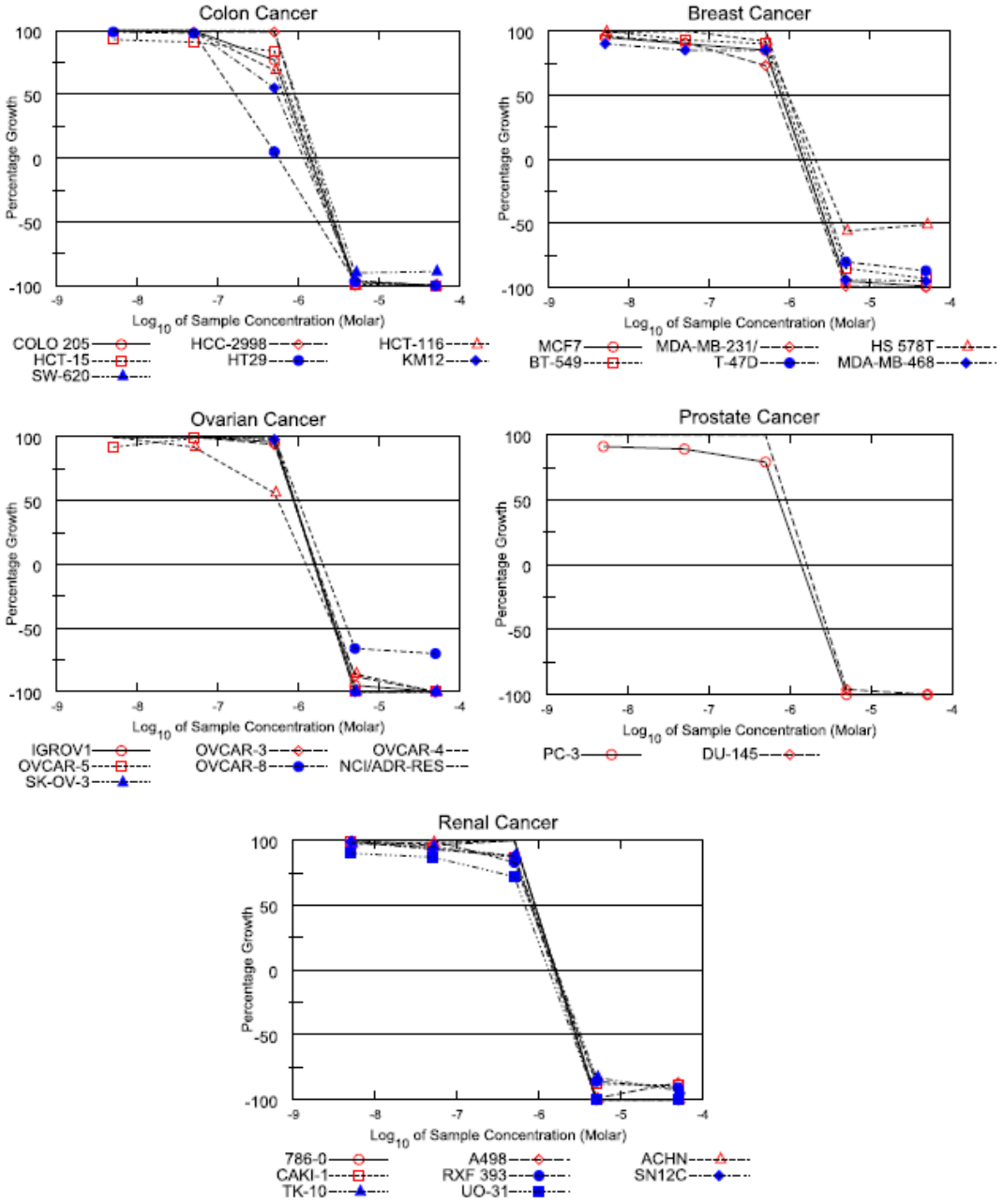


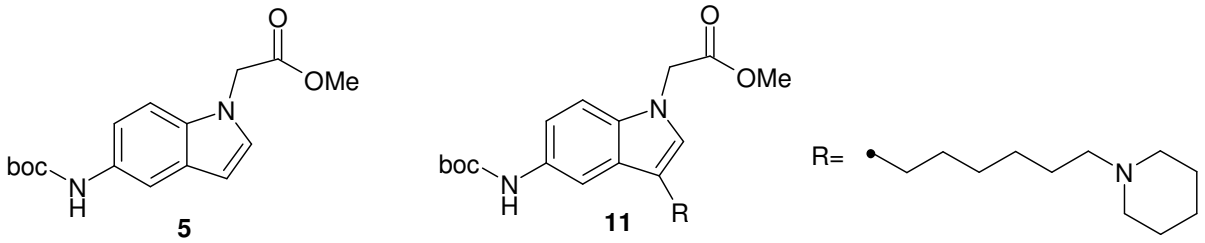
Fig. 59 Dose response curves of AZATRUX (**24a**) on cancer cell lines: colon, ovarian, breast, prostate and renal cancer cell lines.

Panel	Cell line name	$\log_{10}$ GI50
Leukemia	RPMI-8226	-6.14
Leukemia	SR	-8.12
Non-small cell lung cancer	A549/ATCC	-6.03
Non-small cell lung cancer	EKVX	-6.08
Non-small cell lung cancer	HOP-62	-6.07
Non-small cell lung cancer	HOP-92	-6.29
Non-small cell lung cancer	NCI-H226	-6.07
Non-small cell lung cancer	NCI-H23	-6.04
Non-small cell lung cancer	NCI-H322M	-6.06
Non-small cell lung cancer	NCI-H460	-6.05
Non-small cell lung cancer	NCI-H522	-6.05
Colon cancer	COLO 205	-6.15
Colon cancer	HCC-2998	-6.06
Colon cancer	HCT-116	-6.19
Colon cancer	HCT-15	-6.12
Colon cancer	HT-29	-6.79
Colon cancer	KM-12	-6.27
Colon cancer	SW-620	-6.03
CNS Cancer	SF-268	-6.06
CNS Cancer	SF-295	-6.06
CNS Cancer	SF-539	-6.17
CNS Cancer	SNB-19	-6.10
CNS Cancer	SNB-75	-6.10
CNS Cancer	U251	-6.08
Melanoma	LOX IMVI	-6.09
Melanoma	MALME-3M	-6.08
Melanoma	M14	-6.03
Melanoma	MDA-MB-435	-6.07
Melanoma	SK-MEL-2	-6.01
Melanoma	SK-MEL-28	-6.04
Melanoma	SK-MEL-5	-6.09
Melanoma	UACC-257	-6.04
Melanoma	UACC-62	-6.07
Ovarian cancer	IGROV1	-6.06
Ovarian cancer	OVCAR-3	-6.06
Ovarian cancer	OVCAR-4	-6.26
Ovarian cancer	OVCAR-5	-6.04
Ovarian cancer	OVCAR-8	-5.98
Ovarian cancer	NCI/ADR-RES	-6.06
Ovarian cancer	SK-OV-3	-6.01
Renal cancer	786-0	-6.05
Renal cancer	A498	-6.05
Renal cancer	ACHN	-6.05
Renal cancer	CAKI-1	-6.09
Renal cancer	RXF 393	-6.10
Renal cancer	SN12-C	-6.10
Renal cancer	TK-10	-6.08
Prostate cancer	UO-31	-6.17
Prostate cancer	PC-3	-6.14
Breast cancer	DU-145	-6.00
Breast cancer	MCF7	-6.11
Breast cancer	MDA-MB-231/ATCC	-6.17
Breast cancer	HS-578T	-6.02
Breast cancer	BT-549	-6.07
Breast cancer	T-47D	-6.01
Breast cancer	MDA-MB-468	-6.11

Fig. 60 NCI in vitro testing results:
GI50 is the concentration of test drug where 100 × (T - T0)/(C - T0) = 50 and can have values from +100 to -100. The optical density of the test well after a 48-h period of exposure to test drug is T, the optical density at time zero is T0, and the control optical density is C.

4. Conclusions

In this PhD thesis I presented two different indole-based oligomers as new G-quadruplex interactive compounds which I dealt with during my PhD work. In particular, my efforts have been directed towards two synthetic targets represented by: tetrameric macrocycles of indoles, where the monomers are linked by amidic bonds, and the triazatruxene core, a cyclotrimer of indoles were the three monomers are fused in a new aromatic core. For the first synthetic target I designed two new polyamido indole-based tetrameric macrocycles with the same core: one without chains, in order to create a simplified model of the G-tetrad, and one furnished of four hydrophilic chains in order to improve the interaction with the G-quadruplex structure following the proposed model of interaction discussed above (par. 1.4.). In this phase theoretical studies of molecular modeling have been carried out in order to design the two macrocycles. At this point I found the corresponding suitable starting monomers for a synthetic strategy inspired by peptides synthesis in order to obtain the two desired macrocycles. The two resulting monomers (**5** and **11**), which differ for the insertion of the hydrophilic chain in 3-indole position, present the carboxylic and aminic functions suitable arranged in order to obtain the correct geometry of the final macrocycles. Through two synthetic approach, which differed for the insertion of the hydrophilic side chain in 3- indole position through an Heck reaction, I obtained the two suitable monomers (**5**) and (**11**) with the carboxylic and aminic functions suitable arranged and protected (esther and NH-BOC).



The next step in the synthetic pathway towards the designed indole-based macrocycles is the coupling of two monomers. We started from the simplest monomer in our hands: the monomer without chains (**5**). So, in order to coupling this monomer I deprotected selectively the carboxylic and aminic functions and after I carried out the amidic coupling in solution using as a coupling agent classic amidic coupling agent as EDAC (1-ethyl-3-(3-dimethylaminopropyl) carbodiimide) to give the dimer with a 50% of yield. At this point, with an analogous way, I coupled the dimer unit with another monomer unit to give the trimer with only 5% of yield. Probably, the drastic decrease of yield in the coupling step, in the case of the trimer with respect to the dimer, is due to the progressive increasing of the length of the polymer. In fact the increasing of the number of indole

units advantages the possible intramolecular stacking between the different units causing the functional groups to be less reactive in the coupling reaction. For this reason we tried to change the coupling agent to increase the reactivity of the ammine and carboxylic groups. So, I improved the coupling conditions, bypassing the need of an external coupling agent. In fact, in the new strategy, the carboxylic group was activated as an acyl fluoride and the amine group as a silylamine: the two separately activated groups were then directly reacted together. In particular, I increased the yield of the trimer from 5% to 15%. Despite these results, the yield of this compound remained too low to continue the polymerization of the tetrameric compound. So, in order to polymerize the synthesized monomers we are now looking for a collaboration with a research group expert in peptide synthesis.

The second approach towards an indole-based polycyclic aromatic system which I dealt with during my PhD work was represented by the preparation of the triazatruxene core, a cyclotrimer of indoles where the three monomers are fused in a new aromatic core, which presents a wide aromatic surface with three useful points for the attachment of side chains, namely the three indolic NH at 5, 10 and 15 positions substituted triazatruxene derivatives. So, I decided to work on this kind of molecules in order to obtain tri-substituted compounds with optimal features as G-quadruplex ligands.

So I compare several synthetic approaches used to synthesize AZATRUX (**24a**), the first compound of a series, in order to define a versatile pathway that has been then applied in our lab to synthesize other derivatives, with side chains of different length and basicity which are two critical parameters for modulating G-quadruplex ligands properties.

Initially, I followed a synthetic approach reported in the literature in order to obtain N-substituted hydrophilic triazatruxene derivatives, which is represented by the cyclotrimerization of N-substitued-2-indolones with $POCl_3$. Following this way, I obtained in scarce yield both the suitable N-substitued-2-indolone by the first chain insertion step (13.6%), and the desired product AZATRUX (**24a**) by the second cyclotrimerization step (15%). Therefore, considering the scarce yield of the two key steps (chain insertion and cyclotrimerization), this synthetic approach did not result to be useful for the preparation of these triazatruxene derivatives. Consequently, in order to optimize the total yield, we tried to invert the steps order: first, the cyclotrimerization of the unsubstituted triazatruxene core and then its alkylation with the desired side chains. So, we began investigating the different preparations of the triazatruxene core.

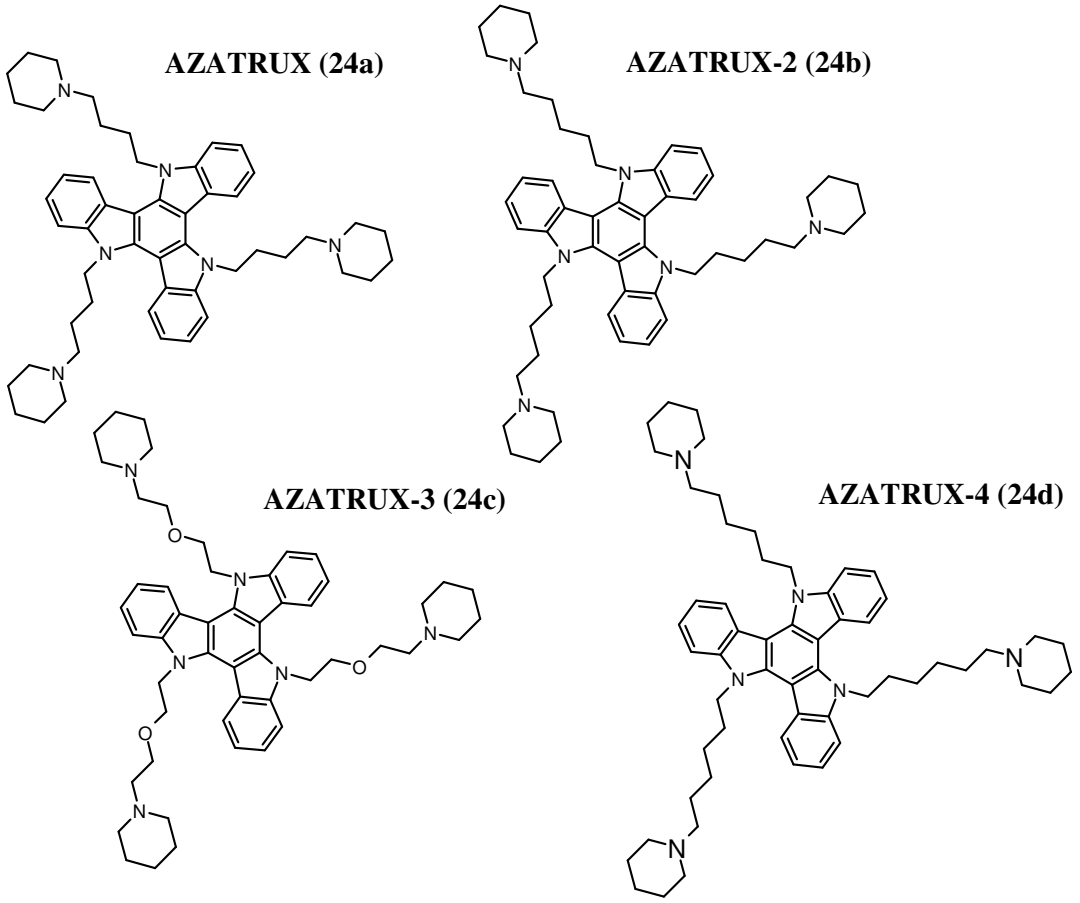
The most logical synthetic approach to obtain this core is the cyclocondensation of indole itself or unsubstituted 2-indolone via halogenation, similarly to what has been previously described for 2-indolone derivatives. So, I compared the cyclotrimerization of indole with bromine in suitable conditions and the cyclotrimerization of 2-indolone with $POCl_3$. This latter route, in our knowledge, represents the best synthetic approach to obtain the unsubstituted triazatruxene core in

terms of time, number of steps and yield. In fact, with the cyclocondensation of the unsibstuted 2-indolone we improved the yield of this step with respect to the suitable N-substituted 2-indolone previously reported from 15% to 48%.

Once optimized the cyclocondensation step with the preparation of the unsubstituted triazatruxene core, we focused our attention on the insertion of the three hydrophilic side chains. In a similar way as described above for the unsubstituted 2-indolone, we inserted the hydrophilic side chains obtaining AZATRUX (24a) with a total yield on the two steps of 36%. In this way I improved the yield of this step with respect to the N-substituted 2-indolone previously reported from 13.6% to 36%.

In conclusion, the best synthetic pathway to obtain AZATRUX (**24a**) is to perform first the cyclotrimerization in $POCl_3$ of the unsubstituted 2-indolone obtaining the triazatruxene core and, then, the insertion of hydrophilic side chains in two steps.

This pathway was applied for the preparation of other three compounds of this series (**24b-d**) with different longer chains with respect to AZATRUX (**24a**) in order to modulate the G-quadrupex interaction (Franceschin *et al.* **2009**).



Subsequently, in order to evaluate the activity of the new synthesized molecules we have studied their ability in inducing and/or stabilizing G-quadruplex DNA structures, as well as their selectivity towards G-quadruplex with respect to duplex DNA. To these aims, several experimental techniques have been used, such as ESI-Mass spectrometry, FRET, absorption spectroscopy. In particular, at the moment complete data by all these techniques have been reported for AZATRUX (**24a**), the first compound of the series, while for the other compounds (**24b-d**) I reported preliminary studies (FRET assays and ESI-MS experiments), since complete studies are still in progress.

The efficiency of AZATRUX in binding the telomeric G-quadruplex and its selectivity with respect to duplex DNA have been studied by electrospray ionization mass spectrometry (ESI-MS). The evaluation of the binding constants obtained by the collected data demonstrates that the studied molecule is a good G-quadruplex ligand able to form both 1:1 and 2:1 drug–DNA complexes and show a good G-quadruplex/duplex selectivity. In particular, competitive experiments show a good selectivity in binding the telomeric G-quadruplex structure with respect to the genomic duplex structure (CT DNA). As an independent confirmation of this quadruplex/duplex selectivity, we have performed titrations of the human telomeric F21T oligonucleotide and the biologically relevant CT DNA with AZATRUX by UV-vis spectroscopy (Ginnari-Satriani *et al.* **2009**) and fluorescence spectroscopy. This different behaviour of AZATRUX towards the two DNA structures confirms a stronger interaction for telomeric G-quadruplex with respect to genomic duplex DNA, also in these conditions (50 mM KCl solution) more similar to physiological conditions than those used in ESI-MS experiments.

Since drug self-aggregation represents a competitive process with binding to G-quadruplex DNA, the self-aggregation in aqueous solution of G-quadruplex ligands must be carefully considered in the model of their interaction with G-quadruplex DNA. For this reason, we have studied in detail the self association of AZATRUX in MES-KCl buffer by two methods (UV-vis absorption spectroscopy and NMR studies). These studies showed a poor tendency of AZATRUX, despite the large aromatic core, to aggregate in water solution both at μM concentration range and at mM concentration range.

In addition, a complete comparative study of all four compounds of the series (**24a-d**) on their selective affinity for G-quadruplex/duplex was carried out by FRET assay and ESI-MS experiments. From the combined reported data these new hydrophilic three side chained triazatruxene derivatives result in potent and selective G-quadruplex ligands and in particular AZATRUX (**24a**) is the best G-quadruplex ligand of this series showing a good selectivity with respect to the duplex DNA in according with the previous experiments. It is interesting to note that FRET assays performed on two G-quadruplexes derived from the *c-kit* proto-oncogene sequences

(*c-kit1* and *c-kit2*) showed that AZATRUX (**24a**) is able to stabilize these structures, suggesting the possible role of AZATRUX also in the regulation of oncogenes.

For all these reasons we decided to test AZATRUX (**24a**) in cell biology studies in vitro. In particular, extensive antitumor in vitro tests performed at the National Cancer Institute (NCI) on 60 cancer cell lines coming from different tissues showed a good anticancer activity of AZATRUX (**24a**) in all cell lines.

These very interesting results on AZATRUX, both from biophysical and biological points of view, encourage further and more specific studies on the mechanisms of action inside cells by this compound and, at the same time, a complete comparative study on all the other derivatives of this series (**24b-d**). Due to the possible pharmacological applications of these compounds, they have been recently included in a Patent as potential anticancer drugs (Ortaggi *et al.* **2009**).

5. Materials and methods.

5.1. Compounds preparations and characterizations.

5.1.1 General

All the commercial reagents and solvents were purchased from Fluka and Sigma-Aldrich, and used without further purification. TLC glass plates (silica gel 60 F_{254}) and silica gel 60 (0.040-0.063 mm) were purchased from Merck. ^{1}H and ^{13}C-NMR spectra were performed with Varian Gemini 200 and Varian Mercury 300 instruments. ESI-MS spectra were recorded on a Micromass Q-TOF MICRO spectrometer. Elemental analyses (C,H,N) were carried out on EA1110 CHNS-O (CE instruments). UV/Vis absorption spectra were registered at 298K using a JASCO V-530 spectrophotometer. MES-KCl buffer is composed by 10 mM MES at pH 6.5 and 50 mM KCl.

5.1.2 Indole-based macrocycles

methyl(5-nitro-1*H*-indol-1-yl)acetate (3): NaH (1.48 g, 61,65 mmol) was carefully added to a solution of commercially availables 5-nitro-indole (**1**) (2g, 12,33 mmol) and methyl bromoacetate (**2**) (2.44, 24.66 mmol) in THF anhydrous (100 ml) at 0°C. The reaction was stirred for two hours at room temperature. Then 20 ml of water were slowly added to the mixture at the temperature of 0°C and the aqueous layer was extracted with ethyl acetate (3X100 ml). The combined organic layers obtained were washed with brine solution, dried with Na_2SO_4 and evaporated under reduced pressure. Then, the crude product was treated with 30 ml of n-hexane and the solid was separated from the organic layer by decantation to give the compound **3**, as a pale yellow solid (2.68g, yield 93%). ^{1}H NMR (200 MHz, chloroform-d) δ 8.59 (d, 1H, J_m= 2.2 Hz, aromatic H), 8.14 (dd, 1H, J_o=9 Hz, J_m=2.2 Hz, aromatic H), 7.28 (d, 1H, J_o=9 Hz, aromatic H), 7.24 (d, 1H, J_o=3.3 Hz, aromatic H), 6.75 (dd, 1H, J_o=3.3 Hz, J=0.8 Hz, aromatic H), 4.92 (1s, 2H, N_{indole}-CH_2-COOMe), 3.78 (1s, 3H, OCH_3) ppm; ^{13}C NMR (chloroform-d6) δ: 168.0 (C=O), 142.1 ($C_{ar.}$), 139.3 ($C_{ar.}$), 131.6 ($C_{ar.}$), 127.9 ($C_{ar.}$), 118.2 ($C_{ar.}$), 117.7 ($C_{ar.}$), 109.0 ($C_{ar.}$), 105.0 ($C_{ar.}$), 52.8 (OCH_3), 47.91, (CH_2) ppm; MS (ESI) m/z 234 [$(M+H)^+$] (Calcd for $C_{11}H_{10}N_2O_4$: 234).

methyl(5-amino-1*H*-indol-1-yl)acetate (4): A mixture of Pd/C (10%) (100mg) and ethanol (10 ml) was stirred in presence of hydrogen gas at atmospheric pressure and at room temperature for fifteen minutes. Then the compound **3** (1 g, 4.27 mmol) was dissolved in 10 ml of THF and then was added to the mixture and the reaction was carried out for 6 hours at room temperature, keeping hydrogen atmosphere. Then, nitrogen was bubbled into the mixture for some minutes to leave up the hydrogen gas and the catalyser was filtred off on celite. After the organic layer was evaporated under reduced pressure to give the pure product **4**, as a dark viscous oil (820 mg, 94% yield). ^{1}H NMR (200 MHz, chloroform-d) δ 7.04 (d, 1H, J_o=8.7 Hz, aromatic H), 6.99 (d, 1H, J_o=3.3 Hz, aromatic H), 6.92 (d, 1H, J_m =2,1 Hz, aromatic H), 6.68 (dd, 1H, J_m =2,1 Hz, aromatic H), 6.37 (d, 1H, J_o=3.1 Hz, aromatic H), 4.78 (s, 2H, N_{arom}-CH_2-COOMe), 3.72 (s, 3H, OCH_3), 3.46 (br, 2H, -NH_2) ppm; ^{13}C NMR (chloroform-d6) δ: 169.0 (C=O), 139.4 ($C_{ar.}$), 131.3 ($C_{ar.}$) ,129.4 ($C_{ar.}$), 128.7 ($C_{ar.}$), 112.7 ($C_{ar.}$), 109.2 ($C_{ar.}$), 105.9 ($C_{ar.}$), 101.2 ($C_{ar.}$), 52.1 (OCH_3), 47.6, (CH_2) ppm; MS (ESI) m/z 204 [$(M+H)^+$] (Calcd for $C_{11}H_{12}N_2O_2$: 204).

methyl{5-[(*tert*-butoxycarbonyl)amino]-1*H*-indol-1-yl}acetate (5): A mixture of compound **4** (4.354 g, 21.34 mmoli) THF (40 ml) and commercially available $(BOC)_2O$ (5.40 ml, 23.47 mmol) was stirred under reflux condition for six hours. Then, the organic layer was evaporated under reduced pressure to give the crude product. After the crude product was treated with n-hexane (15 ml) and the organic layer was separated from the pure product **5**, as a orange solid (3.765 g, 58% of yield), through decantation. ^{1}H NMR (200 MHz, chloroform-d) δ 7.68 (s, 1H, J_o=8.7 Hz, aromatic H), 7.14 (s, 2H, aromatic H), 7.06 (d, 1H, J_o =3,3 Hz, aromatic H), 6.49 (d, 1H, J_o =3,3 Hz, aromatic H), 6.37 (d, 1H, J_o=3.1 Hz, aromatic H), 4.82 (s, 2H, N_{indole}-CH_2-COOMe), 3.72 (s, 3H, OCH_3), 1.54 (s, 9H, C-$(CH_3)_3$) ppm.

3-iodo-5-nitroindole (6): A solution of I_2 (1.57 g, 6.18 mmol) in DMF (5mL) was dropped into a solution of commercially available 5-nitroindole (**1**) (1 g, 6.17 mmol) and KOH (850 mg, 15.2 mmol) in DMF (5 mL) and was stirred at 80°C for 6 h. The reaction mixture was poured ice and water (100 mL) containing ammonia (0.5%) and sodium metabisulphite (0.1% aqueous solution). The precipitate was filtered and washed with water to give the pure product **6** (1.58 g, 88% of yield), as a dark yellow solid. ^{1}H NMR (200 MHz, acetone-d6) δ 11.39 (br, 1H, NH_{indole}), 8.25 (d, 1H, J_m= 2.2 Hz, aromatic H), 8.07 (dd, 1H, J_o=7 Hz, J_m=2.2Hz, aromatic H), 7.75 (s, 1H, aromatic H), 7.62 (d, 1H, J_o=7 Hz, aromatic H) ppm; MS (ESI) m/z 289 [$(M+H)^+$] (Calcd for $C_8H_5IN_2O_2$: 289).

methyl(3-iodo-5-nitro-1*H*-indol-1-yl)acetate (7): NaH (0.42 g, 8.75 mmol) was carefully added to a solution of 3-iodo-5-nitro-indole (**6**) (1.01 g, 3.5 mmol), and commercially available methyl bromoacetate (0.66 ml, 7 mmol) in THF anhydrous (50 ml) at the temperature of 0°C. The reaction was stirred for two hours at room temperature. Then 5 ml of water were slowly added to the mixture at the temperature of 0°C and the aqueous layer was extracted with ethyl acetate (3X40 ml). The combined organic layers were washed with brine solution, dried with Na_2SO_4 and evaporated under reduced pressure. Then, the crude product was treated with 30 ml of n-hexane and the solid is separated from the organic layer by decantation to give the product **7**, as a pale yellow solid (1,160 g, 92% of yield). ^{1}H NMR (200 MHz, acetone-d6) δ 8.26 (d, 1H, J_m= 0.8 Hz, aromatic H), 8.11 (dd, 1H, J_o=9 Hz, J_m=0.8 Hz, aromatic H), 7.75 (s, 1H, aromatic H), 7.66 (d, 1H, J_o=9 Hz, aromatic H), 5.27 (s, 2H, N_{indole}-*CH_2*-COOMe), 3.69 (s, 3H, OCH_3) ppm; ^{13}C NMR (acetone-d6) δ: 167.9 (C=O), 144.8 ($C_{ar.}$), 141.3 ($C_{ar.}$), 137.4 ($C_{ar.}$), 132.1 ($C_{ar.}$), 129.4 ($C_{ar.}$), 119.2 ($C_{ar.}$), 118.7 ($C_{ar.}$), 111.6 ($C_{ar.}$), 51.2 (OCH_3), 48.1, (CH_2) ppm; MS (ESI) m/z 361 [$(M+H)^+$] (Calcd for $C_{11}H_{10}IN_2O_4$: 361).

methyl{3-[6-(1-piperidino)-hex-1-en-1-yl]-5nitro-1H-indol-1-yl}acetate (9): $Pd(OAc)_2$ (3 mg) and triphenilphosphine (15 mg) were added to a solution of compound 5 (0.46 g, 1.28 mmol) in CH_3CN (5ml) was stirred at 100°C. Then, a solution of esenil piperidine (**8**) (426 mg, 2.55 mmol) in Et_3N (5ml) was added drop wise in to the mixture and the reaction was carried out for 24 h under argon at 100°C. Then, the reaction mixture was diluted with chloroform (100 ml), washed with water and after with brine solution. After the organic layer was dried with Na_2SO_4 and evaporated under vacuum to give the crude product. The latter was purified by flash chromatography (chloroform/n-hexane = 50:50) to give the compound of formula **9**, as a dark yellow solid (150 mg, yield 30%). ^{1}H NMR (200 MHz, chloroform-d) δ 8.58 (d, 1H, J_m= 2.2 Hz, aromatic H), 8.32-8.12 (m, 2H, aromatic H), 7.47 (1d, 1H, J_o=7.2 Hz, aromatic H), 5.26 (1d, 1H, J_{trans}=15.2 Hz, arom-*CH*=CH), 4.85 (1d, 1H, J_{trans}=15.2 Hz, arom-CH=*CH*-), 4.66 (1s, 2H, N_{arom}-*CH_2*-COOMe), 3.65 (1s, 3H, OCH_3), 3-1 (m, 18H) ppm; ^{13}C NMR (chloroform-d) δ: 170.3 (C=O), 144.9, 141.0, 136.1, 133.4, 129.3, 128.1, 125.0, 112.1, 111.1, 107.6, 58.8 (*CH_2*-$N_{piperidine}$), 54.1 (*CH_2*-$N_{piperidine}$), 51.2 (OCH_3), 46.8 (N_{indole}-*CH_2*), 28.5, 26.0, 25.0, 23,9, 19.9 ppm; MS (ESI) m/z 400 [$(M+H)^+$] (Calcd for $C_{22}H_{30}N_3O_4$: 400).

methyl{3-[6-(1-piperidino)-hexyl]-5-amino-1H-indol-1-yl}acetate (10): A mixture of Pd/C (10%) (20mg) and ethanol (2 ml) was stirred in presence of hydrogen gas at atmospheric pressure and room temperature for fifteen minutes. Then the compound **9** (0.05g, 0.12 mmol) was dissolved in THF (1ml) and then was added to the mixture. The reaction was carried out for 6 hours

at room temperature, keeping hydrogen atmosphere. Then, nitrogen was bubbled into the mixture for some minutes to leave up the hydrogen gas and the catalyst was filtred on celite. After the organic layer was evaporated under reduced pressure to give the pure product **10**, as a dark viscous oil (43.1 mg, 92.8 % of yield). ^{1}H NMR (chloroform-d): δ 8.56 (s, 1H, aromatic H), 8.11 (d, 1H, J_o =8.8 Hz, aromatic H), 7.03 (d, 1H, J_o=3, aromatic H), 6.80 (d, 1H, J=8.8 Hz, aromatic H), 4.72 (s, 2H, N_{indole}-*CH*$_2$-COOMe), 3.72 (s, 3H, OCH_3), 3-1 (m, 22H) ppm; ^{13}C NMR (chloroform-d) δ: 173.0 (C=O), 136.3 ($C_{ar.}$), 133.4 ($C_{ar.}$), 131.3 ($C_{ar.}$), 128.7 ($C_{ar.}$), 119.3 ($C_{ar.}$), 112.0 ($C_{ar.}$), 107.5 ($C_{ar.}$), 104.2 ($C_{ar.}$), 56.3, 53.3, 48.7, 46.3, 36.4, 33.6, 31.0, 27.1, 26.4, 24.0, 10.0 ppm; MS (ESI) m/z 372 [$(M+H)^+$] (Calcd for $C_{22}H_{33}N_3O_2$: 372).

methyl{5-[(*tert*-butoxycarbonyl)amino]-3-[(1-piperidino)-hexyl]-1*H*-indol-1-yl) } acetate (11): A solution of compound **10** (35 mg, 0.1 mmol) and $(BOC)_2O$ (23 mg, 0.1mmol) in CH_3CN (3 ml) was stirred under reflux condition for six hours. Then, the organic layer was evaporated under reduced pressure to give the crude product. After the crude product was treated with n-hexane (15 ml) and the organic layer was separated from the pure product **11**, as a dark yellow viscous oil (28 mg, 63% of yield), through decantation. ^{1}H NMR (chloroform-d): δ 8.51 (d, 1H, J_o=9 Hz, aromatic H), 7.89 (d, 1H, J_m=2.4 Hz, aromatic H), 6.99 (d, 1H, J_o=5 Hz, aromatic H), 6.77 (dd, 1H, J_o=5 Hz, J_m=2.4 Hz aromatic H), 4.86 (s, 2H, N_{indole}-*CH*$_2$-COOMe), 3.72 (s, 3H, COO-*CH*$_3$), 3-1 (m, 22H), 1.46 (s, 9H, C-*(CH*$_3$*)*$_3$) ppm; ^{13}C NMR (chloroform-d) δ: 173.0 (C=O), 156.3 (C=O) 136.3 ($C_{ar.}$), 133.4 ($C_{ar.}$), 131.3 ($C_{ar.}$), 128.7 ($C_{ar.}$), 119.3 ($C_{ar.}$), 112.0 ($C_{ar.}$), 107.5 ($C_{ar.}$), 104.2 ($C_{ar.}$), δ 80.6, 56.2, 53.2, 48.7, 46.2, 36.3, 33.5, 31.1, 28.9, 27.2, 26.3, 24.1, 10.1 ppm; MS (ESI) m/z 472 [$(M+H)^+$] (Calcd for $C_{27}H_{41}N_3O_4$: 472).

{5-[(*tert*-butoxycarbonyl)amino]-1*H*-indol-1-yl) } acetic acid (12): An aqueous solution of KOH 2N (1.17 ml, 2.34mmol) was added to a mixture of compound **5** (357 mg, 1.17 mmol) and methanol (9ml). The reaction was stirred at reflux condition for 40 minutes. Then, the volume of solvents was concentrated under reduced pressure, were added 10 ml of water and HCl in acqueous solution (37%) until the aqueous layer was acid. So, the aqueous layer was extracted with ethyl acetate (3X50ml) and the combined organic layers was washed with water (3X3ml) and after with brine solution (2X3ml). Then the organic layer was dried with Na_2SO_4 and evaporated under reduced pressure to give the pure product **12**, as a yellow-orange solid (310 mg, 91,4% of yield). ^{1}H NMR (200 MHz, chloroform-d) δ 8.47 (br, 1H, arom-*NH*-COO), 7.61 (s, 1H, aromatic H), 7.08 (s, 2H, aromatic H), 7.02 (d, 1H, J_o =3,3 Hz, aromatic H), 6.47 (d, 1H, J_o =3,3 Hz, aromatic H), 4.77 (s, 2H, N_{indole}-CH_2-COOMe), 1.52 (C-*(CH*$_3$*)*$_3$) ppm; ^{13}C NMR (chloroform-d6) δ: 172.0 (C=O),

154.2 (C=O), 133.5 ($C_{ar.}$), 130.5 ($C_{ar.}$), 129.4 ($C_{ar.}$), 128.6 ($C_{ar.}$), 116.0 ($C_{ar.}$), 112.0($C_{ar.}$), δ 109.1 ($C_{ar.}$), δ 102.1 ($C_{ar.}$), 80.2 ($C(CH_3)_3$), 47.5, (CH_2) 28.3 ($C(CH_3)_3$), ppm; MS (ESI) m/z 313 [$(M+Na)^+$] (Calcd for $C_{15}H_{18}N_2O_4Na$: 313).

Amidic coupling by EDAC - Synthesis of the compounds 13, 15 - general procedure (Scheme 3a): EDAC (1.1 eq.) was added to a mixture of appropriates aminic compound (1 eq.) and carboxylic compound (1 eq.) in anhydrous DMF (1ml). The reaction was stirred for 24 hours at room temperature. Then, an appropriate quantity of water was added and the aqueous layer was extracted with ethyl acetate. The combined organic layers were washed with water and then with brine solution, dried with Na2SO4 and after evaporated under vacuum. The crude product was purified by flash chromatography in the appropriate conditions.

.

Dimeric compound 13 (Scheme 3a): Following the procedure described above, 1.48 g of EDAC (1.18 mmol) was added to a mixture of compound **4** (310 mg, 1.07 mmol), compound **12** (218 mg, 1.07 mmol) and anhydrous DMF (1ml) was stirred. The reaction was carried out for 24 hours at room temperature. After the work-up as above, the crude product was dissolved in methanol, adsorbed on silica gel and purified by flash chromatography (ethyl acetate/n-hexane = 50:50) to give the pure compound of formula **13**, as a white solid (257 mg, yield 50,5%). ^{1}H NMR (300 MHz, chloroform-d) δ 7.80 (br, 1H, arom-NH-COO), 7.62 (d, 1H, J_m=1.8 Hz, aromatic H), 7.20-6.99 (m, 6H, aromatic H), 6.58 (s, 1H, aromatic H), 6.57 (d, 1H, J_o=3 Hz, aromatic H), 6.44 (d, 1H, J_o=3 Hz, aromatic H), 4.87 (s, 2H, N_{indole}-CH_2-CO), 4.77 (s, 2H, N_{indole}-CH_2-CO), 3.69 (s, 3H, OCH_3), 1.53 (s, 9H, C-$(CH_3)_3$) ppm; ^{13}C NMR (chloroform-d6) δ: 168.8 (C=O), 166.2 (C=O), 153.4 (C=O), 134.3 ($C_{ar.}$), 133.1 ($C_{ar.}$), 131.9 ($C_{ar.}$), 129.5 ($C_{ar.}$), 129.3 ($C_{ar.}$), 129.3 ($C_{ar.}$), 129.0 ($C_{ar.}$), 116.4 ($C_{ar.}$), 113.4 ($C_{ar.}$), 111.8 ($C_{ar.}$), 109.4 ($C_{ar.}$), 109.0 ($C_{ar.}$), 103.9 ($C_{ar.}$), 103.7 ($C_{ar.}$), 102.7 ($C_{ar.}$), 80.2 (C-$(CH_3)_3$) C, 52.5 (OCH_3), 50.7 (N_{indole}-CH_2), 47.8 (N_{indole}-CH_2), 28.4 (3C, C-$(CH_3)_3$) ppm; MS (ESI) m/z 499 [$(M+Na)^+$] (Calcd for $C_{26}H_{28}N_4O_5Na$: 499).

Trimeric compound 15 (Scheme 3a): Following the procedure described above, 175 mg of EDAC (0.91 mmol) was added to a mixture of compound **4** (170 mg, 0.83 mmol), compound **14** (285 mg, 1.07 mmol) and anhydrous DMF (1ml). The reaction was stirred for 24 hours at room temperature. After the work-up as above the crude product was dissolved in methanol, adsorbed on silica gel and purified by flash chromatography (ethyl acetate/n-hexane = 70:30) to give the compound of formula **15** (30 mg, 4,6%). ^{1}H NMR (300 MHz, DMSO-d6) δ 10.09 (br, 1H, arom-*NH*-CO), 10.06 (br, 1H, arom-*NH*-CO), 8.82 (br, 1H, arom-*NH*-COO), 7.84 (s, 2H, aromatic H), 7.68 (s, 1H, aromatic H),

7.4-7.2 (m, 9, aromatic H), 6.36 (s, 2H, aromatic H), 6.32 (d, 1H, J_o=3.6 Hz, aromatic H), 5.02 (s, 2H, N_{indole}-*CH_2*-CO), 4.87 (s, 2H, N_{indole}-*CH_2*-CO), 4.77 (s, 2H, N_{indole}-*CH_2*-CO), 3.67 (s, 3H, OCH_3), 1.47 (s, 9H, C-*$(CH_3)_3$*) ppm; (ESI) m/z 671 [$(M+Na)^+$] (Calcd for $C_{36}H_{36}N_6O_6Na$: 671).

Amidic coupling between acyl fluorides and N-silylamines - Synthesis of the compounds 13, 15 - general procedure (Scheme 3b):

Preparation of the N-Silylamine: To appropriate amine in acetonitrile was added BSA (0.5 eq.) and the reaction mixture was stirred at room temperature under N_2 for 1h. The volatiles were removed *in vacuo*.

Preparation of the acyl fluorides: A solution of appropriate carboxylic acid, cyanuric fluoride (2 eq.) and pyridine (1 eq.) in dichloromethane was stirred at room temperature under N_2 for 3h. Then dichloromethane was added to a reacted solution and then the organic phase was extracted with ice water, dried with $MgSO_4$ and evaporated under vacuum.

Coupling between the N-Silylamine and the acyl fluoride: To a solution of the N-silylamine previously obtained in acetonitrile under N_2 was dropped a solution of the appropriate acyl fluoride (1 eq.) in acetonitrile and then was added a catalytic amount of tetrabutylammonium fluoride. The reaction mixture was stirred at room temperature for 24h under N_2. Then, the volatiles were removed *in vacuo* and the crude product was purified by flash chromatography in the appropriate conditions.

Dimeric compound 13 (Scheme 3b): Following the procedure described above were formed the related silylamine and acyl fluoride respectively starting from the compound **4** (500 mg, 2.4 mmol) and the compound **12** (710 mg, 2.4 mmol). Then, carried out the amidic coupling between the two activated compounds and the related work up as previously described, is obtained the crude product. The latter was dissolved in methanol, adsorbed on silica gel and purified by flash chromatography (ethyl acetate/n-hexane = 50:50) to give the pure compound of formula **13**, as a white solid (860 mg, yield 75%) with the same characterization reported above.

Trimeric compound 15 (Scheme 3b): Following the procedure described above were formed the related silylamine and acyl fluoride respectively starting from the compound **4** (29 mg, 0.14 mmol) and the dimeric compound **14** (68 mg, 0.14 mmol). Then, carried out the amidic coupling between the two activated compounds and the related work up as previously described, is obtained the crude product. The latter was dissolved in methanol, adsorbed on silica gel and purified by flash

chromatography (ethyl acetate/n-hexane = 70:30) to give the pure compound of formula **15**, as a white solid (14 mg, yield 15%) with the same characterization reported above.

Dimeric compound 14: An aqueous solution of KOH 2N (2.78 ml, 5.56 mmol) was added to a mixture of compound **13** (1.37 g, 2.88 mmol) and methanol (25 ml). The reaction was stirred at reflux condition for 40 minutes. Then, the volume of solvents was concentrated under reduced pressure, were added 10 ml of water and HCl in acqueous solution (37%) until the aqueous layer was acid. So, the aqueous layer was extracted with ethyl acetate (3X50ml) and the combined organic layers was washed with water (3X3ml) and after with brine solution (2X3ml). Then the organic layer was dried with Na_2SO_4 and evaporated under reduced pressure to give the pure product **14**, as a white solid (1.08 g, 83.8% of yield). ^{1}H NMR (300 MHz, chloroform-d) δ 7.79 (br, 1H, arom-NH-COO), 7.59 (d, 1H, J_m=1.8 Hz, aromatic H), 7.20-6.95 (m, 7H, aromatic H), 6.59 (s, 1H, J_o=3.2 Hz, aromatic H), 6.58 (s, 1H, aromatic H), 6.45 (d, 1H, J_o=3.2 Hz, aromatic H), 4.97 (s, 2H, N_{indole}-*CH_2*-CO), 4.77 (s, 2H, N_{indole}-*CH_2*-CO), 3.69 (s, 3H, *OCH3*), 1.53 (s, 9H, $C(CH_3)_3$) ppm; ^{13}C NMR (methanol-d4) δ: 172.5 (C=O), 168.7 (C=O), 156.2 (C=O), 135.7 ($C_{ar.}$), 135.0 ($C_{ar.}$), 132.7 ($C_{ar.}$), 131.3 ($C_{ar.}$), 131.1 ($C_{ar.}$), 131.0 ($C_{ar.}$), 130.7 ($C_{ar.}$), 130.3 ($C_{ar.}$), 117.3 ($C_{ar.}$), 117.2 ($C_{ar.}$), 114.2 ($C_{ar.}$), 113.4 ($C_{ar.}$), 112.8 ($C_{ar.}$) , 110.3 ($C_{ar.}$), 102.8 ($C_{ar.}$), 80.4 (*C*-$(CH_3)_3$), 50.7, (CH_2), 48.2 (CH_2), 28.81 (3C, C-*$(CH_3)_3$*) ppm; MS (ESI) m/z 499 [$(M+Na)^+$] (Calcd for $C_{25}H_{26}N_4O_5Na$: 499).

5.1.3 Triazatruxene derivatives

1-(4-iodobutyl)-indolin-2-one (17): Commercially available 2-Indolone **16** (200mg, 1.5 mmol) was added to a suspension of sodamide (58 mg, 1.5 mmol) in dry dioxan (4 ml); the mixture was stirred under reflux until evolution of ammonia ceased. 1,4-Diiodobutane **22a** (1 ml, 7.5 mmol) was then added and the mixture was heated under reflux for 5h. The mixture was diluted with AcOEt, washed with water and then with brine solution. The organic layer was dried with Na_2SO_4 and evaporated under vacuum. The crude product was dissolved in chloroform, adsorbed on silica gel and purified by flash chromatography (chloroform/n-hexane = 70:30) to give compound **17** (70 mg, yield 15%), as a colourless viscous oil. ^{1}H NMR (300 MHz, chloroform-d) δ 7.27 (m, 2H, aromatic H), 7.03 (t, 1H, J=7.8 Hz, aromatic H), 6.84 (d, 1H, J=7.8 Hz, aromatic H), 3.73 (t, 2H, J=6.6 Hz, N-CH_2), 3.52 (s, 2H, CH_2-CO), 3.22 (t, 2H, I-CH_2), 1.86 (m, 4H, NCH_2-*CH_2*-*CH_2*-I) ppm; ^{13}C NMR (chloroform-d) δ: 175.0 (C=O), 144.4 ($C_{ar.}$), 127.9 ($C_{ar.}$), 124.6 ($C_{ar.}$), 124.5 ($C_{ar.}$), 122.3 ($C_{ar.}$), 108.3 ($C_{ar.}$), 38.7, 35.7, 30.5, 28.3, 5.8 ppm.

1-[4-(1-piperidino)-butyl]-indolin-2-one (18): A mixture of **17** (150mg, 4.76 mmol) and piperidine (0.69 ml, 7.0 mmol) was heated in THF (3ml) under reflux for 2h. The mixture was evaporated under reduced pressure and purified by flash chromatography (chloroform/methanol/30% aqueous ammonia solution = 95:5:0.5) to give the compound of formula **18** (116 mg, yield 90%), as a colourless viscous oil. ^{1}H NMR (300 MHz, chloroform-d) δ 7.23 (m, 2H, aromatic H), 6.99 (t, 1H, J=7.8 Hz, aromatic H), 6.83 (d, 1H, J=7.8 Hz, aromatic H), 3.70 (t, 2H, J=6.6 Hz, N-CH_2), 3.52 (s, 2H, CH_2-CO), 3.15 (m, 2H, $N_{piperidine}$-CH_2), 2.01 (m, 4H) 1.77 (m, 4H), 1.65 (m, 6H) ppm; ^{13}C NMR (chloroform-d) δ: 175.0 (C=O), 144.2 ($C_{ar.}$), 127.8 ($C_{ar.}$), 124.5 ($C_{ar.}$), 124.4 ($C_{ar.}$), 122.2 ($C_{ar.}$), 108.3 ($C_{ar.}$), 57.9, 54.0, 39.3, 35.7, 25.1, 24.8, 23.5, 23.0 ppm; HR-MS (ESI) m/z 295.1800 [$(M+Na)^+$] (Calcd for $C_{17}H_{24}N_2$ ONa: 295.1786)

5,10,15-tris[4-(1-piperidino)-butyl]diindolo[3,2-a:3',2'-c]carbazole (24a, AZATRUX) - Scheme 4: A mixture of **18** (100 mg, 0.37 mmol) and 2 ml of $POCl_3$ was heated at 100°C for 24 h. Then, the reaction mixture was poured into ice and basified carefully with KOH and the resulting precipitate washed with H_2O and then taken up in $CHCl_3$. This solution was washed with H_2O, dried on Na_2SO_4 and evaporated under reduced pressure to give the crude product as a brown solid. The latter was purified by flash chromatography (chloroform/methanol/30% aqueous ammonia solution = 90:10:1) to give the pure compound of formula **24a**, as a as a dark yellow viscous oil (14 mg, yield 15%). Product **24a** crystallized slowly cooling a saturated methanol solution to obtain white

needles (mp 122-123 °C). ^{1}H NMR (200 MHz, chloroform-d) δ 8.26 (d, 3H, J=8.4 Hz, aromatic H), 7.68 (d, 3H, J=8.2 Hz, aromatic H), 7.5-7.3 (m, 6H, aromatic H), 4.95 (t, 6H, J=7.4 Hz, N_{arom}-CH_2), 2.25 (m, 18H, $N_{piperidine}$-CH_2), 1.99 (m, 6H, $N_{arom}CH_2$-*CH_2*), 1.51 (m, 18H, $N_{piperidine}CH_2$-*CH_2*), 1.39 (m, 6H, γ-$CH_{2piperidine}$) ppm; ^{13}C NMR (chloroform-d) δ: 140.8 ($C_{ar.}$), 138.6 ($C_{ar.}$), 123.2 ($C_{ar.}$), 122.7 ($C_{ar.}$), 121.3 ($C_{ar.}$), 119.7 ($C_{ar.}$), 110.6 ($C_{ar.}$), 103.0 ($C_{ar.}$), 58.2, 54.2, 46.5, 27.5, 25.5, 24.1, 23.4 ppm; UV (DMSO) λ max, nm (ε x10^{-4}, $M^{-1}cm^{-1}$): 267 (1.9), 318 (6.1), 354 (1.0).

2,3,7,8,12,13-hexabromo-10,15-dihydro-5H-diindolo[3,2-a:3',2'-c]carbazole (20): A solution of Br_2 (4.10 g, 25.5 mmol) in CH_3CN (5 ml) was added at a mixture of commercially available indole **19** (1g, 8.5 mmol) and 15 ml of CH_3CN, stirring at room temperature, over 5 min. The mixture was left stirring overnight and then the resulting dark green solid was filtered and washed with acetonitrile (150 ml). The obtained product **18** was recrystallised by addition of acetone (80 ml) to a DMSO (3ml) solution of the crude product (181 mg, 10%). ^{1}H NMR (200 MHz, DMSO-d6) δ 12.30 (br, 3H, N_{arom}-H), 8.59 (s, 3H, J=7, aromatic H), 7.89 (s, 3H, aromatic H) ppm; HR-MS (ESI) m/z 811.5857 [$(M-H)^-$] (Calcd for $C_{24}H_8N_3{}^{79}Br_6$: 811.5818).

10,15-dihydro-5H-diindolo[3,2-a:3',2'-c]carbazole (21) - Scheme 5A: A mixture of **20** (175 mg, 0.21 mmol), Et_3N (0.24 ml, 1.71 mmol), HCOOH (0.06 ml, 1.71 mmol) and 10% Pd/C (68 mg, 0.06 mmol) in MeOH (15 ml) was heated for 30 min under reflux. The mixture was filtered through celite and the filtrate was diluted with CH_2Cl_2, washed with aqueous HCl sol. (10%), dried (Na_2SO_4) and evaporated under reduced pressure to give **21**, as a pale yellow solid (44 mg, yield 60%). ^{1}H NMR (200 MHz, acetone-d6) δ 11.08 (s, 3H, N_{arom}-H), 8.52 (d, 3H, J=7 Hz, aromatic H), 7.68 (d, 3H, J= 7.4 Hz, aromatic H), 7.4-7.2 (m, 6H, aromatic H) ppm; ^{13}C NMR (acetone-d6) δ: 141.0 ($C_{ar.}$), 136.4 ($C_{ar.}$), 124.8 ($C_{ar.}$), 124.5 ($C_{ar.}$), 121.5 ($C_{ar.}$), 121.4 ($C_{ar.}$), 112.9 ($C_{ar.}$), 103.3 ($C_{ar.}$) ppm; UV (DMSO) λ max, nm (ε x10^{-4}, $M^{-1}cm^{-1}$) : 274 (3.3), 309 (6.8), 325 (3.6) , 342 (1.5); MS (ESI) m/z 344 [$(M-H)^-$] (Calcd for $C_{24}H_{14}N_3$: 344).

10,15-dihydro-5H-diindolo[3,2-a:3',2'-c]carbazole (21) - Scheme 5B: Following the initial procedure described above, a mixture of indole **19** (1g, 8.5 mmol) and 15 ml of CH_3CN was stirred at room temperature and a solution of Br_2 (4.10 g, 25.5 mmol) in CH_3CN (5 ml) was added over 5 min. The mixture was stirred overnight and the resulting dark green solid was filtered and washed with acetonitrile (150 ml). At this point, without further purification, the crude product (2 g) was mixed with Et_3N (4.20 ml, 30.23 mmol), HCOOH (1.14 ml, 30.23 mmol) and 10% Pd/C (200 mg, 0.18 mmol) in MeOH (30 ml), and the resulting mixture was heated for 30 min under reflux. After a

work-up similar to that described above, the crude product was dissolved in methanol, adsorbed on silica gel and purified by flash chromatography (ethyl acetate/n-hexane = 15:85) to give the pure compound of formula **21**, as a pale yellow solid (128 mg, yield 13%), with the same characterization reported above.

10,15-dihydro-5H-diindolo[3,2-a:3',2'-c]carbazole (21) - Scheme 5C: Following a similar procedure to that described above, a mixture of 2-indolone **16** (2.0 g, 15 mmol) and 10 ml of $POCl_3$ was heated at 100°C for 8 h. Then, the reaction mixture was poured into ice and neutralized carefully with KOH until pH 7-8. After neutralization, the precipitate was filtered to give the crude product as a brown solid. The latter was dissolved in methanol, adsorbed on silica gel and purified by flash chromatography (ethyl acetate/n-hexane = 15:85) to give the pure compound of formula **21**, as a pale yellow solid (830 mg, yield 48%), with the same characterization reported above.

Synthesis of the intermediate compounds 23a, b, c and d - general procedure (Scheme 6): A mixture of the triazatruxene core **21** and KOH (10 eq.) in THF was heated under reflux for 10 min. An excess of the appropriate dihalide was then added and the mixture was heated under reflux for 6h. The mixture was diluted with AcOEt, washed with 10% aqueous HCl and then with brine solution. The organic layer was dried with Na_2SO_4 and evaporated under vacuum. The crude product was dissolved in chloroform, adsorbed on silica gel and purified by flash chromatography in the appropriate conditions.

5,10,15-tris(4-iodobutyl)diindolo[3,2-a:3',2'-c]carbazole (23a): Following the procedure described above, a mixture of **21** (355 mg, 1.03 mmol) and KOH (576 mg, 10.3 mmol) in THF (20 ml) was reacted with 1,4-diiodobutane **22a** (2.0 ml, 15 mmol). After the work-up as above, the crude product was purified by flash chromatography (ethyl acetate/n-hexane = 5:95) to give the compound of formula **23a** (373 mg, yield 40%), as a dark yellow viscous oil. ^{1}H NMR (200 MHz, acetone-d6) δ 8.30 (d, 3H, J=7.8 Hz, aromatic H), 7.79 (d, 3H, J=7.8 Hz, aromatic H), 7.5-7.3 (m, 6H, aromatic H), 5.03 (t, 6H, J=7.4 Hz, N_{arom}-CH_2), 3.07 (t, 6H, J=7.0 Hz, I-CH_2), 1.89 (m, 6H, $N_{arom}CH_2$-*CH_2*), 1.61 (1m, 6H, ICH_2-*CH_2*) ppm; ^{13}C NMR (acetone-d6) δ: 141.4 ($C_{ar.}$), 138.8 ($C_{ar.}$), 123.5 ($C_{ar.}$), 123.3 ($C_{ar.}$), 121.8 ($C_{ar.}$), 120.3 ($C_{ar.}$), 111.3 ($C_{ar.}$), 103.6 ($C_{ar.}$), 45.7, 30.6, 30.4, 6.1 ppm; MS (ESI) m/z 892 [$(M+H)^+$] (Calcd for $C_{36}H_{37}I_3N_3$: 892).

5,10,15-tris(5-iodopentyl)diindolo[3,2-a:3',2'-c]carbazole (23b): Following the procedure described above, a mixture of **21** (100 mg, 0.29 mmol) and KOH (81 mg, 1.40 mmol) in THF (2 ml)

was reacted with 1,5-diiopentane **22b** (0.64 ml, 4.35 mmol). After the work-up as above, the crude product was purified by flash chromatography (ethyl acetate/n-hexane = 10:90) to give the compound of formula **23b** (97 mg, yield 36%), as a dark yellow viscous oil. ^{1}H NMR (300 MHz, chloroform-d) δ 8.23 (d, 3H, J= 7.5 Hz, aromatic H), 7.61 (d, 3H, J=8.1 Hz, aromatic H), 7.5-7.3 (m, 6H, aromatic H), 4.89 (t, 6H, J=7.8 Hz, N_{arom}-CH_2), 3.03 (t, 6H, J=6.6 Hz, I-CH_2), 1.95 (m, 6H, $N_{arom}CH_2$-*CH_2*), 1.74 (m, 6H, ICH_2-*CH_2*), 1.35 (m, 6H, ICH_2CH_2-*CH_2*) ppm; ^{13}C NMR (chloroform-d) δ: 140.9 ($C_{ar.}$), 138.6 ($C_{ar.}$), 123.4 ($C_{ar.}$), 122.9 ($C_{ar.}$), 121.4 ($C_{ar.}$), 119.9 ($C_{ar.}$), 110.5 ($C_{ar.}$), 103.3 ($C_{ar.}$), 46.5, 32.8, 28.5, 27.4, 6.3 ppm; MS (ESI) m/z 934 [$(M+H)^+$] (Calcd for $C_{39}H_{43}I_3N_3$: 934).

5,10,15-tris(5-bromo-3-oxapentyl)diindolo[3,2-a:3',2'-c]carbazole (23c): Following the procedure described above, a mixture of **21** (130 mg, 0.38 mmol) and KOH (0.71 ml, 5.7 mmol) in THF (2 ml) was reacted with 2-bromoethyl ether **22c** (1.0 ml, 15 mmol). After the work-up as above, the crude product was purified by flash chromatography (ethyl acetate/n-hexane = 15:85) to give the compound of formula **23c** (120 mg, 40 yield %), as a yellow viscous oil. ^{1}H NMR (200 MHz, chloroform-d6) δ 8.32 (d, 3H, J= 8.4 Hz, aromatic H), 7.74 (d, 3H, J=8.2 HZ, aromatic H), 7.5-7.3 (m, 6H, aromatic H), 5.19 (t, 6H, J=6.6 Hz, N_{arom}-CH_2), 3.99 (t, 6H, J=6.6 Hz, $N_{arom}CH_2$-*CH_2*-O), 3.61 (t, 6H, J=5.8 Hz, O-*CH_2*-CH_2Br), 3.25 (t, 6H, J=5.8 Hz, OCH_2-*CH_2*-Br) ppm; ^{13}C NMR (chloroform-d) δ: 141.2 ($C_{ar.}$), 138.7 ($C_{ar.}$), 123.2 ($C_{ar.}$), 123.1 ($C_{ar.}$), 121.6 ($C_{ar.}$), 120.4 ($C_{ar.}$), 110.6 ($C_{ar.}$), 103.5 ($C_{ar.}$), 71.0, 69.3, 46.1, 30.0 ppm; MS (ESI) m/z 800 [$(M+H)^+$] (Calcd for $C_{36}H_{38}Br_3N_3O_3$: 800).

5,10,15-tris(6-iodohexyl)diindolo[3,2-a:3',2'-c]carbazole (23d): Following the procedure described above, a mixture of **21** (130 mg, 0.38 mmol) and KOH (211 mg, 3.80 mmol) in THF (5 ml) was reacted with 1,6-diiodohexane **22d** (0.94 ml, 5.7 mmol). After the work-up as above, the crude product was purified by flash chromatography (ethyl acetate/n-hexane = 10:90) to give the compound of formula **23d** (150 mg, yield 40%), as a dark yellow viscous oil. ^{1}H NMR (200 MHz, chloroform-d) δ 8.26 (d, 3H, J=7.8 Hz, aromatic H), 7.63 (d, 3H, J=8.2 Hz, aromatic H), 7.5-7.3 (m, 6H, aromatic H), 4.92 (t, 6H, J=7.4 Hz, N_{arom}-CH_2), 3.02 (t, 6H, J=7.4 Hz, I-CH_2), 1.95 (m, 6H, $N_{arom}CH_2$-*CH_2*), 1.66 (m, 6H, ICH_2-*CH_2*), 1.27 (m, 12H, ICH_2CH_2-*CH_2*) ppm.

Synthesis of the final compounds 24a, b, c and d - general procedure (Scheme 6): A mixture of the intermediate compound **23** and piperidine was heated in THF under reflux for 2h. The mixture

was evaporated under reduced pressure and purified by flash chromatography (ethyl acetate saturated with 30% aqueous ammonia solution) to give the corresponding compound of formula **24**.

5,10,15-tris[4-(1-piperidino)-butyl]diindolo[3,2-a:3',2'-c]carbazole (24a, AZATRUX): Following the general procedure, **23a** (209 mg, 0.23 mmol) was reacted with piperidine (0.69 ml, 7.0 mmol) in THF (5ml). The purification of the crude product gave the compound of formula **24a** (161 mg, yield 91%), as a dark yellow viscous oil. A part of compound **24a** (110 mg) was precipitated in the form of its hydrochloride salt by dissolving it in a mixture of methanol/HCl (methanol/aqueous HCl 37%=95:5) and adding diethyl ether: 81 mg of a white solid was obtained with a yield of 58%. ^{1}H NMR (300 MHz, D_2O) δ 7.61 (d, 3H, J=7.8 Hz, aromatic H), 7.40 (m, 6H, aromatic H), 7.25 (m, 3H, aromatic H), 3.98 (m, 6H, N_{arom}-CH_2), 2.89 (br, 6H, $N_{piperidine}$-CH_2), 2.53 (m, 6H, $N_{piperidine}$-CH_2), 2.29 (br, 6H, $N_{piperidine}$-CH_2), 1.49 (m, 24H, $N_{piperidine}CH_2$-*CH_2*, $N_{arom}CH_2$-*CH_2*), 1.06 (m, 6H, $CH_{2piperidine}$) ppm; UV (aq. MES-KCl) λ max, nm (ε x 10^{-4}, $M^{-1}cm^{-1}$): 255 (2.6), 313 (4.8); HR-MS (ESI) m/z 763.5393 [$(M+H)^+$] (Calcd for $C_{51}H_{67}N_6$: 763.5427). Anal. found C 63.5% , H 8.3%, N 8.4%; (Calcd for [$C_{51}H_{66}N_6$·3HCl·5H_2O]: C63.6 %, H8.3%, N 8.7%).

5,10,15-tris[5-(1-piperidino)-pentyl]diindolo[3,2-a:3',2'-c]carbazole (24b, AZATRUX-2): Following the general procedure, **23b** (100 mg, 0.29 mmol) was reacted with piperidine (0.31 ml, 3.1 mmol) in THF (2ml). The purification of the crude product gave the compound of formula **24b** (73 mg, yield 91%), as a dark yellow viscous oil. ^{1}H NMR (300 MHz, chloroform-d) δ 8.20 (d, 3H, J=8.1 Hz, aromatic H), 7.60 (d, 3H, J=8.4 Hz, aromatic H), 7.5-7.3 (m, 6H, aromatic H), 4.92 (t, 6H, J=6.6 Hz, N_{arom}-CH_2), 2.5-2.3 (m, 18H, $N_{piperidine}$-CH_2), 1.9-1.7 (m, 18H, $N_{arom}CH_2$-*CH_2*-*CH_2*-*CH_2*-$CH_2N_{piperidine}$), 1.47 (m, 12H, $N_{piperidine}CH_2$-*CH_2*), 1.02 (m, 6H, γ-$CH_{2piperidine}$) ppm;. δ: 140.8 ($C_{ar.}$), 138.5 ($C_{ar.}$), 123.1 (2C, $C_{ar.}$), 121.4 ($C_{ar.}$), 120.2 ($C_{ar.}$), 110.9 ($C_{ar.}$), 103.3 ($C_{ar.}$), 57.5, 53.2, 46.3, 28.7, 24.0, 23.8, 23.7, 22.5 ppm; UV (DMSO) λ max, nm (ε x10^{-4}, $M^{-1}cm^{-1}$): 264 (2.2), 317 (6.1), 354 (0.9). HR-MS (ESI) m/z 805.5864 [$(M+H)^+$] (Calcd for $C_{54}H_{73}N_6$: 805.5897).

5,10,15-tris[5-(1-piperidino)-3-oxapentyl]diindolo[3,2-a:3',2'-c]carbazole (24c, AZATRUX-3): Following the general procedure, **23c** (120 mg, 0.15 mmol) was reacted with piperidine (0.38 ml, 3.8 mmol) in THF (2ml). The purification of the crude product gave the compound of formula **24c** (94 mg, yield 89%), as a dark yellow viscous oil. ^{1}H NMR (300 MHz, chloroform-d) δ 8.31 (d, 3H, J=7.8 Hz, aromatic H), 7.72 (d, 3H, J=9 Hz, aromatic H), 7.5-7.3 (m, 6H, aromatic H), 5.16 (t, 6H, J=6.6 Hz, N_{arom}-CH_2), 3.93 (t, 6H, J=6.0 Hz, $N_{arom}CH_2$-*CH_2*-O), 3.47 (t, 6H, J=6.6 Hz, O-*CH_2*-$CH_2N_{piperidine}$), 2.40 (t, 6H, J=6.3 Hz, OCH_2-*CH_2*-$N_{piperidine}$) 2.27 (m, 12H, $N_{piperidine}$-CH_2), 1.49 (m,

12H, $N_{piperidine}CH_2$-*CH_2*), 1.33 (m, 6H, γ-$CH_{2piperidine}$) ppm; ^{13}C NMR (chloroform-d). δ: 141.1 ($C_{ar.}$), 138.8 ($C_{ar.}$), 123.1 ($C_{ar.}$), 123.0 ($C_{ar.}$), 121.6 ($C_{ar.}$), 120.1 ($C_{ar.}$), 110.6 ($C_{ar.}$), 103.4 ($C_{ar.}$), 69.3, 69.1, 58.3, 54.7, 46.2, 25.6, 24.0 ppm. UV (DMSO) λ max, nm (ε $x10^{-4}$, $M^{-1}cm^{-1}$): 263 (2.3), 317 (6.1), 352 (1.0). HR-MS (ESI) m/z 811.5270 [$(M+H)^+$] (Calcd for $C_{51}H_{67}N_6O_3$: 811.5275).

5,10,15-tris[6-(1-piperidino)-hexyl]diindolo[3,2-a:3',2'-c]carbazole (24d, AZATRUX-4): Following the general procedure, **23d** (150 mg, 0.15 mmol) was reacted with piperidine (0.46 ml, 4.6 mmol) in THF (2ml). The purification of the crude product gave the compound of formula **24d** (114 mg, yield 90%), as a dark yellow viscous oil. ^1H NMR (300 MHz, chloroform-d) δ 8.25 (d, 3H, J=7.8 Hz, aromatic H), 7.61 (d, 3H, J=8.7 Hz, aromatic H), 7.5-7.3 (m, 6H, aromatic H), 4.89 (m, 6H, N_{arom}-CH_2), 2.5-2.2 (m, 18H, $N_{piperidine}$-CH_2), 1.7-1.2 (m, 42H) ppm. ^{13}C NMR (chloroform-d). δ: 140.9 ($C_{ar.}$), 138.8 ($C_{ar.}$), 123.4 ($C_{ar.}$), 122.7 ($C_{ar.}$), 121.4 ($C_{ar.}$), 119.6 ($C_{ar.}$), 110.5 ($C_{ar.}$), 103.1 ($C_{ar.}$), 59.1, 54.4, 46.8, 29.6, 27.2, 26.4, 25.7, 24.3 ppm. HR-MS (ESI) m/z 847.6355 [$(M+H)^+$] (Calcd for $C_{57}H_{79}N_6$: 847.6366).

5.2. Biophysical and biological assays.

5.2.1. Molecular modelling.

All the experiments were performed using the InsightII package on an SGI workstation. This software is widely used for the modelling of protein-ligand complexes, and more rarely for DNA-ligand complexes (Osiadacz *et al.* **2000**).

The G-quadruplex structure used in all the simulations is the X-ray derived monomeric structure of the 22-mer human telomeric DNA sequence $AGGG(TTAGGG)_3$, (par. 1.1; Parkinson *et al.* **2002**) 1KF1 code in the PDB. The PDB coordinates file was imported into the InsightII modelling package; the potassium ions in the central channel between the planes of each G-quartet were preserved, while all water molecules were deleted. Hydrogen atoms were added to the structure and potentials and partial charges were assigned according to the CVFF force-field, considering a pH of 7.0. A +1 partial charge was assigned to each potassium ion.

Ligand molecules were built using the Builder module in InsightII. Nitrogen atoms in the side chains were protonated assigning a +1 formal charge and CVFF force-field atom and bond types were assigned to all the structures. All the structures were energy-minimized (2000 steps, Polak-Ribiere conjugate gradient) using the Discover3 module.

Docking was performed in two phases with the Affinity Docking module of InsightII. The binding pocket was defined as all and only H atoms on the guanine bases of the external G-quartet planes. In the first phase 200 ligand orientations were randomly centred on the G-quadruplex structure. In this phase charges were not considered, non-bonded cut-offs were set to 8Å and Van der Waals radii to 10% of the full value. The complexes were minimized for 500 steps using Polak-Ribiere conjugate gradient method; energy tolerance and energy range were set respectively to 10000 and 40kcal/mol. In the second phase, the 75 lowest energy structures were used to perform simulated annealing. During this phase Van der Waals radii were adjusted to their full values and a distance-dependent dielectric constant of 4.0 was used. Each system was again minimized for 500 steps of conjugate gradient and then molecular dynamics was performed, starting at a temperature of 800K and cooling the system to 200K over a period of 10ps. The resulting structures were minimized for 2000 steps of conjugate gradient and the 25 structures with the lowest total energies were evaluated with the Analysis module of InsightII.

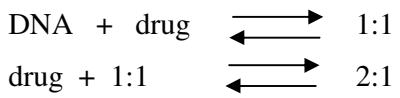
5.2.2. ESI-mass experiments.

*Oligonucleotide*s.

Single-stranded oligonucleotides were purchased from Eurofins MWG Operon (Ebersberg,Germany). Their sequences are as follows: TG_4T (5'-TGGGGT-3'); F21T (5'-GGGTTAGGGTTAGGGTTAGGG-3'); F21TTT (5'-GGGTTAGGGTTAGGGTTAGGGTT-3'); DK66 (5'-CGCGAATTCGCG-3'). Calf thymus DNA, used in competition experiments, was purchased from Sigma-Aldrich and subjected to sonication (Sonyprep 150 sonicator) for 8 minutes to obtain an average length of 500bp (according to gel electrophoresis analysis with Mass Ruler DNA ladder mix-low range). Calf thymus DNA, used in competition experiments, was purchased from Sigma-Aldrich.

Evaluation of the binding constants and percentage of bound DNA.

Considering drug-DNA complexes in 1:1 and 2:1 stoichiometry, the formation of such complexes can be represented by two distinct equilibria:



which are in turn described by the following two equations:

$$K_1 = [1:1] / ([DNA]\,[drug]) \qquad \textbf{1}$$

$$K_2 = [2:1] / ([1:1]\,[drug]) \qquad \textbf{2}$$

where [DNA], [drug], [1:1] and [2:1] represent respectively the concentrations of the different species in solution: DNA (duplex or quadruplex depending on the oligonucleotide used), the ligand, the 1:1 and 2:1 drug-DNA complexes at equilibrium.

The association constants K_1 and K_2 (equations **1** and **2**, respectively) can be calculated directly from the relative intensities of the corresponding peaks found in the mass spectra, with the assumption that the response factors of the oligonucleotides alone and of the drug-DNA complexes are the same, so that the relative intensities in the spectrum are supposed to be proportional to the relative concentrations in the injected solution (Rosu *et al.* **2002**):

$$I_{(DNA)} / [DNA] = I_{(1:1)} / [1:1] = I_{(2:1)} / [2:1] \qquad \textbf{3}$$

In this way, since DNA and drugs initial concentrations (C_0 and $C_{0'}$ respectively) are known, it is possible to obtain the concentration of each species appearing in **1** and **2**:

$$[j] = C_0 \cdot I_{(j)} / (I_{(DNA)} + I_{(1:1)} + I(_{2:1)}) \qquad \textbf{4}$$

$$[drug] = C_{0'} - [1:1] - 2[2:1] \qquad \textbf{5}$$

where [j] stands for [DNA], [1:1] or [2:1].

The constants were determined at different drug/DNA ratios, depending on the intensity of the signals (Guittat *et al.* **2009**): 2.5:5, 5:5, 7.5:5, 10:5 and 20:5 micromolar concentrations ratios.

A further manipulation of the obtained data leads to calculate the amount of ligand bound, according to an equation developed by de Pauw and his group (Rosu *et al.* **2003**) derived from **4**

$$[ligand\ bound] = C_0\,(I_{(1:1)} + 2I_{(2:1)}) / (\,I_{(DNA)} + I_{(1:1)} + I_{(2:1)}\,) \qquad \textbf{6}$$

This parameter, representing the total amount of the drug bound to DNA, is useful to compare the efficiency of different ligands in DNA binding, when they interact as single molecules. Since perylene derivatives are known to interact with DNA also in self-aggregate forms (Rossetti *et al.* **2005**), we decided to carry out a slightly different approach, according to another equation, described by Brodbelt and co-workers (Mazzitelli *et al.* **2006**), which has been demonstrated to be

more correct in such cases and was specifically applied in the study of the interactions between DNA and some perylene derivatives:

$$\% \text{ bound DNA } (\%b) = 100 \cdot (I_{(1:1)} + I_{(2:1)}) / (I_{(DNA)} + I_{(1:1)} + I_{(2:1)}) \qquad \mathbf{7}$$

This parameter (%b) represents the percentage of DNA engaged in complexation with the ligand (Casagrande *et al.* **2009**).

Evaluation of the percentage of bound quadruplex DNA in the competition experiments:

Since in many cases the starting point (the percentage of bound quadruplex DNA in absence of CT DNA) is very different for the various compounds, it is not straightforward to compare the behavior of the different compounds. In order to make a direct comparison between the tested compounds in terms of selectivity, we have decided to analyze the relative variation of the percentage of quadruplex bound, normalized % bound quadruplex (N%), defined as:

$$N\% = \% \text{ quadruplex bound in presence of CT} / \% \text{ quadruplex bound in absence of CT} \qquad \mathbf{8}$$

This parameter could be considered as a direct measure of how much a molecule is able to remain bound to quadruplex DNA even in presence of increasing amounts of double-stranded DNA (Casagrande *et al.* **2009**).

ESI mass spectra of duplex and quadruplex DNA.

In the negative ion ESI mass spectrum of duplex DNA alone (DK66) three ions are easily identified at a *m/z* ratio in agreement with the calculated values: an intense peak at *m/z* 1457, identified as $[\text{duplex} - 5H]^{5-}$ and two less intense peaks, identified as $[\text{duplex} - 4H]^{4-}$, at *m/z* 1822 and as $[\text{duplex} - 4H]^{6-}$ at *m/z* 1215. In the case of quadruplex DNA, monovalent cations such as ammonium can bind to the folded oligodeoxynucleotide, as an intrinsic component of the itself quadruplex structure (Williamson **1993**, Rhodes & Giraldo **1995**). Each oligodeoxynucleotide has a mass spectrum with a typical pattern, depending on the number of ammonium ions included in the structure: TG_4T presents peaks corresponding to $[(TG_4T)_4 + 3NH_4^+ - 8H^+]^{5-}$ and $[(TG_4T)_4 + 3NH_4^+ - 7H^+]^{4-}$ (respectively at *m/z* 1500 and 1875) as clearly predominant ions; in some cases the peak at *m/z* 1250 corresponding to the 6- charge state is also detectable. F21T and F21TTT, on the other hand, do not show ammonium adducts as dominant species in their mass spectra; this feature has also been observed by Bowers, De Pauw and coworkers: intense peaks are detectable at *m/z* 1329 and 1451,

relative to the ions $[F21T - 5H^+]^{5-}$ and $[F21TTT - 5H^+]^{5-}$ and less intense peaks are present at *m/z* 1662 and 1815 relative to the ions $[F21T - 4H^+]^{4-}$ and $[F21TTT - 4H^+]^{4-}$. The intensity of the peaks is related to the degree of desolvation of the $NH_4^+/PO4^-$ ion pairs: if all ion pairs lose NH_3 to give back the neutral PO_3H, it means that too harsh conditions have been used, leading even to partial fragmentation of the analyte; if too many ion pairs remain, the measurement of the intensities is too imprecise. So, we decided to select a restricted number of charge states for each oligonucleotide with proper tuning of the spectrometer in order to measure the intensities of the peaks as precisely as possible and so determine the binding constants. Proper tuning means that there is enough desolvation with the minimal fragmentation of the ions, giving a peak with a suitable signal-to-noise ratio, making sure to leave unaltered the response factors of oligonucleotides alone and DNA-drug complexes, which are supposed to be identical. We found that, for our studies, two charge states respect these parameters: hence, we have chosen 5^- and 4^- ions for TG4T, while only the 5^- charge state has been selected for F21T, F21TTT and DK66.
As for the instrumental parameters, we properly adjusted the capillary voltage, the source temperature and pressure, the cone voltage and the collision energy, as described below.

Instrumentation.

All the experiments were performed on a Q-TOF MICRO spectrometer (Micromass, now Waters) equipped with an ESI source, in the negative ionization mode. The rate of sample infusion into the mass spectrometer was 10 µl/min. The capillary voltage was set to -2.6 kV, the source temperature was adjusted to 70°C and the cone voltage to 30 V. Full scan MS spectra were recorded in the m/z range between 800 and 2500, with 100 acquisitions per spectrum. Data were analyzed using the MassLynx software developed by Waters.

Samples preparation protocol.

Oligonucleotides were dissolved in bi-distilled water to obtain the starting stock solutions and were annealed in 150 mM ammonium acetate buffer by heating at 90°C for 5 minutes and then cooling slowly to room temperature. The final concentration of oligonucleotides stocks was 50 µM in either duplex or quadruplex units. Ammonium acetate was chosen as the buffer main component for its good compatibility with ESI-MS. Calf thymus DNA (CT) was dissolved in bi-distilled water. Since its average chain length is 13kb, it was subjected to sonication (Sonyprep 150 sonicator) for 8 minutes to obtain an average length of 500bp (according to gel electrophoresis analysis with Mass Ruler DNA ladder mix-low range). Drugs stock solutions were prepared dissolving in bi-distilled water the desired amount of drug hydrochlorides to obtain a final concentration of 100 µM.

Samples were prepared by mixing appropriate volumes of 150 mM ammonium acetate buffer, 50 μM annealed oligonucleotide stock solution, perylene or coronene derivatives 100 μM stock solutions and methanol. The final concentration of DNA in each sample was 5 μM (in duplex or quadruplex unit) and the final volume of the sample was 50 μl. Drugs were added at different drug/DNA ratios, ranging between 0.5 and 4. Methanol was added to the mixture just before injection (in a percentage of 15 % vol.) after the complexation equilibrium in ammonium acetate was established, in order to obtain a stable electrospray signal. As a reference, samples containing only 5 μM DNA with no drug were prepared in each series.

Competion experiments.

Samples for competition experiments were prepared following the procedure described above, adding an appropriate volume of CT solution. Final concentrations of quadruplex DNA and drug solutions were always 5 μM and CT was added at two different duplex/quadruplex ratios (1 and 5), calculated on the basis of the phosphate group concentrations.

In order to minimize casual errors each experiment has been repeated at least five times, in the same experimental conditions, and data were processed and averaged with the SIGMA-PLOT software.

5.2.3. NMR DOSY experiment

The experiment was performed on a Bruker Avance 400 spectrometer, operating at 400.13 MHz for the 1H Larmor frequency, at 298 K by using 5-mm NMR tubes.

Experimental condition:

- spectral width: 12 ppm;
- data points: 64 K;
- digital resolution: 0.07 Hz;
- number of scans:16;
- 90° pulse lenght: 11 μs;
- recycling time: 2s;
- DOSY pulse program: ledgp2s;
- diffusion gradient lenght *little delta* δ: 2 ms;
- diffusion time *big delta* Δ : 150 ms;
- number of gradient increments: 32;
- gradient shape: sine bell;
- recovery time between two gradient pulses: 200μs.

5.2.4 UV/vis absorption and fluorescence spectroscopy

UV/vis absorption:

Absorbance VS drug concentration:

Absorbance of AZATRUX in MES/KCl buffer was registered at an initial concentration of 17 μM, following further dilutions up to 1.7 μM.

Titrations by UV spectroscopy:

F21T was annealed in MES/KCl buffer, heating the sample at 90°C for ten minutes and cooling slowly to room temperature. Both F21T and sonicated CT DNA were diluted in MES/KCl buffer at an initial concentration of 200μM (R=20). Adding increasing volumes of 10μM drug stock the DNA/drug ratio (R) was decreased until a final DNA concentration of 1μM (R=0.1), while taking constant drug concentration. Two series of measures were performed for each DNA to confirm reproducibility.

Fluorescence spectroscopy:

Drugs stocks preparation.

Drugs were dissolved in acqueosus MES/KCl buffer and in DMSO, so that to obtain 10 μM stocks.

Instrumentations:

Fluorescence emission spectra were registered at 298K using a Varian Cary Eclipse Fluorescence spectrophotometer. The excitation and emission slits width was set to 5nm, while the excitation wavelength and the emission range where suitably changed. The emission PMT (PhotoMultiplier Tubes) detector voltage was modulated from 450 to 800 V. In particular for the spectra of **24a**, **24b**, **24c** in DMSO the excitation wavelength was set to 318 nm, the emission range was set from 325 to 550 nm and the emission PMT voltage was fixed to 500 V. For the spectra of **24a**, **24b**, **24c** in MES-KCl buffer the excitation wavelength was set to 310 nm, the emission range was set from 320 to 550 nm and the emission PMT voltage was fixed to 500 V. For the spectra of **21** in DMSO the excitation wavelength was set to 309 nm, the emission range was set from 320 to 550 nm and the emission PMT voltage was fixed to 450 V. For the spectra of **21** in MES-KCl buffer the excitation wavelength was set to 310 nm, the emission range was set from 320 to 550 nm and the emission PMT voltage was fixed to 800 V.

5.2.5. Fluorescence Resonance Energy Transfer (FRET) assays.

Oligonucleotides.

Three different fluorescent conjugated oligonucleotides were used as models for intramolecular G-quadruplex (F21T, c-kit1 and c-kit2) and one as model for duplex DNA (F10D). Both of them were purchased from Oswel (Southampton, UK). DNA was initially dissolved as a stock 20 μM solution in purified water; further dilutions were carried out in the relevant buffer

Oligonucleotide sequences:

G4-TEL "*F21T*": 5'-FAM-dGGG(TTAGGG)$_3$-TAMRA-3';

G4-KIT "*KIT1*": 5'-FAM-d[AGAG$_3$AG$_2$GCGCTG$_3$AG$_2$AG$_3$GCT]-TAMRA-3′;

G4-KIT "*KIT2*": 5′-FAM-d[C$_3$G$_3$CG$_3$CGCGAG$_3$AG$_4$AG$_2$]-TAMRA-3′;

Duplex DNA "*F10D*": 5'-FAM-dTATAGCTATA-(CH$_2$-CH$_2$-O)$_6$-dTATAGCTATA-TAMRA-3'

Drugs stocks preparation.

Drugs were dissolved in bi-distilled water, so that to obtain 1mM stocks; further dilutions were carried out in 50mM potassium cacodylate buffer (pH 7.4). Stocks were used within 72 hours from their preparation. They were kept at +4°C, avoiding light exposure.

FRET stabilisation assay.

The ability of the compounds to stabilize intramolecular G-quadruplex and duplex DNA was investigated using a FRET assay modified to be used in a 96-well format. Oligonucleotides were diluted in 50mM potassium cacodylate buffer (pH 7.4) to the 2x concentration (400nM) and then annealed by heating to 85°C for 5 minutes, followed by cooling to room temperature in the heating block. 96-well plates (MJ Research, Waltham, MA) were prepared by aliquoting 50μL of the annealed DNA to each well, followed by 50μL of the compound solutions. Samples were left equilibrating for half an hour before starting the experiment. Measurements were made on a DNA Engine Opticon (MJ Research) with excitation at 450-495nm and detection at 515-545nm. Fluorescence readings were taken at intervals of 0.5°C over the range 30-100°C, with a constant temperature being maintained for 30 seconds prior to each reading to ensure a stable value. Final analysis of the data was carried out using a script written in the program Origin 7.0 (OriginLab Corp., Northampton, MA).

Drug concentrations were chosen so that to have the final drug/DNA ratios (R) as reported in the following table.

conc.(μM)	0.10	0.50	1.00	2.00	3.00	4.00	5.00
R	**0.50**	**2.25**	**5.00**	**10.00**	**15.00**	**20.00**	**25.00**

All the experiments were carried out in triplicate. Maximum drug concentrations were proven not to have enough fluorescence on their own to disturb fluorescence of the probes. Where necessary methanol stocks were used, with a final methanol concentration less than 1% not affecting DNA stability.

6. References

- Aboul-ela, F.; Murchie, A.L.; Lilley, D.M. "NMR study of parallel-stranded tetraplex formation by the hexadeoxynucleotide d(TG_4T)" *Nature* **1992**, *360*, 280.
- Agrawal, P.; Barthwal, S. K.; Barthwal, R. "Studies on self-aggregation of anthracycline drugs by restrained molecular dynamics approach using nuclear magnetic resonance spectroscopy supported by absorption, fluorescence, diffusion ordered spectroscopy and mass spectrometry" *Europ. J. of Med. Chem.* **2009**, *44(4)*, 1437-1451.
- Alvino, A.; Franceschin, M.; Cefaro, C.; Borioni, S.; Ortaggi, G.; Bianco, A. "Synthesis and spectroscopic properties of highly water-soluble perylene derivatives" *Tetrahedron* **2007**, *63*, 7858.
- Ambrus, A.; Chen, D.; Dai, J.X.; Bialis, T.; Jones, R.A.; Yang, D.Z. "Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution" *Nucleic Acids Res.* **2006**, *34*, 2723-35.
- Ambrus, A.; Chen, D.; Dai, J.; Jones, R. A.; Yang, D. "Solution structure of the biologically relevant G-quadruplex element in the human c-myc promoter. Implications for G-quadruplex stabilization" *Biochemistry* **2005**, *44*, 2048–58.
- Arora, A.; Dutkiewicz, M.; Scaria, V.; Hariharan, M.; Maiti, S.; Kurreck, J. "Inhibition of translation in living eukaryotic cells by an RNA G-quadruplex motif " *RNA* **2008**, 14, 1290-6.
- Azzalin, C. M.; Reichenbach, P.; Khoriauli, L.; Giulotto, E.; Lingner, J. "Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends" *Science* **2007**, 318, 798-801.

- Bang, I. “Untersuchungen u¨ber die Guanylsäre“ *Bioch. Ztschr.* **1910**, 26, 293-311.
- Binz, N.; Shalaby, T.; Rivera, P.; Shin-ya, K.; Grotzer, M. A. “Telomerase inhibition, telomere shortening, cell growth suppression and induction of apoptosis by telomestatin in childhood neuroblastoma cells” *Eur. J. Cancer* **2005**, *41*, 2873–2881.
- Blackburn, E. H. “Telomere states and cell fates” *Nature* **2000**, *408*, 53–56.
- Blackburn, E.H. “Structure and function of telomeres” *Nature* **1991**, 350, 569–73.
- Blasco, M.A. ” Telomeres and human disease: ageing, cancer and beyond” *Nat. Rev. Genet.* **2005**, *6*, 611-622.
- Carpino, L. A.; Sadat-Aalaee, D.; Chao, H. G.; De Selms, R. H. “[(9-Fluorenylmethyl)oxy]carbonyl (FMOC) amino acid fluorides. Convienient new peptide coupling reagents applicable to the FMOC/tert-butyl strategy for solution and solid-phase syntheses” *J. of the Am. Chem. Soc.* **1990**, *112(26)*, 9651-2.
- Casagrande, V.; Alvino, A.; Bianco, A.; Ortaggi, G.; Franceschin, M. "Study of binding affinity and selectivity of perylene and coronene derivatives towards duplex and quadruplex DNA by ESI-MS” *J. Mass Spectrom.* **2009**, *44(4)*, 530-40.
- Cesare, A.J.; Griffith, J.D. ” Telomeric DNA in ALT cells is characterized by free telomeric circles and heterogeneous t-loops” *Mol. Cell. Biol* **2004**, *24*, 9948-57.
- Chen, Q.; Kuntz, I.D.; Shafer, R.H. “Spectroscopic recognition of guanine dimeric hairpin quadruplexes by a carbocyanine dye” *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 2635-9.
- Dai, J.; Carver, M.; Yang, D. ”Polymorphism of human telomeric quadruplex structures” Biochimie **2008**, *90*,1172-83.
- Dai, J.; Chen, D.; Jones, R.A.; Hurley, L.H.; Yang, D. “NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region” *Nucleic Acids Res.* **2006**, 34, 5133–44.
- Dai, J.X.; Carver, M.; Punchihewa, C.; Jones, R.A.; Yang, D.Z. “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*, 4927-40.
- Davis, J. T.; Spada, G. P. “Supramolecular architectures generated by self-assembly of guanosine derivatives” *Chem. Soc. Rev.* **2007**, *36*, 296–313.
- De Armond, R.; Wood, S.; Sun, D.; Hurley, L.H.; Ebbinghaus, S.W. “Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1alpha promoter” *Biochemistry* **2005**, 44, 16341–50.

- De Cian, A.; Grellier, P.; Mouray, E.; Depoix, D.; Bertrand, H.; Monchaud, D.; Teulade-Fichou, M.P.; Mergny, J.L.; Alberti, P. "Plasmodium telomeric sequences: structure, stability and quadruplex targeting by small compounds" *Chembiochem* **2008**, 9 ,2730-9.
- De Cian, A.; Gros, J.; Guédin, A.; Haddi, M.; Lyonnais, S.; Guittat, L.; Riou, J.F.; Trentesaux, C.; Saccà, B.; Lacroix, L.; Alberti, P.; Mergny, J.L.; "DNA and RNA quadruplex ligands" *Nucleic Acids Symp Ser (Oxf.)* **2008**, *52*, 7-8.
- De Cian, A.; Lacroix, L.; Douarre, C.; Temime-Smaali, N.; Trentesaux, C.; Riou, J.F.; Mergny, J.L. " A new steroid derivative stabilizes g-quadruplexes and induces telomere uncapping in human tumor cells" *Biochimie*, **2007**, *72*, 631-40 .
- De Lange, T. " Shelterin: the protein complex that shapes and safeguards human telomeres" *Genes and Development* **2005,** *19*, 2100-10.
- Dexheimer, T.S.; Sun, D.; Hurley, L.H. "Deconvoluting the structural and drug-recognition complexity of the G-quadruplex-forming region upstream of the bcl-2 P1 promoter" J. Am. Chem. Soc. **2006**, 128, 5404–15.
- Doi, T.; Yoshida, M.; Shin-ya, K.; Takahashi, T. "Total synthesis of (R)-telomestatin" *Org. Lett.* **2006**, *8*, 4165–4167.
- Eissenstat, M. A.; Bell, M. R.; D'Ambra, T. E.; Alexander, E. J.; Daum, S. J.; Ackerman, J. H.; Gruett, M. D.; Kumar V.; Estep, K. G. " Aminoalkylindoles: structure-activity relationships of novel cannabinoid mimetics" *J. Med. Chem.* **1995**, *38*, 3094.
- Evans, D.; I. Lockhart, M. "The Schmidt reaction with aromatic ketones." *J. Chem. Soc.*, **1965**, 4806-12.
- Feng, J.; Funk, W.D.; Wang, S.S.; Weinrich, S.L.; Avilion, A.A, Chiu, C.P.; Adams, R.R.; Chang, E.; Allsopp, R.C.; Yu, J.; Le, S.Y.; West, M.D.; Harley, C.B.; Andrews, W.H.; Greider, C.W.; Villeponteau, B. " The RNA component of human telomerase" *Science* **1995**, *269*, 1236-41.
- Franceschin, M. " G-Quadruplex DNA structures and organic chemistry: more than one connection" *Eur. J. Org. Chem.* **2009**, *(14)*, 2225-38.
- Franceschin, M.; Alvino, A.; Casagrande, V.; Mauriello, C.; Pascucci, E.; Savino, M.; Ortaggi, G.; Bianco A. "Specific interactions with intra- and intermolecular G-quadruplex DNA structures by hydrosoluble coronene derivatives: a new class of telomerase inhibitors" *Bioorg. Med. Chem.* **2007**, *15*,1848-58.
- Franceschin, M.; Ginnari-Satriani, L.; Alvino, A.; Ortaggi,G.; Bianco, A. "Study of a convenient method for the preparation of hydrosoluble fluorescent triazatruxene derivatives" *Eur. J. Org. Chem.*, **2009**, DOI: 10.1002/ejoc.200900869.

- Franceschin, M.; Lombardo, C.M.; Pascucci, E.; D'Ambrosio, D.; Micheli, E.; Bianco, A.; Ortaggi, G.; Savino, M." The number and distances of positive charges of polyamine side chains in a series of perylene diimides significantly influence their ability to induce G-quadruplex structures and inhibit human telomerase" *Bioorg. Med. Chem.* **2008**, *16*, 2292-2304.
- Franceschin, M.; Pascucci, E.; Alvino, A.; D'Ambrosio, D.; Bianco, A.; Ortaggi, G.; Savino, M. "New highly hydrosoluble and not self-aggregated perylene derivatives with three and four polar side-chains as G-quadruplex telomere targeting agents and telomerase inhibitors" *Bioorg. Med. Chem. Lett.* **2007**, *17*, 2515–22.
- Franceschin, M.; Rossetti, L.; D'Ambrosio, A.; Schirripa, S.; Bianco, A.; Ortaggi, G.; Savino, M.; Schultes, C.; Neidle, S. "Natural and synthetic G-quadruplex interactive berberine derivatives" *Bioorganic & Medicinal Chemistry Letters* **2006**, *16*, 1707-11.
- Gabelica, V.; De Pauw, E.; Rosu, F. "Interaction between antitumor drugs and a double-stranded oligonucleotide studied by electrospray ionization mass spectrometry. *Journal of Mass Spectrometry* **1999**, *34*, 1328.
- Gabelica, V.; Rosu, F.; Houssier, C.; De Pauw, E. "Gas phase thermal denaturation of an oligonucleotide duplex and its complexes with minor groove binders" *Rapid Communications in Mass Spectrometry* **2000**, *14*, 464.
- Garcia-Frutos, E. M.; Gomez-Lor, B. " Synthesis and Self-Association Properties of Functionalized C3-Symmetric Hexakis(p-substituted-phenylethynyl)triindoles. *J. Am. Chem. Soc.* **2008**, *130*, 9173-7.
- Garcia-Frutos, E. M.; Gutierrez-Puebla, E.; Monge, M. A.; Ramirez, R.; de Andres, P.; de Andres, A.; Ramirez, R.; Gomez-Lor, B. "Crystal structure and charge-transport properties of N-trimethyltriindole: Novel p-type organic semiconductor single crystals" *Organic Electronics* **2009**, *10(4)*, 643-52.
- Gellert, M.; Lipsett, M. N.; Davies, D. R. "Helix formation by guanylic acid" *Proc. Natl Acad. Sci. USA* **1962**, *48*, 2013–2018.
- Gillis, A.J.; Schuller, A. P.; Skordalakes, E. "Structure of the Tribolium castaneum telomerase catalytic subunit TERT" *Nature*" **2008**, *455*, 633-637.
- Ginnari-Satriani, L.; Casagrande, V.; Bianco, A.; Ortaggi, G.; Franceschin, M. "A hydrophilic three side-chained triazatruxene as a new strong and selective G-quadruplex ligand" *Org. Biomol. Chem.* **2009**, *7(12),* 2513-6.
- Gomez, D.; Aouali, N.; Renaud, A.; Douarre, C.; Shin-Ya, K.; Tazi, J.; Martinez, S.; Trentesaux, C.; Morjani, H.; Riou, J. F. "Resistance to senescence induction and telomere

shortening by a G-quadruplex ligand inhibitor of telomerase" *Cancer Res.* **2003**, *63*, 6149–53.

- Gomez, D.; Mergny, J. L.; Riou, J. F. "Detection of telomerase inhibitors based on g-quadruplex ligands by a modified telomeric repeat amplification protocol assay" *Cancer Res.* **2002**, *62*, 3365–8.
- Gomez-Lor B.; Echavarren, A. M. "Synthesis of a triaza analogue of crushed-fullerene by intramolecular palladium-catalyzed arylation" *Org. Lett.* **2004**, *6*, 2993-6.
- Gomez-Lor, B.; Hennrich, G.; Alonso, B.; Monge, A.; Gutierrez-Puebla E.; Echavarren, A. M. "A redox-active C3-symmetric triindole-based triazacyclophane" *Angew. Chem., Int. Ed.* **2006**, *45*, 4491.
- Gornall, K. C.; Samosorn, S.; Talib, J.; Bremner, J. B.; Beck, J. L. "Selectivity of an indolyl berberine derivative for tetrameric G-quadruplex DNA" *Rapid Commun. Mass Spectrom.* **2007**, *21*, 1759–66.
- Grand, C.L.; Han, H.; Mu~noz, R.M.; Weitman, S.; Von Hoff, D.D.; Hurley, L.H.; Bearss, D.J. "The cationic porphyrin TMPyP4 downregulates c-MYC and human telomerase reverse transcriptase expression and inhibits tumor growth in vivo" *Mol. Cancer Ther.* **2002**, *1*, 565-73
- Grayson, I. "Water-soluble carbodiimide - an efficient coupling agent for synthesis" *Speciality Chemicals* **2000**, *20(3)*, 86-88.
- Griffith, J.D.; Comeau, L.; Rosenfield, S.; Stansel, R.M.; Bianchi, A.; Moss, H.; de Lange, T." Mammalian telomeres end in a large duplex loop" *Cell* **1999**, *97*, 503-514.
- Guittat, L.; De Cian, A.; Rosu, F.; Gabelica, V.; De Pauw, E.; Delfourne, E.; Mergny J. L. "Ascididemin and meridine stabilise G-quadruplexes and inhibit telomerase in vitro" *Biochimica et Biophysica Acta* **2005**, *1724*, 375.
- Guyen, B.; Schultes, C. M.; Hazel, P.; Mann J.; Neidle S. "Synthesis and evaluation of analogues of 10*H*-indolo[3,2-*b*]-quinoline as G-quadruplex stabilising ligands and potential inhibitors of the enzyme telomerase" *Org Biomol Chem.* **2004**; *2(7)*, 981-8.
- Haider, S.M.; Parkinson, G.N.; Neidle, S. *J. Mol. Biol.*" Crystal structure of the potassium form of an Oxytricha nova G-quadruplex" **2002**, *320*, 189-200.
- Han, H.; Cliff, C.L.; Hurley, L.H. "Accelerated assembly of G-quadruplex structures by a small molecule" *Biochemistry* **1999**, *38*, 6981-6.
- Harley, C.B.; Futcher, A.B.; Greider, C.W. "Telomeres shorten during ageing of human fibroblasts" *Nature* **1990**, *345*, 458-60.

- Held, D. M.; Kissel, J. D.; Patterson, J. T.; Nickens, D. G.; Burke, D. H. "HIV-1 inactivation by nucleic acid aptamers" *Front. Biosci.* **2006**, *11*, 89–112.
- Hiyama, E.; Hiyama, K. " Telomere and telomerase in stem cells" *British Journal of Cancer* **2007**, *96*, 1020-4.
- Hiyoshi, H.; Kumagai, H.; Ooi, H.; Sonoda, T.; Mataka, S. "Donor-π-acceptor type symmetric cyclic triindoles: synthesis and properties" *Heterocycles* **2007**, *72*, 231-8.
- Hodes, R."Molecular targeting of cancer: telomeres as targets" *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 7649-51.
- Howell, R.M.; Woodford, K.J.; Weitzmann, M.N.; Usdin, K. "The chicken b-globin gene promoter forms a novel 'cinched' tetrahelical structure", *J. Biol. Chem.* **1996**, *271*, 5208-14.
- Hsu, S.T.; Varnai, P.; Bugaut, A.; Reszka, A.P.; Neidle, S., Balasubramanian, S. A "G-rich sequence within the c-kit oncogene promoter forms a parallel G-quadruplex having asymmetric G-tetrad dynamics" *J. Am. Chem. Soc*. **2009**, *131(37)*, 13399-409.
- Huppert, J. L.; Bugaut, A.; Kumari, S.; Balasubramanian, S. "G-quadruplexes: the beginning and end of UTRs" *Nucleic Acids Res.* **2008**, *36*, 6260-8.
- Huppert, J.L. "Hunting G-quadruplexes" *Biochimie* **2008** ,*90*, 1140-8.
- Huppert, J.L. "Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes" *Chem. Soc Rev.* **2008**, *37*, 1375-84.
- Jantos, K.; Rodriguez, R.; Ladame, S.; Shirude, P. S.; Balasubramanian, S." Oxazole-based peptide macrocycles: a new class of G-quadruplex binding ligands" *J. Am. Chem. Soc.* **2006**, *128*, 13662–3.
- Jing, N.; Marchand, C.; Liu, J.; Mitra, R.; Hogan, M. E.; Pommier, Y. *J.* "Mechanism of inhibition of HIV-1 integrase by G-tetrad-forming oligonucleotides in Vitro" *Biol. Chem.* **2000**, *275*, 21460–67.
- Karimata, H.; Miyoshi, D.; Fujimoto, T.; Koumoto, K.; Wang, Z. M.; Sugimoto, N. " Conformational switch of a functional nanowire based on the DNA G-quadruplex" *Nucleic Acids Symp. Ser.* **2007**, *51*, 251– 252.
- Kaucher, M. S.; Harrell Jr., W. A.; Davis, J. T.; *J. Am. Chem. Soc.* "A unimolecular G-quadruplex that functions as a synthetic transmembrane Na+ transporter" **2006**, *128*, 38–39.
- Kim, K.K.; Jang, J.G. "Carbazole and indoline derivatives and organic light emitting device using same" *U.S. Pat. Appl. Publ.* **2006**, US2006063037.
- Kim, N.W.; Piatyszek, M.A.; Prowse, K.R.; Harley, C.B.; West, M.D.; Ho, P.L.C.; Coviello, G.M.; Wright, W.E.; weinrich, S.L.; Shay, J.W. "Specific association of human telomerase activity with immortal cells and cancer" *Science* **1994**, *266*, 2011-5.

- Lai, W. Y.; He, Q. Y.; Zhu, R.; Chen Q. Q.; Huang, W. ”Kinked star-shaped fluorene/triazatruxene co-oligomer hybrids with enhanced functional properties for high-performance, solution-processed, blue organic light-emitting diodes” *Adv. Funct. Mater.*, 2008, *18*, 265.
- Li, J.; Correia, 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, 4649-59.
- Lingner, J.; Cech, T. R. “Telomerase and chromosome end maintenance” *Curr. Opin. Genet. Dev.* **1998**, *8*, 226–32.
- Lingner, J.; Cech, T.R. “Telomerase and chromosome end maintenance” *Curr. Opin Genet. Dev.* **1998**, *8(2)*, 226-32.
- Mashima, T.; Matsugami, A.; Nishikawa, F.; Nishikawa, S.; Katahira, M. “Unique quadruplex structure and interaction of an RNA aptamer against bovine prion protein”, *Nucleic Acids Research* **2009**, *in press*.
- Mazzitelli, C. L.; Brodbelt, J. S.; Kern, J. T.; Rodriguez, M.; Kerwin, S. M. “Evaluation of binding of perylene diimide and benzannulated perylene diimide ligands to DNA by electrospray ionization mass spectrometry” *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 593–604.
- Mergny, J. L.; De Cian, A.; Ghelab, A.; Sacca, B.; Lacroix, L.; “Kinetics of tetramolecular quadruplexes” *Nucleic Acids Res.* **2005**, *33*, 81-94.
- Mergny J. L.; Maurizot J. C. “Fluorescence resonance energy transfer as a probe for G-quartet formation by a telomeric repeat” *Chembiochem.* **2001**; 2(2), 124-32.
- Micheli, E.; Lombardo, C.M.; D’Ambrosio, D.; Franceschin, M.; Neidle, S.; Savino, M. ” Selective G-quadruplex ligands: the significant role of side chain charge density in a series of perylene derivatives” *Bioorg. Med. Chem. Lett.* **2009**, *19*, 3903-8.
- Monks, A.; Scudiero, D. A.; Skehan, P.; Shoemaker, R. H.; Paull, K. D.; Vistica, D. T.; Hose, C.; Langley, J.; Cronice, P.; Vaigro-Wolf, M.; Gray-Goodrich, M.; Campbell, H.; Mayo, M. R. Feasibility of a high-flux anticancer drug screen using a diverse panel of cultured human tumor cell lines“ *JNCI, J. Natl. Cancer Inst.* **1991**, *83*, 757-766.
- Müller, S. ; Pantos¸ G. D. ; Rodriguez, R. ; Balasubramanian, S. “Controlled-folding of a small molecule modulates DNA G-quadruplex recognition” *Chem. Commun.* **2009**, 80–82.
- Naasani, I.; Seimiya, H.; Yamori, T.; Tsuruo, T. "FJ5002: a potent telomerase inhibitor identified by exploiting the disease-oriented screening program with COMPARE analysis” *Cancer Res.* **1999**, *59*, 4004–11.

- Neidle S. “The structures of quadruplex nucleic acids and their drug complexes” *Curr. Opin. Struct. Biol.* **2009**, 19(3), 239-50.
- Neidle, S.; Harrison, R. J.; Reszka, A. P.; Read, M. A. “Structure-activity relationships among guanine-quadruplex telomerase inhibitors” *Pharmacol. Ther.* **2000**, *85*, 133–9.
- Neidle, S.; Parkinson, G.N. “Quadruplex DNA crystal structures and drug design” *Biochimie* **2008**; 90, 1184-96.
- Norton, J.C.; Piatyszek, M.A.; Wright, W.E.; Shay, J.W.; Corey, D.R. ” Inhibition of human telomerase activity by peptide nucleic acids” *Nature Biotechnol.* **1996**, *14*, 615-9.
- Ortaggi, G.; Bianco, A.; Franceschin, M.; Ginnari-Satriani, L.; Casagrande V. “Ciclotrimeri simmetrici di eterocicli aromatici loro procedimenti di preparazione e loro uso in campo farmaceutico” *Patent no.* **RM2009A000197**, 24 Aprile 2009.
- Otero, G. ; Biddau, G.; Sanchez-Sanchez, C. ; Caillard, R. ; Lopez, M. F. ; Rogero, C.; Palomares, F. J.; Cabello, N.; Basanta, M. A.; Ortega, J.; Mendez, J.; Echavarren, A.M.; Perez, R.; Gomez-Lor B.; Martin-Gago, J. A. “Fullerenes from aromatic precursors by surface-catalysed cyclodehydrogenation” *Nature* **2008**, *454*, 865-8.
- Pagano, B.; Giancola, C. ”Energetics of quadruplex-drug recognition in anticancer therapy” *Curr. Cancer Drug Targ.* **2007**, *7*, 520-40.
- Pagano, B.; Mattia, C. A.; Virno, A.; Randazzo, A.; Mayol, L.; Giancola, C. “Thermodynamic analysis of quadruplex DNA-drug interaction” *Nucleosides, Nucleotides & Nucleic Acids* **2007**, *26(6-7)*, 761-5.
- Palumbo, S.L.; Memmott, R.M.; Uribe, D.J.; Krotova-Khan, Y.; Hurley, L.H.; Ebbinghaus, S.W. “A novel G-quadruplex forming GGA repeat region in the c-myb promoter is a critical regulator of .promoteractivity” *Nucleic Acids Res.* **2008**, 36(6), 1755-69
- Parkinson, G. N.; Ghosh, R.; Neidle, S.” Structural basis for binding of porphyrin to human telomeres” *Biochemistry* **2007**, *46*, 2390–2397.
- Parkinson, G.N.; Lee, M.P.; Neidle, S. “Crystal structure of parallel quadruplexes from human telomeric DNA” *Nature* **2002**, *417*, 876-80.
- Phan, A.T.; Kuryavyi, V.; Burge, S.; Neidle, S.; Patel, D.J. “Structure of an unprecedented G-quadruplex scaffold in the human c-kit promoter” *J. Am. Chem. Soc.* **2007**, 129, 4386–92.
- Phan, A.T.; Modi, Y.S.; Patel, D.J. “Propellor-type parallel-stranded Gquadruplexes in the human c-myc promoter” *J. Am. Chem. Soc.* **2004**, *126*, 8710-16.
- Phillips, K.; Dauter, Z.; Murchie, A.J.; lilley, D.M.; Luisi, B. “The crystal structure of a parallel-stranded guanine tetraplex at 0.95 A resolution” *J. Mol. Biol.* **1997**, *273*, 171.

- Qin, Y.; Rezler, E.M.; Gokhale, V.; Sun, D.; Hurley, L.H. "Characterization of the G-quadruplexes in the duplex nuclease hypersensitive element of the PDGF-A promoter and modulation of PDGF-A
- Rajeswari, S. ; Jones, R. J.; Cava, M. P. "A new synthesis of amides from acyl fluorides and N-silylamines" *Tetrahedron Letters* **1987**, *28(43)*, 5099-102.
- Randazzo, A.; Galeone, A.; Esposito, V.; Varra, M.; Mayol, L." Interaction of distamycin A and netropsin with quadruplex and duplex structures: a comparative 1H-NMR study" *Nucleosides Nucleotides Nucleic Acids* **2002**, *21*, 535–45.
- Rankin, S.; Reszka, A.P.; Huppert, J.; Zloh, M.; Parkinson, G.N.; Todd, A.K.; Ladame, S.; Balasubramanian, S.; Neidle, S. "Putative DNA quadruplex formation within the human c-kit oncogene" *J. Am. Chem. Soc.* **2005**, 127, 10584–89.
- Read, M. A.; Harrison, R. J.; Romagnoli, B.; Tanious, F. A.; Gowan, S. H.; Reszka, A. P.; Wilson, W. D.; Kelland, L. R.; Neidle, S. "Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors" *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 4844–9.
- Rezler, E. M.; Seenisamy, J.; Bashyam, S.; Kim, M. Y.; White, E.; Wilson, W. D.; Hurley, L. H Telomestatin and diseleno sapphyrin bind selectively to two different forms of the human telomeric G-quadruplex structure" *J. Am. Chem. Soc.* **2005**, *127*, 9439–9447.
- Rhodes, D.; Giraldo, R. "Telomere structure and function" *Curr. Opin. Struct. Biol.* **1995**, *5*, 31.
- Robertson, N.; Parsons, S.; MacLean, E. J.; Coxall R. A.; Mount, A. R." Preparation, X-ray structure and properties of a hexabrominated, symmetric indole trimer and its TCNQ adduct: a new route to functional molecular systems" *J. Mater. Chem.* **2000**, *10*, 2043-7.
- Rossetti, L.; Franceschin, M.; Schirripa, S.; Bianco, A.; Ortaggi, G.; Savino, M. "Selective interactions of perylene derivatives having different side chains with inter- and intramolecular G-quadruplex DNA structures. A correlation with telomerase inhibition" *Bioorg. Med. Chem. Lett*. **2005**, *15*, 413.
- Rosu, F.; De Pauw, E.; Gabelica, V. "Electrospray mass spectrometry to study drug-nucleic acids interactions" *Biochimie* **2008**, *90*, 1074– 87.
- Rosu, F. ; De Pauw, E. ; Guittat, L. ; Alberti, P. ; Lacroix, L. ; Mailliet, P. ; Riou, J. F. ; Mergny J. L. "Selective interaction of ethidium derivatives with quadruplexes: an equilibrium dialysis and electrospray ionization mass spectrometry analysis" *Biochemistry* **2003**, *42*, 10361.

- Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; D. Pauw E. "Triplex and quadruplex DNA structures studied by electrospray mass spectrometry" *Rapid Communications in Mass Spectrometry* **2002**,*16*, 1729.
- Rubinstein, L. V.; Shoemaker, R. H.; Paull, K. D.; Simon, R. M.; Tosini, S.; Skehan, P.; Scudiero, D. A.; Monks, A.; Boyd, M. R. "Comparison of in vitro anticancer-drug-screening data generated with a tetrazolium assay versus a protein assay against a diverse panel of human tumor cell lines" *JNCI, J. Natl. Cancer Inst.* **1990**, *82,* 1113-1118.
- Saretzki, G. " Telomerase inhibition as cancer therapy" *Cancer Lett.* **2003**, *194*, 209-19.
- Seenisamy, J.; Bashyam, S.; Gokhale, V.; Vankayalapati, H.; Sun, D.; Siddiqui-Jain, A.; Streiner, N.; Shin-Ya, K.; White, E.; Whilson, W. D.; Hurley, L. H. "Design and synthesis of an expanded porphyrin that has selectivity for the c-MYC G-quadruplex structure." *J. Am. Chem. Soc.* **2005**, *127*, 2944–59.
- Sen, D.; Gilbert, W. "Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis" *Nature* **1988**, 334, 364–366.
- Shammas, M. A.; Shmookler Reis, R. J.; Li, C.; Koley, H.; Hurley, L. H.; Anderson, K. C.; Munshi, N. C. "Telomerase inhibition and cell growth arrest after telomestatin treatment in multiple myeloma" *Clin. Cancer Res.* **2004**, *10*, 770–6.
- Shay, J.W; Wright, W.E. "Telomerase therapeutics for cancer: challenges and new directions" *Nat. Rev. Drug Discov.* **2006**, *5*, 577-84.
- Shirude, P. S.; Gillies, E. R.; Ladame, S.; Godde, F.; Shin-Ya, K.; Huc, I.; Balasubramanian, S. "Macrocyclic and helical oligoamides as a new class of G-quadruplex ligands" *J. Am. Chem. Soc.* **2007**, *129*, 11890– 91.
- Shirude, P.S.; Okumus, B.; Ying, L.; Ha, T., Balasubramanian, S. "Single-molecule conformational analysis of G-quadruplex formation in the promoter DNA duplex of the proto-oncogene c-kit" *J. Am. Chem. Soc.* 2007 ,129, 7484–5.
- 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 USA* **2002**, 99, 11593–8.
- Simonsson, T. "LinksG-quadruplex DNA structures: variations on a theme" **Biol Chem. 2001**, 382, 621-8.
- Simonsson, T.; Pecinka, T.; Kubista, M. " DNA tetraplex formation in the control region of c-myc" *Nucleic Acid Res.* **1998**, *26*, 1167-72.
- Songyang, Z.; Liu, D. "Inside the mammalian telomere interactome: regulation and regulatory activities of telomeres" *Crit. Rev. Eukaryot. Gene Expr.* **2006**, *16*, 103-118.

- Sun, D.; Guo, K.; Rusche, J.J.; Hurley, L.H. "Facilitation of a structural transition in the polypurine/
- 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*, 2113-6.
- Talarico, M.; Termine, R.; Garcia-Frutos, E. M.; Omenat, A.; Serrano, J. L.; Gomez-Lor, B.; Golemme, A." New Electrode-Friendly Triindole Columnar phases with High Hole Mobility. *Chem. Mater.*, 2008, **20**, 6589.
- Tauchi, T.; Shin-ya, K.; Sashida, G.; Sumi, M.; Okabe, S.; Ohyashiki, J. H.; Ohyashiki, K. " Telomerase inhibition with a novel G-quadruplex-interactive agent, telomestatin: in vitro and in vivo studies in acute leukemia" *Oncogene* **2006**, *25*, 5719–25.
- Teng, Y.; Girvan, A. C.; Casson, L. K.; Pierce Jr., W. M.; Qian, M.; Thomas, S. D.; Bates, P. J. "AS1411 alters the localization of a complex containing protein arginine methyltransferase 5 and nucleolin" *Cancer Res.* **2007**, 67, 10491-500.
- Tera, M.; Iida, K.; Ishizuka, H.; Takagi, M.; Suganuma, M.; Doi, T. Shin-ya, K.; Nagasawa, K. "Synthesis of a potent G-quadruplex-binding macrocyclic heptaoxazole" *Chembiochem* **2009**; *10*, 431-5.
- Tera, M.; Ishizuka, H.; Takagi, M.; Suganuma, M.; Shin-ya, K.; Nagasawa, K. " Macrocyclic hexaoxazoles as sequence- and mode-selective G-quadruplex binders" *Angew. Chem. Int. Ed.* **2008**, *47*, 5557–60.
- Ulaner, G.A. " Telomere maintenance in clinical medicine" *Am. J. Med.* **2004**, *117*, 262-9.
- Urata, H.; Kumashiro, T.; Otake, T.; Kawahata, T.; and Akagi, M. "Anti-HIV-1 activity of L-DNA quadruplex" *Nucleic Acids Research* **2002**, *2*, 163-4.
- Wang, Y.; Patel, D.J. " Solution structure of the human telomeric repeat d[$AG_3(T_2AG_3)_3$] G-tetraplex" *Structure* **1993**, *1*, 263-82.
- Wei, C.; Jia, G.; Yuan, J.; Feng, Z.; Li, C. "A Spectroscopic Study on the Interactions of Porphyrin with G-Quadruplex DNAs" *Biochemistry* **2006**, *45*, 6681–6691.
- White, L. K.; Wright, W. E.; Shay, J. W. " Telomerase inhibitors" *Trends Biotechnol.* **2001**, *19*, 114-20.
- Williamson, J. R. **"**Guanine quartets" *Curr. Opin. Struct. Biol.* **1993**, 3(3), 357-62.
- Wright, W. E.; Pereira-Smith, O. M.; Shay, J. W. "Reversible cellular senescence: a two-stage model for the immortalization of normal human diploid fibroblasts" *Mol. Cell. Biol.* **1989**, 9, 3088–92.

- Wright, W.E.; Pereira-Smith, O.M.; Shay, J.W. ”Reversible cellular senescence: implications for immortalization of normal human diploid fibroblasts” *Mol. Cell. Biol.* **1989**, *9*, 3088-92.
- Yang, D.; Hurley, L.H. “Structure of the biologically relevant G-quadruplex in the c-MYC promoter” *Nucleosides, nucleotides & nucleic acids* **2006**, *25(8)*, 951-68.
- Yang, J.; Du, X.; Lazar, A. J.; Pollock, R.; Hunt, K.; Chen, K.; Hao, X.; Trent, J.; Zhang, W. “Genetic aberrations of gastrointestinal stromal tumors” *Cancer* **2008**, *113*, 1532–43.
- Zahler, A.M.; Williamson, J.R.; Cech, T.R.; Prescott, D.M. “Inhibition of telomerase by G-quartet DNA structures” *Nature* **1991**, *350*, 718-20.