UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Università degli Studi di Padova

Dipartimento di Scienze del Farmaco

SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE MOLECOLARI
INDIRIZZO: Scienze Farmaceutiche - XXVII CICLO

SYNTHESIS, CHARACTERIZATION AND BIOPHYSICAL EVALUATION OF NOVEL G-QUADRUPLEX STABILIZING AGENTS

Direttore della Scuola: Ch.mo Prof. Antonino Polimeno
Coordinatore d'indirizzo: Ch.mo Prof. Stefano Moro
Supervisore: Ch.mo Prof. Giuseppe Zagotto

Dottorando: Giovanni Ribaudo

Alla mia famiglia

INDEX

ABSTRACT [ITALIAN]

Sequenze di DNA particolarmente ricche di guanine e potenzialmente in grado di formare strutture di tipo G-quadruplex sono diffuse sia nel genoma umano che in quello di altre specie.

Il telomero ne rappresenta un esempio ormai largamente discusso e costituisce tuttora un interessante obiettivo nella strategia antitumorale basata sulla inibizione indiretta della telomerasi tramite stabilizzazione di G-quadruplex. Va comunque considerato che nel genoma umano sono state individuate oltre 376000 sequenze di questo tipo, ossia con la peculiarità di essere ricche in guanine, con una localizzazione preferenziale in alcune regioni rappresentate dai proto-oncogeni. In questo ambito i G-quadruplex potrebbero agire come interruttori di accensione e spegnimento, o di regolazione, della trascrizione di tali sequenze; il DNA strutturato generalmente non viene infatti processato dagli enzimi coinvolti. I G-quadruplex sono stati descritti in letteratura quali regolatori di molti processi cellulari di rilievo come l'allineamento cromosomico, la replicazione, la trascrizione e la ricombinazione del genoma. Recentemente, inoltre, la ricerca nell'ambito della stabilizzazione del G-quadruplex si è affacciata all'ambito antivirale. Ad esempio BRACO-19, uno stabilizzatore di G-quadruplex a struttura acridinica, ha mostrato effetti anti-HIV-1 ed in generale la capacità di interagire con G-quadruplex costituiti da acidi nucleici a base di DNA, RNA o ibridi.

Lo scopo del progetto di ricerca è costituito dalla sintesi di piccole molecole che agiscono come potenziali stabilizzatori di acidi nucleici strutturati in G-quadruplex. Le molecole che sono state sintetizzate nel corso di questo progetto condividono, in generale, alcuni motivi strutturali comuni a composti riportati in letteratura quali antrachinoni, antraceni, naftalenediimmidi, acridine; gli schemi di sintesi sono inoltre stati progettati per riprendere le caratteristiche chimico-strutturali di un composto precedentemente sintetizzato dal

gruppo di ricerca del Prof. Zagotto che ha mostrato una notevole capacità nello stabilizzare il DNA G-quadruplex.

Tecniche avanzate e complementari quali modellistica molecolare, *fluorescence melting*, studi di legame ESI-MS e *ion mobility* MS sono state impiegate per lo studio della capacità delle molecole sintetizzate di interagire con il DNA e di stabilizzare tale particolare struttura. Gli esperimenti hanno messo in luce il ruolo di alcuni aspetti strutturali dei composti sintetizzati, come ad esempio le proprietà conformazionali, nell'influenzare l'efficacia nella stabilizzazione del quadruplex. La ricerca ha permesso infine di ottenere interessanti informazioni di relazione struttura-attività e di inviduare composti con una promettente capacità di stabilizzare il G-quadruplex.

Riferimenti

Neidle S. Stephen Neidle on cancer therapy and G-quadruplex inhibitors. Interview by Joanna De Souza. Drug Discov Today. 2004, 9(18), 778-81

Rezler EM, Bears DJ, Hurley LH. DNA tetraplex-binding drugs. Annu. Rev. Pharmacol. 2003, 43, 359-379

Read M, Harrison RJ, Romagnoli B et al . Structure-based design of selective and potent G-quadruplex-mediated telomerase inhibitors. Proc Natl Acad Sci USA 2001, 98, 4844-9

Perrone R, Butovskaya E, Daelemans D, Palu` G, Pannecouque C and Richter SN. Anti-HIV-1 activity of the G-quadruplex ligand BRACO-19. J Antimicrob Chemother. 2014 doi:10.1093/jac/dku280

ABSTRACT [ENGLISH]

The so-called quadruplex forming DNA sequences, with their peculiar feature of being particularly rich in guanines, can be found in many portions of both human and non-human genome. Considering the human genetic information, while one of the traditional localization of these sequences, the telomere, still represents an appealing target in anticancer therapy through indirect telomerase inhibition, many novel involvements are emerging in other portions of the genome. It has to be considered, first of all, that more than 376000 guanine rich sequences were identified in the human genome, with a preferential localization in some regions represented by proto-oncogenes. In this context G-quadruplexes could act as switches turning on and off, or regulating, the transcription of some sequences, according to the fact that structured DNA usually is not processed by the involved enzymes. G-quadruplexes have also been described over the years to be involved in many other key cellular processes such as chromosomal alignment, replication, transcription, genome recombination. This year another important piece of information was added to the quest for G-quadruplex stabilizers as antiviral agents. BRACO-19, an already described acridine-based stabilizing agent, was reported to show anti HIV-1 effects. These attractive targets boost the interest for the discovery of novel G-quadruplex stabilizer and for the investigation of their binding properties with different nucleic acids (DNA, RNA or hybrids), expanding their possible application from the anticancer to the antiviral field.

The research project is aimed to the synthesis of small molecules acting as potential stabilizers of this peculiar super molecular arrangement reported to be relatively easily formed by guanine-rich sequences, such as the ones in telomeres. The molecules that were synthesized during this project share, in general, the common structural motifs of previously reported G-quadruplex stabilizing agents (athraquinones, anthracenes, naphtalenediimides, acridines) and are inspired to a compound previously synthesized by the research group of Prof. Zagotto that showed a remarkable activity in stabilizing G-quadruplex DNA.

Enhanced techniques such as molecular modeling, fluorescence melting, ESI-MS binding studies and ion mobility MS were then employed for the investigation of the capability of the synthesized molecules of interacting with, and stabilizing, G-quadruplex DNA. These complementary techniques enlightened the relevance of some structural aspects of the synthesized compounds, such as conformational properties, in influencing the efficacy of the DNA stabilization. The results allow to describe preliminary structure-activity relationship data and some promising compounds were finally disclosed.

References

Neidle S. Stephen Neidle on cancer therapy and G-quadruplex inhibitors. Interview by Joanna De Souza. Drug Discov Today. 2004, 9(18), 778-81

Rezler EM, Bears DJ, Hurley LH. DNA tetraplex-binding drugs. Annu. Rev. Pharmacol. 2003, 43, 359-379

Read M, Harrison RJ, Romagnoli B et al . Structure-based design of selective and potent G-quadruplex-mediated telomerase inhibitors. Proc Natl Acad Sci USA 2001, 98, 4844-9

Perrone R, Butovskaya E, Daelemans D, Palu` G, Pannecouque C and Richter SN. Anti-HIV-1 activity of the G-quadruplex ligand BRACO-19. J Antimicrob Chemother. 2014 doi:10.1093/jac/dku280

1. INTRODUCTION

1.1 AIM OF THE PROJECT

This research project is aimed to the design, synthesis, characterization and evaluation of novel G-quadruplex stabilizing agents. Among other non-conventional arrangements of DNA, G-quadruplexes are reported to be involved in many key cellular processes. One of the classical and most described location of these structures is in telomeres: vertebrates telomeres contain the repeated sequence TTAGGG, with all the bases except the terminal 3' end 15-200 nucleotides are in the duplex form.[1] The guanine-rich single stranded overhang is very likely to fold into a G-quadruplex structure generally associated with proteins.[2] This conformation prevents the telomere elongation process by telomerase, due to the fact that structured DNA is not processed by the enzyme. As a result, stabilization of G-quadruplex structures represents an interesting strategy in preventing cell immortalization.

Telomeres and telomerase are interesting targets in the search of potential anticancer agents, first of all because of the peculiar G-quadruplex structure that could represent the key for a selective therapy target. Some existing compounds are known to be G-quadruplex stabilizers, through the mechanism of intercalation or end stacking: porfirines, perilenes, anthraquinones, fluorenones, acridines, dibenzophenanthrolines, telomestatin. In addition to this, as long as it has been reported that guanine-rich sequences can be found both in telomeres and in promoter regions of DNA (proto-oncogenes) but also in viral DNA/RNA genome (SARS, Epstein Barr), stabilization of G-quadruplexes could be involved in the viral infection evolution,[3] encouraging the growing interest in stabilizers as antiviral agents.

Our research group previously synthesized a promising compound with excellent G-quadruplex stabilizing activity.[4] In this PhD project the attention was focused primarily synthesis of novel compounds to better set up a structure-activity relation study. Structure based and ligand based drug design approaches were also introduced in order to support the

design of the compounds and to figure out their interaction motif with the quadruplex. During the research activity different scaffolds such as anthracene, anthraquinone, bisanthrapyrazole and acridine were used for further modifications. On the other hand, novel structurally constrained derivatives were designed to investigate the role of the flexibility of the side chains in the interaction with the guanines or the loops, even overcoming their chemical properties. In this part of the project, NMR experiments were carried out in order to support the hypothesis concerning the flexibility constrains.

1.2 QUADRUPLEXES: NOT ONLY TELOMERASE

An overview of quadruplexes in human genome

G-quadruplex is one of the most interesting super molecular, non canonical arrangement of DNA. The so called quadruplex forming sequences, with their peculiar feature of being rich in guanines, can be found in many portions of the human and non human genome. In fact, while one of the traditional localization of these sequences, the telomere, still represents the key for an appealing target in anticancer therapy through indirect telomerase inhibition, many novel involvements are emerging in other portions of the genome.

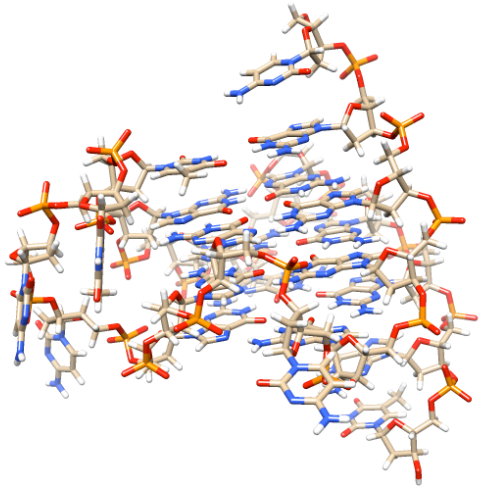


FIG1. G-quadruplex structure formed by human VEGF promoter.[5]

It has to be considered, first of all, that more than 376000 G-quadruplexes are estimated to be present in the whole human genome, with a preferential localization in some regions of proto-oncogenes.[1] To be more specific, then, it has been reported that in more than 40% of human protein-coding genes at least one G-quartet motif can be described in their promoter regions, and this represents an additional suggestion about the role of this arrangement as a trigger in transcriptional processes.[6] G-quadruplexes have also been described over the years to be involved in many other key cellular processes in addition to the interference

with telomerase such as chromosomal alignment, replication, transcription, and genome recombination.[7]

Introduction to G-quadruplexes in other organisms

In addition to this, G-quadruplex forming regions are also reported to be present in several other organisms.[8]

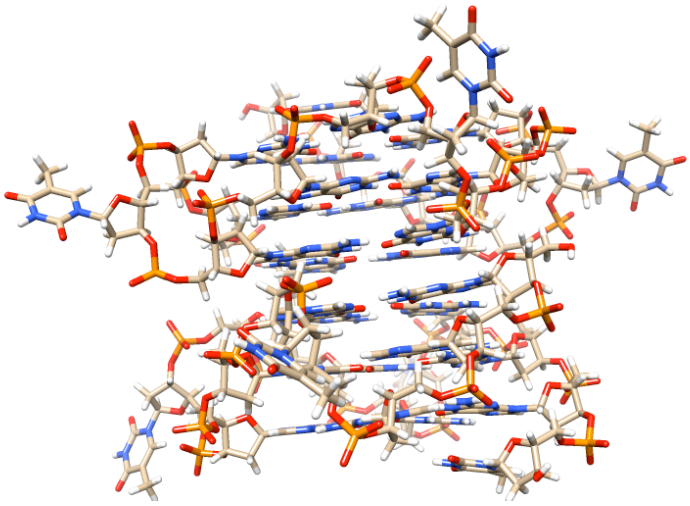


FIG2. G-quadruplex sequences as the one reported inhibits HIV replication.[9]

Expanding then the area of interest from the antiproliferative field to potential antiviral targets, G-rich sequences were shown to inhibit the replication in culture of human immunodeficiency virus (HIV-1) and, moreover, as will be later described, NCp, a multifunctional protein from HIV-1, has been show to interact with quadruplexes.[10]

1.3 THE ROLE OF TELOMERES

A brief introduction

At the beginning of XXI century more than 100 different types of cancers and, partially, their physiological pathways were already been described. Despite, or in addition to, this even today, cancer remains a leading cause of death in the world and the quest for new therapeutical strategies remains open. Back in the 90s Shay and Harley discovered in 90 out of 101 tumor cell lines the presence and the activity of an enzyme, telomerase, instead normally inactive in somatic cells.[11] Later on additional studies have shown a clear correlation between telomerase reactivation and nearly 90% of tumor cells.[11] Telomerase, as its usual and natural role, is involved in maintaining the length of the most external portion of the human chromosome. In fact, the outer ends of the single stranded molecules of DNA have a high degree of instability due to various mechanisms including, for example, degradation by exonucleases, control systems that identify breaks in DNA and that send signal leading to the apoptosis of the cell and, the most intriguing, the loss of bases at each cell cycle, due to the mechanism of replication.[12] To overcome these problems, the outer portions of our genome present peculiar sequences, called telomeres, which consist of hundreds of copies of a sequence made by repeated non-coding bases. The role of telomeres is to act as a protection of the genetic information, preventing its degradation by enzymes, but especially allowing its maintenance through the cell cycles.[12] Telomere shortening due to the mechanism of replication leads anyway to a natural senescence of the cell after a number of cycles, setting a limit in the cell lifetime. In some cells, however, the telomere happens to be elongated by telomerase, overcoming this time limit. This is true in stem cells, while telomerase is inactive in most somatic cells.[12] The pathologic scenario is when the presence of errors in the regulation of the enzyme telomerase occurs, leading to its reactivation that causes a deregulation in the balance

between telomere shortening and lengthening. This turns out in the inhibition of the signal of cellular aging.[11] The observation that almost 90% of the tumor cells was found to maintain the length of telomere by this enzyme is then consistent with the resulting immortalization of the cell.

Replication and "end replication problem"

DNA replication is the mechanism that allows the reproduction of the entire chromosome(s) during a cell cycle, that is the time between the birth of the cell and the moment in which the cell itself splits into two. This part of its life is named mitosis or M phase. The intermediate time in which the cell grows and prepares for mitosis, the interphase, cell is divided into the G1, S and G2 phases.[13] The synthesis of the DNA sequences occurs in the S phase and is operated by an enzyme called DNA polymerase. To be more specific many kinds of DNA polymerase are involved. Among these, DNA Pol δ, the DNA Pol ε and DNA Pol α are the enzymes reported to be involved in the addition of nucleotides to the terminal 3'-OH of the 5'-3' growing sequence.[14]

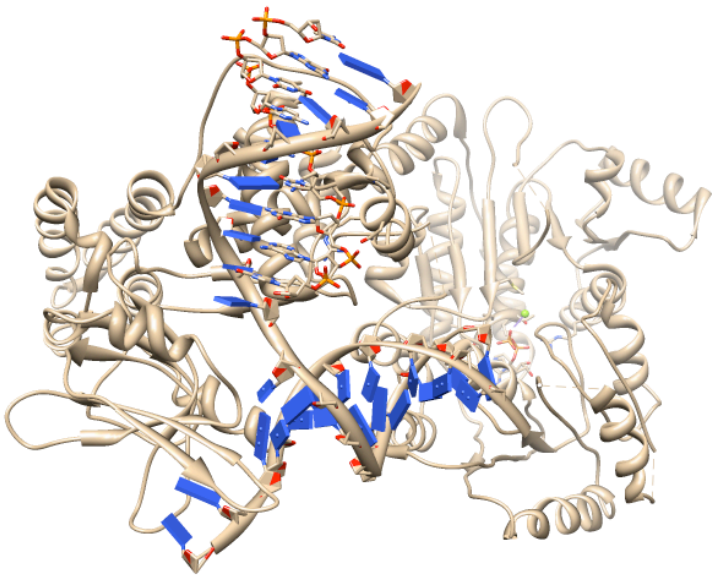


FIG3. The fork[15]

Both DNA strands of the double helix are replicated simultaneously. To allow this, at the beginning of the replication process the double chain is forced into a structure called the replication fork, so that the two single strands become both available for the action of DNA polymerase.[14]

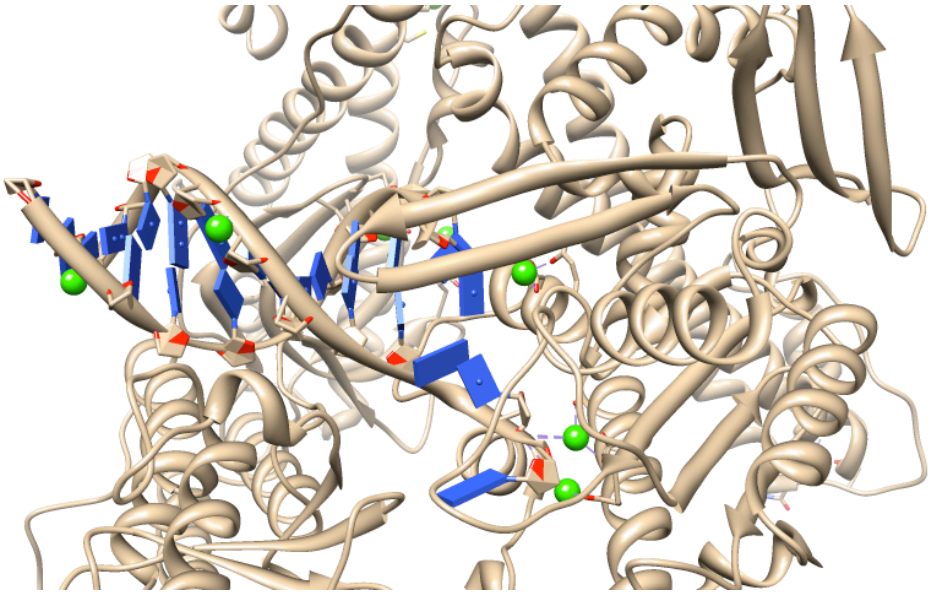


FIG4. Dna polymerase I bound to duplex DNA.[16]

Anyway, as long as DNA polymerase can smoothly proceed with the synthesis of the nucleic acid only in the 5'-3' direction, the replication takes place with two different mechanisms on the two strands. In fact, one strand is synthesize continuously while the original sequence is processed, while the other one is copied in a discontinuous manner, through the so called Okazaki fragments.[14] These shorter sequences are supposed then to be rejoined by another pool of enzymes, the ligase.[14] As previously introduced, the DNA polymerases involved in the elongation are DNA Pol δ and DNA pol ε, which need a free 3'-OH free to start the process of polymerization. The action is directed by DNA Pol α that binds to the pre-replicative complex. This complex, also known as pre-RC, is formed in the G1 phase of the cell cycle.[14] The DNA Pol α (primase) synthesizes an RNA primer, a sequence of 5-10 nucleotides which ends with the free 3'-OH.[14] This event gives rise to all the following synthetic process. The leading strand (the one that is smoothly replicated) needs only one primer sequence and after its recognition replication proceeds without gaps

up to 5' end. On the other hand primers represent a crucial issue for the lagging strand: the need for the primer creates the "end replication problem" encountered at the terminal end of the chain. To be more specific, the discontinuous chain instead requires multiple primers, resulting at the end in impossibility for the polymerases machinery to completely copy the 3' end. All of this results in the shortening of the lagging strand at each cell cycle, which could potentially cause loss of genetic information throughout the cycles.[14] Given these considerations, it is clearly understandable how the presence of the telomeric sequence at the 3' end of the chromosomic DNA is relevant.

Telomeres and the "Hayflick limit"

The peculiar structure of telomeres consists in the repetition of a sequence of DNA bases: 5'-TTAGGG-3'. As long as it is a non-coding sequence, the loss of a portion of the telomere during each replication cycle does not lead to loss of genetic information.[17] This is anyway true until a certain limit is reached. Somatic cells replicate in vitro a limited number of times, and this event can directly be connected to what stated above: even if the loss of bases during the various cell cycles doesn't affect coding sequences, after that a certain number of bases are lost and the "Hayflick limit" is reached the cell is not in the condition to replicate anymore.[18] In other words, the exceeding of this limit, that would be a further shortening of the telomere, even if theoretically possible as long as are still present hundreds of nucleotides, induces an arrest of growth called senescence (M1 phase).[17] In some cases some cells could exceed this stage of senescence, due to inactivation of the genes that normally codify for critical control points of the cell cycle. This would turn out in a further shortening of the telomere, leading the cell to a second critical stage (M2), usually quickly followed by its death.[17] The intriguing point is that a cell somehow overcoming phase M2 acquires the ability to replicate indefinitely, becoming virtually immortal. This event can be connected, in the great majority of the cases, to the reactivation of the enzyme telomerase, as previously introduced.[17]

1.4 TELOMERASE: THE ENZYME

The structure: hTR, hTERT and other proteins involved

Telomerase is a ribonucleic/proteic complex and has the role of elongating the 3' portion of the terminal chromosome due to the presence in its interior of a portion of RNA, which acts as a primer itself.

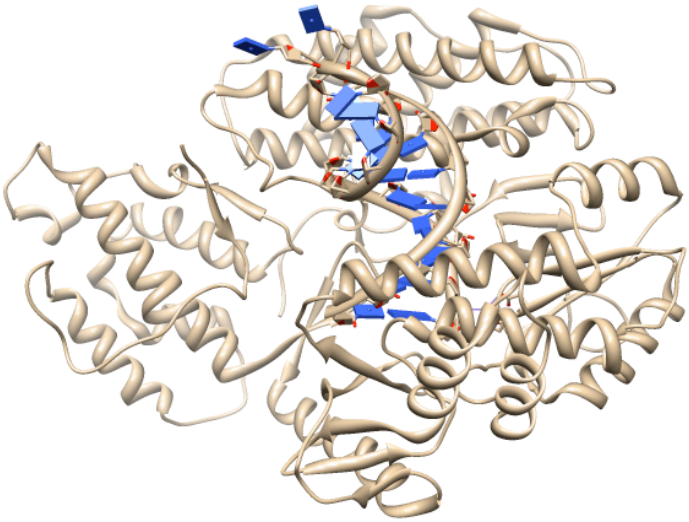


FIG5. Crystal structure of telomerase from *Tribolium castaneum*[19, 20]

The enzyme is traditionally described as composed of two subunits: the region that carries the RNA named hTR (human telomerase RNA) and the catalytic subunit hTERT (human telomerase reverse transcriptase).[17] Some proteins help then the formation of the complex and support the overall stability.

Concerning hTR, the RNA strand, that allows the recognition and the synthesis of the telomeric DNA, is a strand of 451 nucleotides transcribed by RNA polymerase II. [21] Near the 5' end this portion has a peculiar and well conserved among many different species base repeat: the sequence of bases 3'-AUCCCAAUC-5' allows the interaction with the telomere.[17] Inside this subunit the remaining RNA is in the double stranded form. Among its roles there are the regulation of the addition of nucleotides during the transcription of

the single stranded sequence and it provides a linkage with the catalytic subunit hTERT.[22] Also peculiar is a pseudoknot domain, known to be relevant in maintaining the structural stability of the nucleic acid and of the whole enzyme.[22]

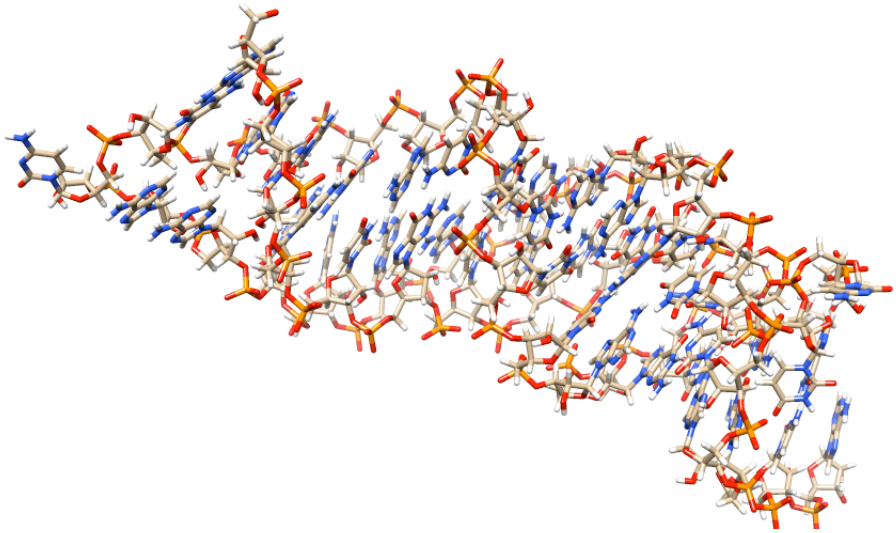


FIG6. NMR structure of human telomerase RNA pseudoknot.[23]

This kind of domain, from a structural point of view, can be seen as a loop region that folds and further interacts with a contiguous chain. Another important constituent is represented by the transactivation domain, made by a loop hairpin and other internal loops. It contributes to the correct folding of the pseudoknot domain and, in general, to the activity of the enzyme.[22] Many different double stranded or peculiarly arranged domains have been deeply described and investigated, all of them giving contributes to the stability and overall activity of the subunit and of the enzyme.[22] In addition to this, the structural stability is ensured also by a good number of associated proteins fundamental to the stability, the maturation and localization of the hTR subunit in the complex telomerase and to the interaction with the telomere.[17]

The catalytic subunit, hTERT, has a highly conserved structure and carries similarities with the functional domains of the reverse transcriptase and DNA polymerase. The main domains are the TRBD (RNA binding domain), the RT (reverse transcriptase) domain, the N-terminal domain and the C-terminal domain.[24] The reverse transcriptase domain, in

particular, is organized in the two subdomains named “palm” and “fingers”, connected by a loop.

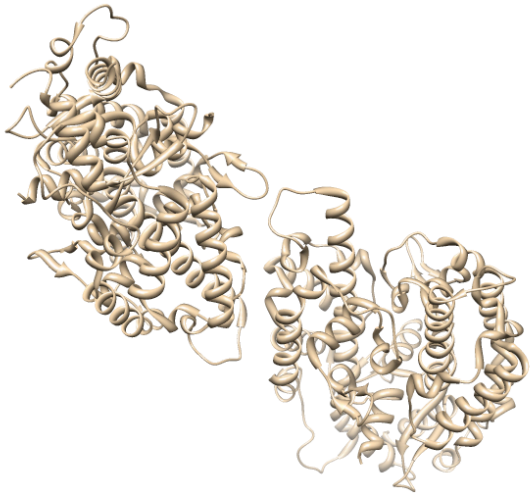


FIG7. RNA binding domain from vertebrate.[25]

This sequence seems to be involved in the exact placement of the 3’ chromosomic end within the active site of the enzyme.[22] The domain IFD (insertion finger domain) seems to be involved in protein-protein intramolecular interactions that ensure the stability of the RT domain.[26] The size of the catalytic cavity is enough to allow the insertion of seven or eight pairs of bases. The catalytic activity that takes place in the active site is then carried out thanks to the presence of three residues of aspartic acid and a lysine, which promotes the activation (acid catalysis) of the moiety that acts as a leaving group during the synthesis, the pyrophosphate. The N-terminal domain TEN is not detectable in every species, it can not in fact be found in the crystallographic structure of the catalytic subunit of telomerase from *T. castaneum*, a recently reported well considered form (FIG5).[27] When present, it is described to help the enzyme to add multiple sequence repetitions to a single primer. It also contains a portion with weak affinity for the RNA-based portion thanks to the RID1 (RNA interaction domain 1).[28] The C-terminal domain, then, is involved in the process of formation and stabilization of the DNA/RNA heteroduplex in the active site of the enzyme.

As previously introduced, one of the most peculiar features of human telomerase is the presence of the RNA subunit, a distinctive region because of which this enzyme differs

from most of the reverse transcriptase. This is made possible also by the many proteins that support and help the overall structural stability. Proteins binding hTERT and remaining stably associated in a complex are hsp90 and p23 and in addition to their structural functions probably have an influence in promoting an adaptation of the catalytic subunit to allow the interaction with the strand.[17] Furthermore, the C-terminal residue, happens to be associated to other proteins through residues of serine and threonine. These interactions are reported to be extremely important for the regulation of the nuclear localization of the enzyme.[17]

Mechanism of action and regulation

While synthesizing DNA, telomerase proceeds following two mechanisms. Processivity type I allows the displacement of telomerase along the DNA template after the addition of each nucleotide; processivity type II allows the translocation of the telomerase complex after the addition of an entire sequence.[29]

The replication itself takes place in three basic steps. The first step consists in bonding of the RNA sequence to the 3' end of the telomere. The heteroduplex formed in this way is generally more than 7-8 base pairs long (FIG8). This average lenght is maintained because the base pair bonds between the distal couplets of bases are broken as soon as the synthesis proceeds.[26] The weak interactions between RNA and DNA bases in the heteroduplex are not enough to obtain the required stability of the nucleic acid-protein-nucleic acid complex, and additional weak interactions are present.[26] To be more specific, three different points of interaction between the enzymatic complex and the telomeric DNA were enlightened. As previously said the 3' end of the DNA forms a hybrid heteroduplex with the enzymatic RNA. Two parts of the DNA sequence closet o this first interact then with a proximal site (TEN, N-terminal domain) and a distal site.[26]

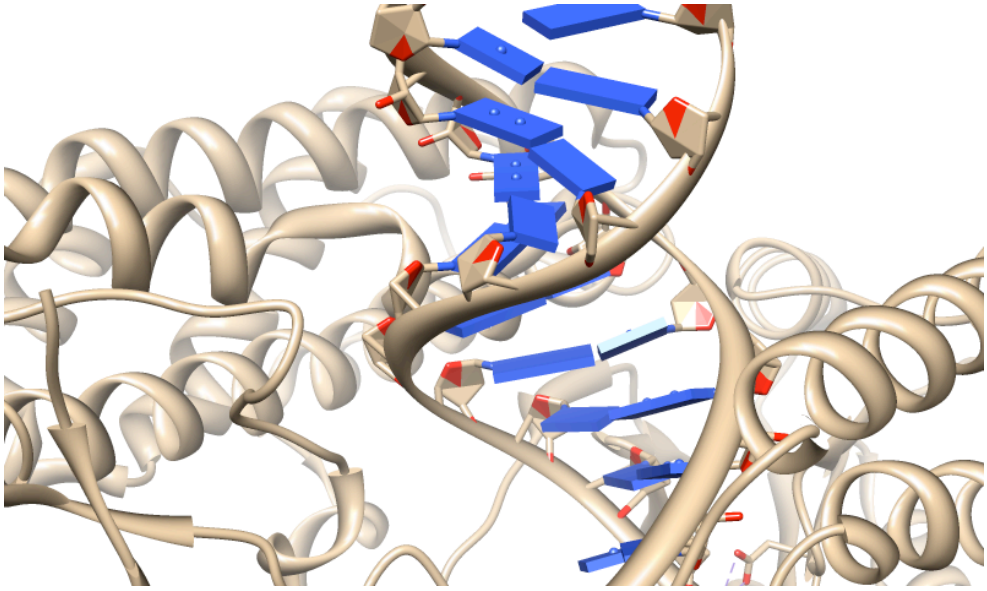


FIG8. Detail of RNA/DNA heteroduplex in telomerase from *Tribolium castaneum*[20]

The second step consists in the addition of nucleotides to the growing chain. The enzyme proceeds as a DNA polymerase. The process is catalyzed by two Mg^{2+} ions that are involved in the binding to the substrate. Those ions are coordinated by the three residues of aspartic acid. A lysine provides the activation of the acid catalysis process to promote the cleavage of the leaving group pyrophosphate.[22] At one point, the limit of the RNA template is reached. Telomerase can then dissociate (if the process is completed) or move toward the end of the newly synthesized DNA chain and begin to synthesize another DNA sequence to continue the elongation of the telomere. This feature of telomerase is called processivity.[22] The process that allows the translocation of the enzyme has not been completely enlightened yet, but it seems to be allowed by a conformational change of the protein complex involved in the telomere anchoring sites.[30]

As previously introduced, the activity of telomerase is repressed in most somatic cells during the embryonic differentiation. On the other hand, it can be still found in its active form in some tissues such as the haematopoietic stem cells and lymphocytes.[31] Basing on the considerations reported above, the expression and the activity of this enzyme are described to be fundamental for the immortalization of a cell pursued through an unlimited number of replication. The regulation of telomerase activity can occur at several stages of

the process in which the enzyme itself is involved: transcription, mRNA splicing, mutations and modifications of hTR and hTERT structures, localization inside the cell and assembly of the ribonucleoprotein complex.[17] The key step anyway is reported to be the expression (and its regulation) and the correct assembly of the hTERT catalytic component. The gene encoding for hTERT is present as a single copy in the chromosome band 5p15.33.[17] It has been found that, as can be expected, a functional telomerase requires a complete transcription of the gene. In addition to this gene expression regulation factors are very important in the tuning process of transcription; hTERT gene has binding sites for these factors, classified as activators or repressors, in proximal or distal regions.[17]

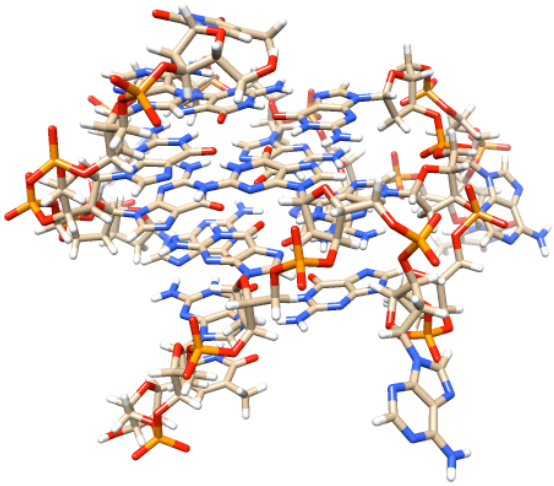


FIG9 NMR structure of a c-myc G-quadruplex.[32]

Among the activation factors (or genes encoding for factors) there are:

- *c-myc*, an oncogene involved in the mechanisms of cell proliferation, growth and apoptosis. This gene encodes a protein, Myc, which form heterodimers able to recognize and bind to a site in the promoter.[17,33] In addition to this an over expression of c-myc gene boosts the activity of the hTERT gene promoter.[31]

- *Sp1*, a transcription factor that binds to the GC-box and is consecutively involved in the expression and amplification of some genes relevant for cell survival in general. The hTERT gene carries five GC-box located between two E-boxes and they seem to be involved in the gene promotion.[17,33]

- VEGF promoter, as reported at the beginning of this chapter.[5]

Reported repression factors are:

- *Mad1*, the gene coding for the Mad protein. This protein can form Mad/Max heterodimers like also Myc does. In this case, otherwise, an interaction with E-box promotes a repressive effect.[17,33]
- *p53*, a protein involved in repression of the growth of tumor cells. This can happen through an interruption of the cell cycle or by an induced apoptotic event. Recent studies propose connections between this protein and the transcriptional repression of hTERT. [17]

A lack of expression or a down regulation of these genes turns out, and it has been shown in many different tumor cell lines where a reactivated telomerase that guarantees the maintenance of telomere length can be enlighted.[17,33] Some examples reported in literature comprehend cells belonging to tumor tissues from lung, stomach, colon and rectum. These cells, supporting this, showed minor or no expression of mRNA coming from the *Mad1* gene, justifying its role in the regulatory mechanisms of tumor.[34]

In addition to the above reported regulation processes, the activity of telomerase can be influenced by other adjustment mechanisms. In particular the activity of telomerase can be also tuned by the telomere itself, depending on its structural properties and the proteins bounded to the DNA sequence.

1.5 THE TELOMERES

A quite comprehensive definition of telomere is, from a structural point of view, "ribonucleoprotein complex". As previously introduced it is located in the outer part of the chromosome and the nucleic acid forming the sequence is mainly arranged in a double strand. The most peculiar feature is that the chain comprehends 5'- TTAGGG-3' repeated sequences. In humans this motif can be repeated in a range between 9000 and 15000 bases. The structure is then completed by a set of proteins and a single stranded sequence at the 3' end with a size of 50-300 bases. This is usually named G-overhang.[35] The telomere is shortened at each cycle of a number of bases between 50 and 150.[36] It has been proved that several thousand bases are required to allow an entire life cycle facing the progressive shortening.[36] Telomeres are also reported to protect against the activation of cellular response to DNA damage (DDR, DNA damage response): thanks to their particular structure they can arrange to form a loop called *t-loop*, which acts as a terminal capping of chromosomes. The single-stranded chain is situated in a position which allows its interaction with the double helix and form another loop, called the *d-loop*.[35]

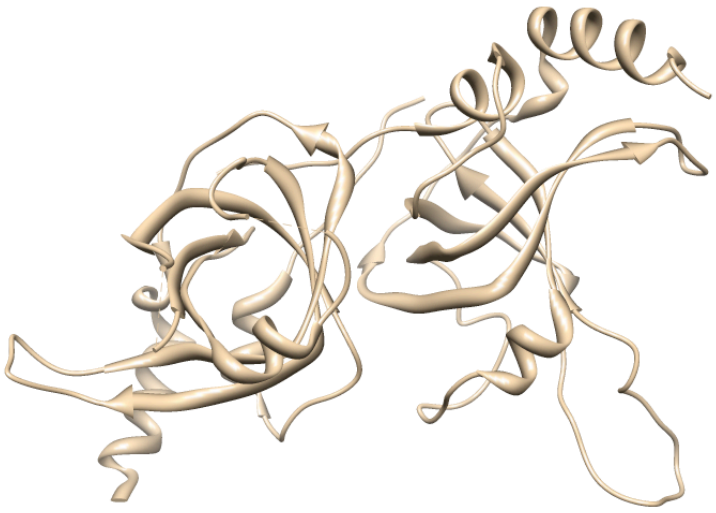


FIG10. Proteins from the *shelterin* complex.[37]

The structure is stabilized thanks to the interaction with many different proteins, up to more

than 200 according to what literature reports.[35] The main interaction motif is anyway the one with the *shelterin* complex, a system formed by six interlinked proteins.[38] Protein-protein, protein-single stranded DNA and protein-duplex interactions are present. From a molecular point of view, the system consists of TRF1 and TRF2 (Telomeric Repeat Binding factor 1 and 2), TIN2 (TRF1-and TRF2-Interacting Nuclear Factor 2), POT1 (Protection Of Telomeres 1), TPP1 (Also known as the set composed by TINT1, PTOP and PIP1) and RAP1 (Repressor/Activator Protein 1).[38] TRF1 and TRF2 bind the telomere on the double-stranded portion, while POT1 binds the G-overhang (the single stranded portion). TIN2, Rap1 and TPP1 are not directly binding DNA, but give respectively interaction with TRF1, TRF2 and POT1.[39] Even if structurally similar, TRF1 and TRF2 act differently within the complex. TRF1 seems to be directly involved in the regulation of the enzymatic activity of telomerase. The more extended is the interaction between TRF1 and the telomere, the lower is the probability of binding of telomerase for a sterical hindrance due to the arrangement obtained.[36] TRF1 influences by a negative feedback the activity of telomerase, that it binds in greater proportions when the enzyme is processing longer telomeres.[40] TRF2 is on the other hand involved in telomere protection (it helps the *t-loop* formation) and stabilization of the G-quadruplex. Protein Rap1 is connected to TRF2 and plays a role down regulating telomeric length. TIN1 holds then together the protein complex, binding to TRF1 and TRF2.[40] The POT1/TPP1 heterodimer is otherwise directly involved in balancing the activity of telomerase: it competes with the enzyme for the binding to the single stranded telomere.[40] POT1 itself also possesses a binding domain for DNA which allows a strong interaction with the telomeric sequence with high affinity. Binding the single stranded chain it inhibits telomerase activity.[40]

The correct arrangement and the stability of the *shelterin* complex is essential for a balanced functionality. A failure in even a single subunit affects heavily both the activity of telomerase and the various repair mechanisms.[38] Concerning this, the DDR system involves two particular PI3K-related protein kinases, ATM and ATR. When these proteins are in touch with the damaged site a signaling pathway is switched on and all of this leads to the

activation of repair mechanisms that could affect the outer portion of the chromosome.[39] Two subunits of the shelterin turned out to be particularly of interest in this field: TRF2 interferes with the ATM pathway, probably through the stabilization of the *t-loop* structure, while POT1 is involved in the suppression of the ATR pathway.[39]

Given all these consideration, telomerase is then regulated by the spatial arrangement of telomeric DNA. A peculiar conformation of the nucleic acid could in fact act against the enzyme making the DNA sequence itself not processable.

1.6 G-QUADRUPLEX

Structural features of quartet and quadruplex

The DNA sequences rich in guanine tend to form special structures called quartets guanine, or G-quartet. These are composed of four guanines arranged in a square planar motif and each couple of bases is linked by two hydrogen bonds involving N1, N2 and N7 nitrogen atoms and O6 oxygen atom of guanine.[19]

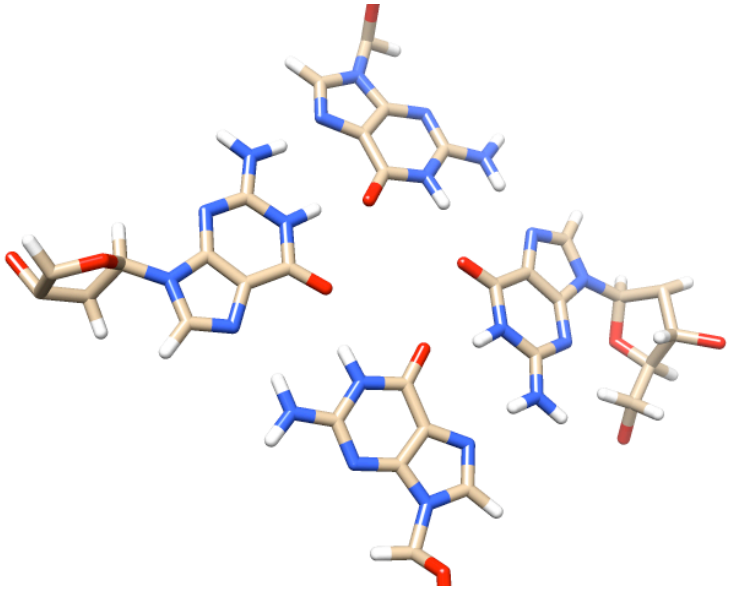


FIG11. Detailed view of G-quartet.[41]

Chains containing a sufficient number of guanines can arrange then in compact structures, given from the superimposition of different quartets named G-quadruplexes. The base motif is composed by at least two stacked quartets and the loops, sequences of nucleotides connecting the guanines but not involved in the structure.[19] The G-quadruplex is a structure of a remarkable stability thanks to the stacking of the bases (π-π.), and the hydrogen-bonding motif, together with electrostatic interactions, improve the overall arrangement. The heteroatoms directed toward the center of the quartet (and of the quadruplex) require a positively charged atom as a compensation for the concentration of negative charge given

by the electrons.[40] The distance between the N1 nitrogen atom and the O6 oxygen atom, between 2.85 and 2.95 Å,[42] gives a clear idea about the size of the central cavity and so, about the ion that could be involved in this arrangement.

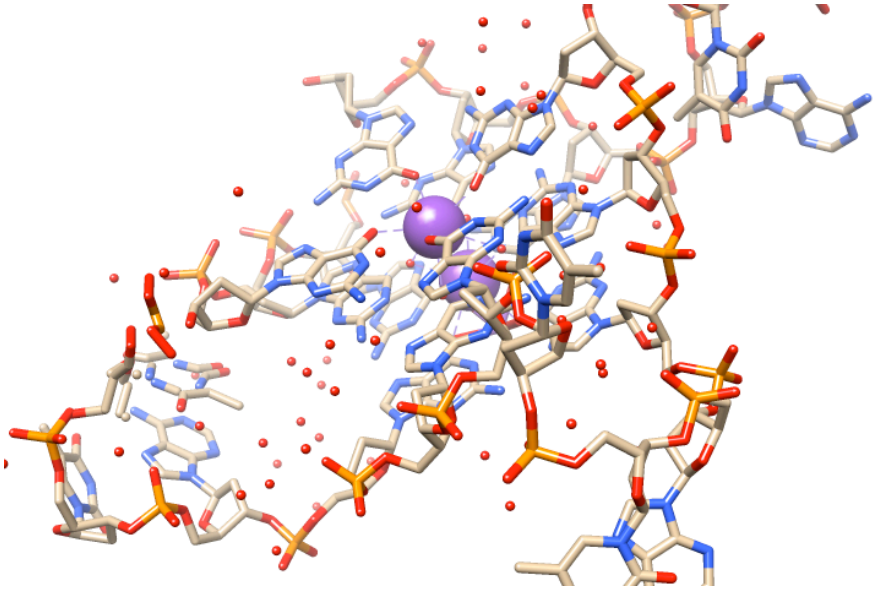


FIG12. Detailed view of G-quadruplex.[43]

Atomic radius is a key factor as long as the whole structure is arranged all around this "tunnel" of ions. Na^+ is sufficiently small to be coordinated by the guanines and be maintained in the same surface in a co-planar motif. Given this, it is not unusual the collocation of this ion also between the quartets along the G-quadruplex.[19,44] K^+ is otherwise reported to be localized between two quartets of guanine and creates a bipyramidal tetragonal configuration being coordinated by the carbonylic oxygens of the guanines.[44] Other potential stabilizers of G-quadruplex are Rb^+, Cs^+ or divalent cations such as Ca^{2+} (in this case the structure is similar to the one reported for K^+).

The quadruplex arrangement is supported by the four strands of nucleotides connecting the guanines and they can belong to one DNA chains or be the result of the winding of a single chain. According to this, G-quadruplexes can be intramolecular, bimolecular or tetramolecular. An intramolecular G-quadruplex is an arrangement of a single chain of DNA. Concerning the sequence of bases, it can be described as $X_n\ G_m\ G_m\ G_m\ X_o\ G_m\ X_p$,

where m is the number of guanines involved in the formation of the G-quartet, while X_n, X_o and X_p represent any combination of nucleotides forming an anse.[44] Intermolecular G-quadruplexes are otherwise made by two (bimolecular) or four (tetramolecular) different DNA chains. In the first case, i. e. the bimolecular arrangement, the sequence of bases follows the motif $X_n\ G_m\ X_o\ G_m\ X_p$, while the tetramolecular structure it can be represented as $X_n\ X_n\ G_m\ X_o\ G_m$.[44] The four strands may also have the same or different orientation: if all the sequences are directed towards the same direction they (and the structure) are called parallel, while antiparallel arrangements are more various. Examples are:

- three parallel strands and one with directed towards the opposite direction
- two parallel closet o each other and two opposite strands closet o each other
- two parallel and two opposite strands alternatively disposed[45]

The arrangement of these sequences is connected with the conformation of the bond between guanine and deoxyribose. This dihedral angle may be in the syn conformation, if the angle is between 0° and 90°, or in the anti conformation if it is between -120° and 180°. When all four strands are parallel all the bases are in the anti conformation while the presence of one or more antiparallel strands forces bases to the syn conformation.[45] The width of the grooves resulting from these different conformations can vary from narrow to wide.[45]

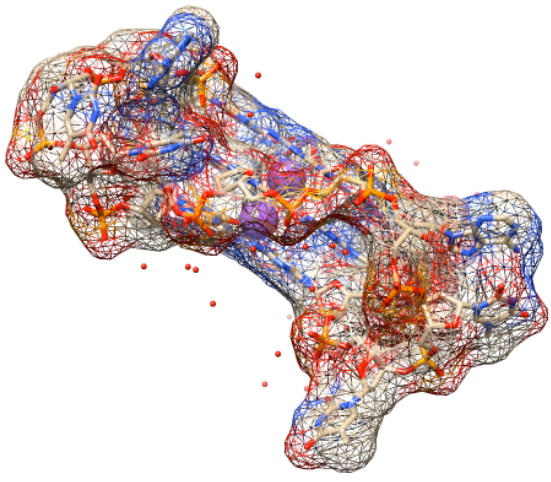


FIG13. G-quadruplex: its volume and surface.[57]

Another factor that can influence this parameter is the nature of loops in terms of sequence

of bases.[44] The loops can be divided into four groups: side loops (connecting two adjacent antiparallel chains), diagonal loops (connecting two alternate antiparallel chains), propeller loops (connecting two adjacent parallel chains), V-shaped loops (connecting the corners of two G-quartets where one of the two is lacking of a connection to a guanine).[46] The number of quartets stacked, the length of the strands and their sequence and in some cases also the ion involved can drive the structure to one or another loop arrangement.

Basing on all these considerations, G-quadruplexes cannot be described as a well-defined and rigid structure, given all the variables described above.

The telomeric G-quadruplex

The 3' end of the human telomere, shows an overhang made by a single stranded guanine-rich DNA.

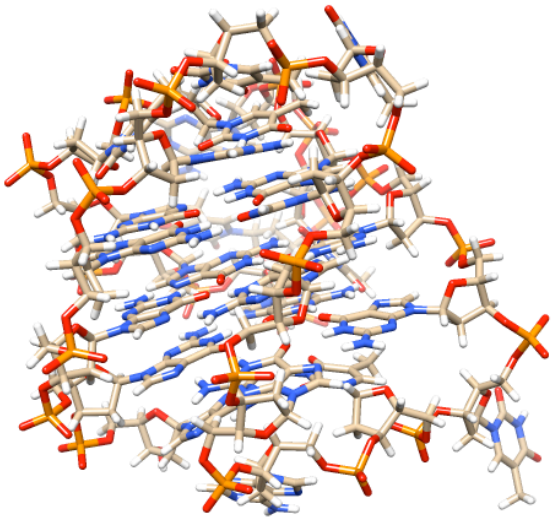


FIG14. A telomeric G-quadruplex, NMR solved structure.[41]

Although a natural tendency to form G-quadruplexes can be expected, as already described above the structure forms the *shelterin* complex together with a number of proteins (hPOT1 in particular is strongly bound to this portion).[19] According to this, G-quadruplex then is very prone to form *in vivo* if induced by the binding to some molecules able to stabilize the

arranged structure, even in comparison to the complex with hPOT1.[19] The investigation of the exact topology of the telomeric G-quadruplex in the cellular environment is representing a challenge since years. In 1993, Wang and Patel used the sequence d[AGGG-(TTAGGG)$_3$] to investigate the structure formed in a NaCl solution through NMR analysis and the intramolecular basket type structure was found: two parallel strands with the guanines in syn-syn-anti-anti conformations.[41]

The study of the G-quadruplex in a Na^+ solution, however, has the limitation of not completely reproducing the cellular environment, where the concentration of Na^+ is 10 mM (the measure was carried out with concentrations over ten folds higher) and the concentration of K^+ is 135 mM.[41] More recent studies were then aimed at reproducing physiological conditions. In 2002, Parkinson's, Lee and Neidle reported a different topology for the same sequence d[AGGG-(TTAGGG)$_3$] obtaining the crystallographic structure of the arrangement in a K^+ solution.[47] This intramolecular quadruplex presents all parallel strains and the anti guanine conformation. Other NMR studies on telomeric sequences in K^+ solutions enlightened the presence of an equilibrium between various forms of G-quadruplex supporting, as previously introduced, that a single representation of the structure is not enough to describe this peculiar and flexible arrangement.[19] The probably most representative structure in these studies is called antiparallel (3+1) quadruplex, in which three chains are oriented in the same direction, while the fourth in the opposite.[41]

Some studies enlightened intermolecular structures with three chains belonging to a strand of DNA and a chain belonging to a different molecule[19] while more recent studies have instead described intramolecular structures formed by human telomeric sequence d[TAGGG(TTAGGG)3] arranged in the (3+1) core.[41]

Coming to the end, the single-stranded chain has proved to be an excellent substrate for the formation of a quadruplexes, but the structure adopted in the actual arrangements of the is still an open field for further investigation.[19]

1.7 OUTSIDE THE TELOMERE: OTHER G-QUADRUPLEXES

Besides telomerase inhibition

The presence of G-quadruplex has been reported or hypothesized to be formed in several specific regions of the genome in addition to telomeres. Recent bioinformatics studies have identified that in human genome, basing on considerations about the sequence of bases, approximately 376,000 sequences could realistically arrange in G-quadruplexes.[48] As previously introduced, these sequences show a peculiar and not random localization: the so-called "putative G-quadruplex-forming regions" are prevalent in proto oncogenes but lack in tumor-suppressor genes and are also co-regulated by the interaction with proteins.[19] In particular, they have been found in 40% of promoter genes and other regulatory regions, and this suggests a possible role in the regulation of the transcription.[48]

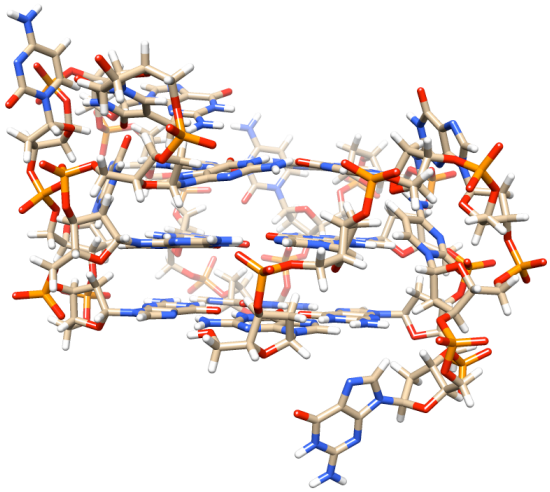


FIG15. G-quadruplex from *bcl-2* promoter.[49]

The genes *c-myc* (see above for a detailed description of its role and of the role of these and other promoters in the regulation of telomerase) and *c-kit* are two outstanding examples. To be more specific, *c-kit* can be connected to cell proliferation, differentiation and apoptosis. A role in the regulation of telomerase itself seems also to be proved.[19] Some other reported

examples of oncogenic promoter regions realistically regulated by G-quadruplex formation are *bcl-2* (mediates chromosomal translocation connected with the onset of lymphomas), *vegf* (stimulates the formation of blood vessels), *hif-1α* (activated in many tumors and associated with local invasion and metastasis) sequences.[50] In some cases the suggestion of the presence of G-quadruplex structures in these sequences has also been confirmed by experimental analysis.[19] It is also important to notice that promoter regions are part of DNA duplexes and they, probably, unwound during transcription to be able to fold in G-quadruplexes.[50] This leads to a novel attractive target in anticancer therapy: a small molecule stabilizing these G-quadruplexes could modify the expression of the genes. Unusual DNA conformations in human genome seem also to be involved in the triple repeat disease. In this disease, manifesting through neuromuscular and neurodegenerative disorders, a dynamic intergenerational expansion of triple repeat d(CGG)n-d(CCG)n, d(CAG)n-d(CTG)n e d(GAA)n-d(CTT)n is reported.[50]

G-quadruplex in HIV, other viruses and RNA sequences

G-quadruplex represents also a promising target in antiviral therapy, as long as many G-rich sequences have been observed in viral genomes.[51] To be more specific, as an example, human parvovirus adeno-associated virus (AAV) and HIV are under study.

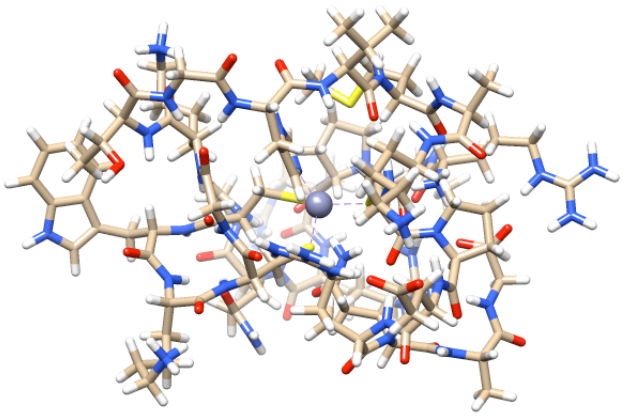


FIG16. HIV-2 NCp protein complexing Zn.[52]

In particular DNA structured in G-quadruplex was found to interfere with the action of HIV-1 integrase.[50] Concerning HIV, other studies report how other fundamental processes for viral replication seem to be involved. To be more specific, it is known that the main structural proteins of HIV-1 are expressed as a 55 kDa polyprotein, later on processed and cleaved by a protease to obtain the mature structural proteins.[53] A class of proteins resulting from this process, called NCps, are reported to bind nucleic acids. In particular the interaction is directed to the phosphodiesteric backbone of the polynucleotide and, even if the binding motif in nonspecific, NCp shows a sequence-specific binding to runs of Gs, UGs or TGs involving zinc fingers in the interaction. The protein, and this activity leads to its definition as a "chaperone", facilitates the arrangement of the nucleic acid to a stable conformation: the G-quadruplex.[53] In addition to this sequences rich in guanines were found in other viral DNA/RNA genomes (SARS, Epstein Barr).[3]

This year another important piece of information was added to the quest for G-quadruplex stabilizers as antiviral agents. BRACO-19, an already described acridine-based G-quadruplex stabilizer,[54] was reported to show anti HIV-1 effects.[55]

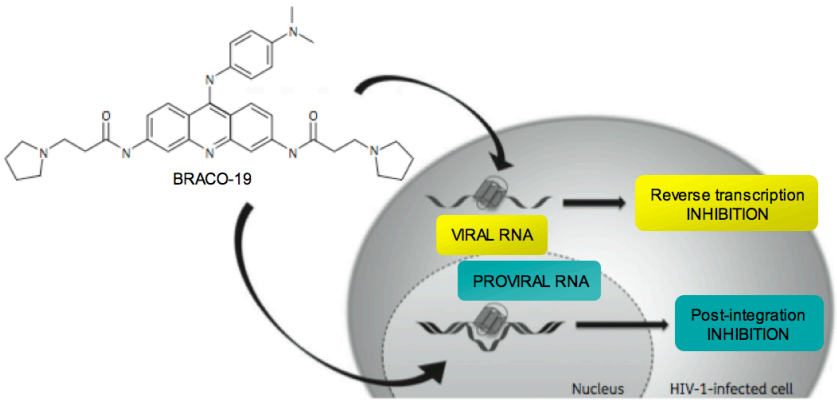


FIG2. Antiviral activity of a G-quadruplex stabilizer, adapted from Perrone *et al.*[55]

In particular this G-quadruplex stabilizer, as reported in FIG2, is proposed to be capable of acting through two different mechanisms. It can inhibit reverse transcription by stabilizing G-quadruplexes in viral RNA infecting the cell or also binding the quadruplexes formed by post-integration proviral RNA. As previously introduced, indeed, a nucleic acid arranged in

G-quadruplex is usually not processed by the enzymes. These attractive targets boost the interest for the discovery of novel G-quadruplex stabilizer and for the investigation of their binding properties with different nucleic acid (DNA, RNA or hybrids), expanding their possible application from the anticancer to the antiviral field

Many efforts are still being made also from another point of view, the structural investigation of G-quadruplex. In general, classical G-quadruplex topologies go nowadays together with novel quadruplex folds and pairing and alignments alternative to the tetrad, containing homo- and mixed- tetrads, triads, pentads, hexads and heptads.[50] This is even truer when considering more peculiar and uncommon (or maybe less popular) sequences that can form super molecular structures. An example is represented by RNA quadruplexes, involving a nucleic acid mostly found in the single stranded arrangement inside the cell.

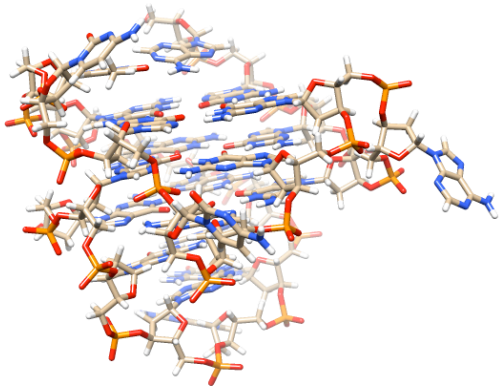


FIG17. RNA G-quadruplex.[56]

These structures are much less characterized even if they are reported to be at least as stable as DNA quadruplexes, and they are ion-dependent the same way. Their presence and role in cell life is a very recent topic, and only little information is available to date.[57]

1.8 QUANTITATIVE VISUALIZATION AND RECENT UPDATES

Inter or intramolecular G-quadruplexes are known to show various, easily interconvertible topologies. This leads to an extreme flexibility and the commonly used modeling-docking strategies can lack of information in this field. On the other hand, recent investigations made use of antibodies to quantitatively visualize G-quadruplexes inside human cells.[58]

In 2013 another important step in G-quadruplex investigation has been made: the crystal structure of human telomeric DNA complexed with berberine (a stacked alkaloid bound with a ratio higher than 1:1) has been reported.[59]

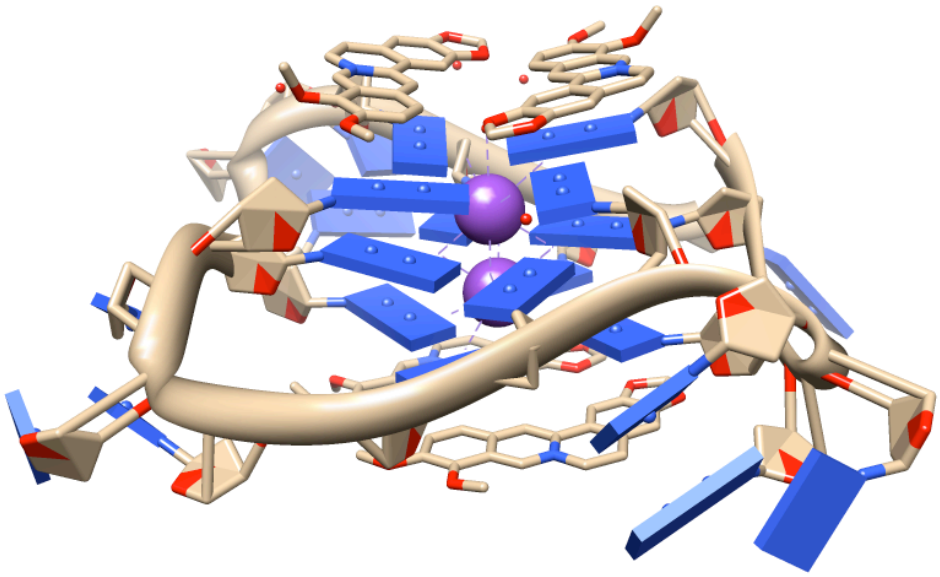


FIG18. Structure of the cited compound bound to the quadruplex in a ratio higher than 1:1.[59]

1.9 STABILIZATION OF THE G-QUADRUPLEX

Inhibition of telomerase

In anticancer therapy two different strategies can be adopted to contrast the activity of telomerase. A first attempt was the development of an action targeted to the enzyme. On the other hand, recent and more successful strategies are focused on the so-called indirect inhibition, obtained by interference with the substrate of the enzyme.[60] As far as the first strategy is concerned both the inhibition of the catalytic subunit hTERT and of the RNA subunit hTR were investigated. This strategy is based on immunotherapy using the hTERT as an antigen.[60] In the second kind of approach the target are instead the guanine-rich sequences in the telomeres. As previously mentioned, these sequences, if arranged in G-quadruplexes, can inhibit telomerase.[60] In the first strategy the action is directed against the enzyme blocking the elongation of the telomere, while the cell still survives for a good number of cell cycles, according to the length of telomere itself.[1] The stabilization of G-quadruplex allows the same long-term effect because prevents the telomerase to process the telomere but also introduces a short-term effect, as long as it is supposed to produce rapid cell senescence together with the activation of DNA damage reactions.[19] In this case the effect is due to the fact that proteins are not able to build the *shelterin* complex when the telomere is structured in G-quadruplex. This turns out in an inefficient protection of the single stranded portion.[19]

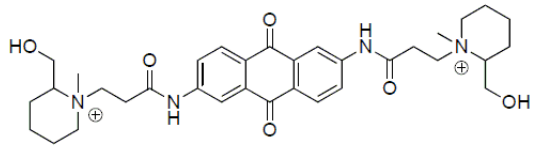
The interaction motif: stacking and beyond

A large number of molecules belonging to various chemical classes have been reported. The design of a good G-quadruplex stabilizer, anyway, still represents a hot issue, and many approaches of rational design still seem to fail. First of all it has to be considered that,

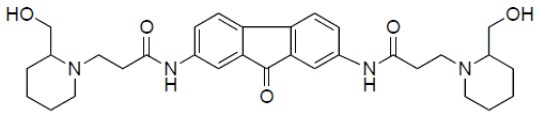
as previously stated, there is not a "one and only" structure describing the G-quadruplex, but many topologies can be identified. On the other hand, luckily, all topologies generally share three possible binding motifs with a potential stabilizer: end-stacking with an external g-quartet, intercalation between two quartets or interaction with the loops.[61] The "sandwich" pattern has been then reported: the molecule is inserted between two quartets of two different G-quadruplexes, or the molecule can be stacked between the quartet and a loop.[19] In terms of energy the most favored is the end stacking, as long as it does not disrupt the arrangement (as may happen in case of intercalation).[19] More recently other interaction patterns received growing attention. In particular, the field of the groove binders seems to be open for new discoveries: binding agents with a good selectivity for G-quadruplexes have been reported and are under study both *in silico* and *in vitro*..[62]

Reported G-quadruplex stabilizing agents

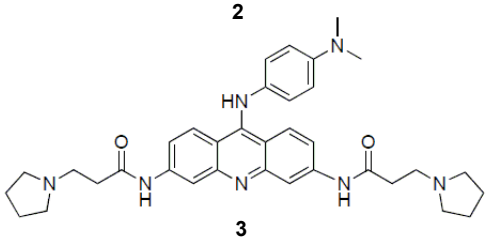
Here are reported some of the classes of compounds already known to be G-quadruplex stabilizing agents.



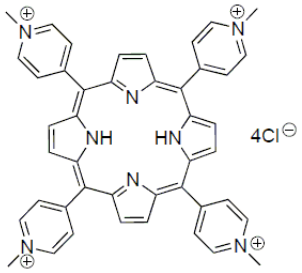
1



2



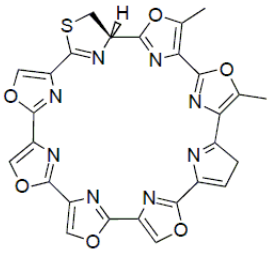
3



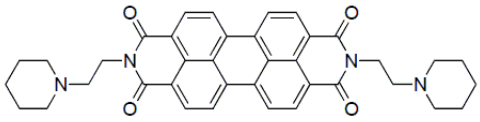
4

As will be enlightened later, most of them share some common features from a structural

point of view to satisfy the requirements for a good interaction with the quartets or, in some cases, with the loops. Some outstanding examples among these classes of compounds there are anthraquinones, fluorenones, acridines, naphthalene diimides, perylenes and telomestatine.



5



6

Anthraquinones are one of the first class of described compounds with G-quadruplex stabilizing properties. They have a polycyclic planar chromophore and generally two or more side chains as substituents of the rings in various positions. In particular the 2,6-disubstituted series showed a good activity. A good side chain can be connected to the aromatic ring through an amide group, carry a spacer and a heterocycle (piperidine or pyrrolidine). The heterocyclic nitrogens are charged because of the protonation in the physiological environment, and this promotes the interaction with the phosphate groups of the nucleic acid, showing otherwise a negative charge. Given that end stacking is supposed to be the binding motif, these compounds have not shown great selectivity for G-quadruplexes when compared to the activity towards double stranded DNA.[19] The compound BSU-1051 is reported to have an IC_{50} around 23 μM against telomerase.[61]

The scaffold of fluorenones, then, is quite similar to the one of anthraquinones but a slight improvement in the stabilization was found, with the activity of the best compound of the 2,7-disubstituted series presenting an IC_{50} of 8 μM.[63]

Acridines, as BRACO-19, show a planar chromophore carrying a nitrogen that can be

protonated at physiological pH, which is a good feature to direct the centre of the molecule towards the centre of the quadruplex, where the channel with negative density is present.[61] Some reported crystallographic studies showed that the binding motif could be a stacking of disubstituted acridine between the external quartet and a diagonal loop.[19]

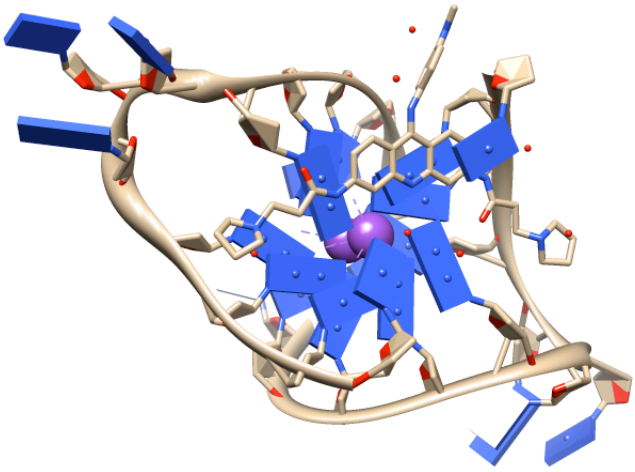


FIG19. Acridine gives stacking interaction with a G-quadruplex[42]

In addition to this, trisubstituted acridines showed a better selectivity for the G-quadruplex compared to the activity towards double stranded DNA.[19]

Porphyrines show an extended planar chromophore that could perfectly match the whole quartet. The tetra-N-methylpyridino derivatives, such as TMPyP4, are selective for the quadruplex. Also porphyrines carrying a coordinated cation (Mn^{3+}) in the center of the structure were reported and showed a high degree of selectivity for the quadruplexes (10000 fold when compared to the tendency of binding a double stranded DNA).[19]

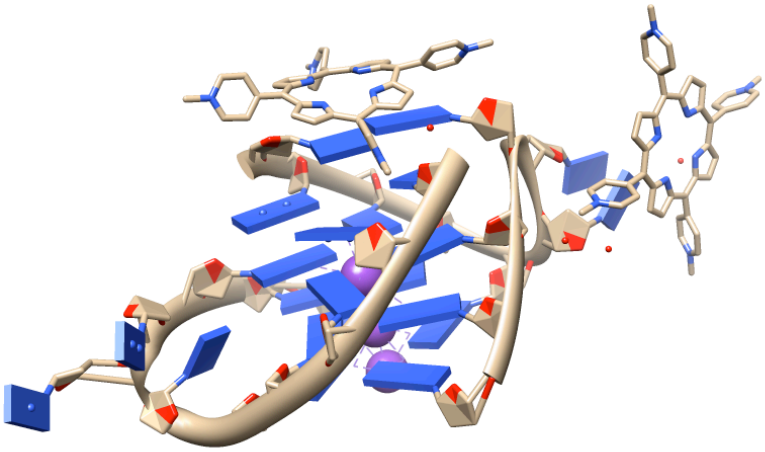


FIG20. Two different binding motifs for TMPyP4.[64]

Naphthalene diimides feature four side chains with a terminal heterocycle that allow a good selectivity for the G-quadruplex.[19] Crystallographic studies showed a remarkable end-stacking binding motif.[65]

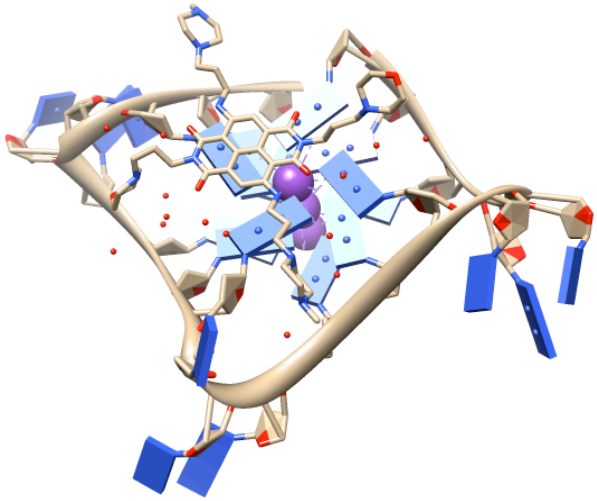


FIG21. Naphtalene diimide binds the G-quadruplex through stacking interaction[43]

Perylenes were designed with the aim of increasing the planar surface of the scaffold. Imide groups and side chains ending with heterocycles were conserved in order to maintain the same overall structural features. PIPER (IC50 0.2 μM) is an outstanding example of this class.[66] NMR studies showed again that the binding motif with G-quadruplex is the end stacking.[67]

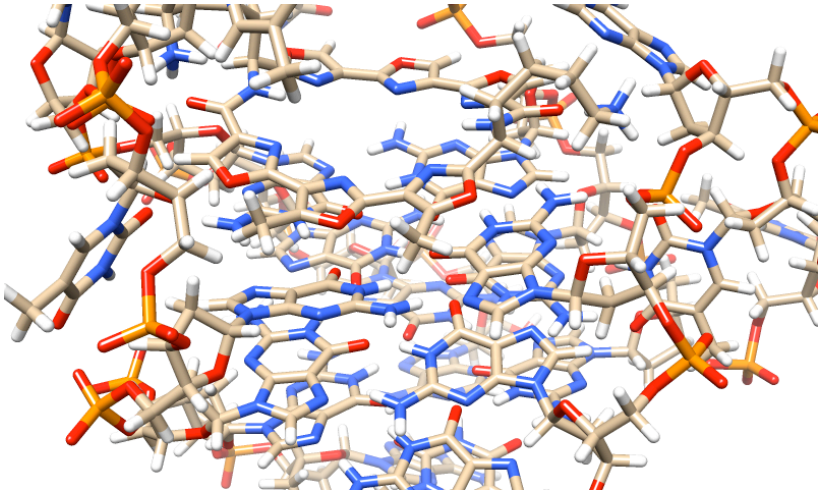


FIG22. A telomestatine derivative bound to the quadruplex.[68]

Telomestatine, with its oxazoles backbone, show san excellent overlap with G-quartet in terms of size.[19] This compound is one of the most effective ligands of G-quadruplex with an IC_{50} against telomerase of 5 nM.

1.10 THE LEAD COMPOUND

A good pharmacophore

The analysis of the already reported stabilizers of the quadruplex allows examining and describing the required features to obtain the anti-telomerase activity and, together with this, the selectivity for the quadruplex against the double stranded DNA. As can be observed in the structures of the compounds cited above, two main features can be enlightened:

- A planar aromatic chromophore, which binds to the planar structure of the G-quartet through π-π interactions. The size of this planar surface can also drive to greater selectivity for the G-quadruplex. In addition to this the presence of a nitrogen atom in the scaffold or a metal ion bound through coordination would lead to an even more successful interaction, as previously described
- Two, three of four side chains that can give interactions with the loops and the grooves. Two important features are the chain length and the presence at their end of heterocycles carrying nitrogen atoms with a positive charge in the physiological environment.

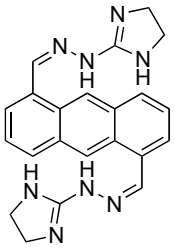
The active lead compound, as will be showed later, presents all the characteristics listed above.

Lead compound and modifications

The compound that was taken as a lead in this research project shows a 1,5-disubstituted anthracene scaffold. In addition to this the side chains are charged with two nitrogens from the hydrazone group and two nitrogens from the 4,5-dihydroimidazole ring resulting in the

possibility of having multiple positive charge states available. The compounds that were newly synthesized in the first part of the project were designed keeping in mind these features reported to be required for a good activity, but some modification (slight or heavier) were applied to the structure to analyze the outcoming effect in terms of binding and stabilization activity. In particular, three main aspects were considered:

- the heterocycle, focusing the attention on the size and the heteroatoms present
- the role (if any) of the Cα connected to the aromatic scaffold and of its hybridization
- the length and the chemical properties of the side chains



7

The project later on expanded with synthetic schemes based on other scaffolds and the subsequent attempts of derivatization. In particular, as will be explained later, bisanthrapyrazole was found to be a good candidate, thanks to its scaffold carrying five condensated rings with four nitrogen atoms included. To provide further information about both the role of the scaffold and the side chains, two well known already cited classes of compounds were eventually considered: anthraquinones and acridines. During the latest part of the project then, some structurally constrained compounds with a structure extremely close to the one of the lead compound were designed and obtained. The aim of this additional synthetic scheme was to enlighten the role of the flexibility of the side chains, a feature that seems to overcome even their chemical nature in terms of relevance for an efficient interaction with the G-quadruplex.

1.11 COMPUTATIONAL STUDIES: LIMITATIONS AND PERSPECTIVES

The synthetic process can be aided by some advices coming from computational studies: although molecular modeling simulations over G-quadruplex-ligand interactions are developing and becoming more and more reliable, only in the very recent period a structural analysis of both the structures of G-quadruplexes (solved thanks to NMR or X-rays) and the 3D models of the ligands and their interactions became reliable in terms of predicting activity properties. In this connection, some recent research works started to consider the combination of ligand-based (the investigation of the structural properties of the binders) and structure-based (an analysis that focuses on the target of the binder) virtual screening approaches in the search for novel G-quadruplex stabilizing agents.[69] While this technology enlighten an innovative workflow for the discovery of novel potential stabilizers through library screening and then *in vitro* assay, it still suffers for some limitations related to the peculiar features of the G-quadruplex itself. Docking experiments have to face the fact that human telomeric sequence has been shown to fold in at least four different structures with completely different motifs of guanine disposition and strand/loop orientation.[69] Considering all the possible viabilities would turn out in an impossibility to represent all the realistically possible folds in a reasonable time, in terms of calculations.[69] At the moment, then, the solution is to summarize all the possible and easily inter-convertible arrangements with one (or more, in the more sophisticated experiments) model structure to run dockings or dynamic simulations.[69] This strategy is eventually, over the time, growing from rough approximation to predictive model.

The synthesis of novel compounds should anyway take advantage of the knowledge that can be obtained by the evaluation of some peculiar structural features of both the ligands and the quadruplex.

A first consideration should be, for example, one that compares the effective extension,

basing on real telomeric sequences, of the already well-described planar surface of the quadruplex that is available for stacking interaction

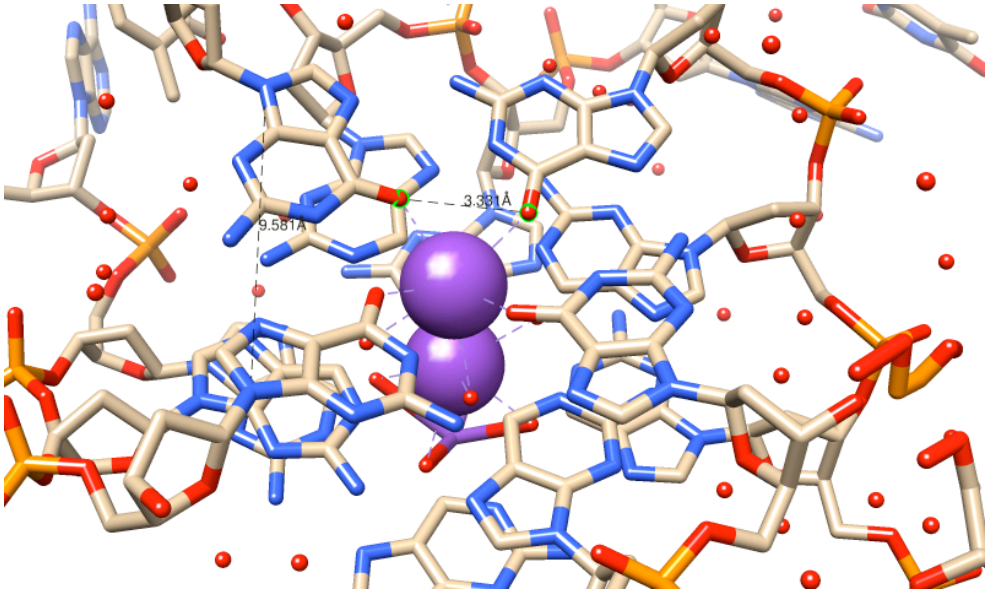


FIG23. A closer view to the planar surface of a telomeric G-quadruplex.[43]

If we assume that the four carbonilic oxygen atoms coming from the four guanines lay on the same plane, they are describing a square with the side of 3.4 Å. This corresponds to what stated in the introduction concerning the size of the tunnel where the ions are inserted. If we otherwise look at the square described by the nitrogen atoms of the guanines connected to the ribose ring, this bigger square has a side that exceeds 9 Å, and this turns out in a wide, planar and prone to aromatic stacking interaction surface. A molecule with a good interaction profile should take advantage of this feature and, in addition to that, be able to give interaction with the bases or the phosphate groups describing the loops.

The compound that is presented as the lead in this study perfectly fit this DNA arrangement.

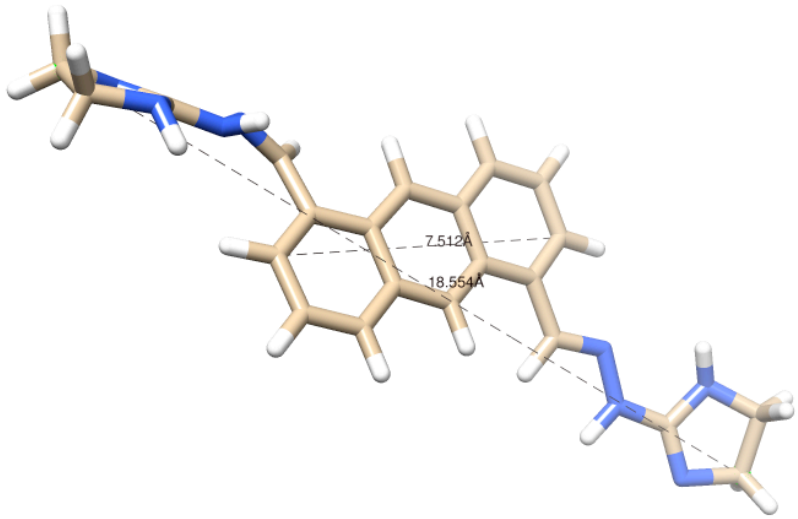


FIG24. Structural analysis on the lead compound.

Focusing on the anthracene scaffold, is evident how the size of the three condensed ring, with a diagonal of more than 7 Å from carbon to carbon (the actual size is higher if the actual bulge of the molecule is considered), gives a surface with the same order of magnitude of the one described for the quartet. On the other hand, it is also noticeable that the lead compound shows two extended side chains able to largely exceed the size of the quartet to give, potentially, an interaction with the loops.

In this regard, the following picture shows also how these general rules concerning the size of the scaffold and of the side chains are well conserved between the lead compound and other already reported and well-known stabilizers. As an example, when compared to BRACO-19[43], the 1,5-disubstituted anthracene derivative shows a perfect superimposition of its scaffold to the acridine system. In addition to this, the side chains are, even if pointing in different directions because of the position in which the substitution of the ring occurs, of a comparable size. The energy of the structures here presented was minimized (see the experimental section for further details). As was already described, at the same way, the side chains shows in this case protonable nitrogen atoms at their ends.

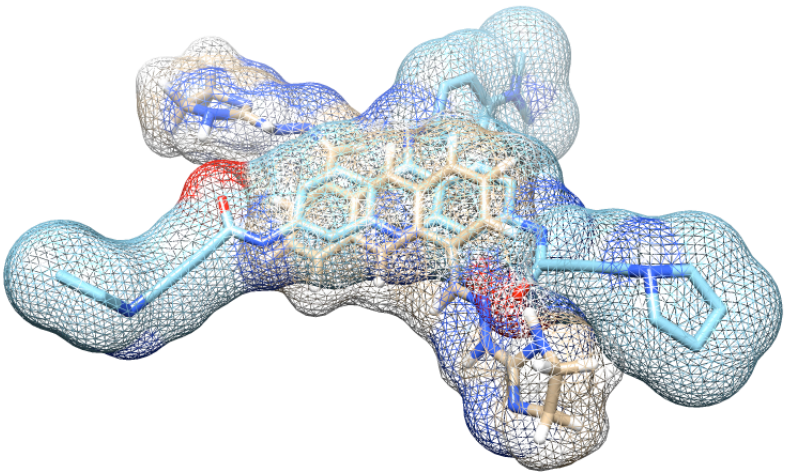


FIG25. Comparison between minimized structures of BRACO-19 and the lead compound.

On the other hand, this comparison also shows also how the overall bulge of the two molecules is, overall the same.

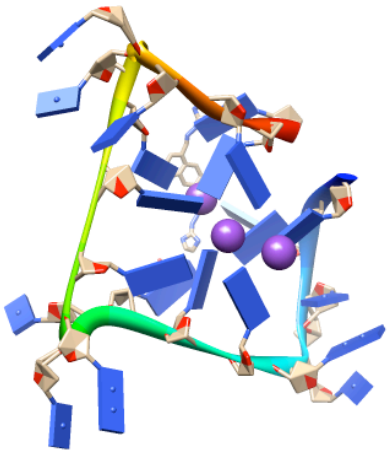


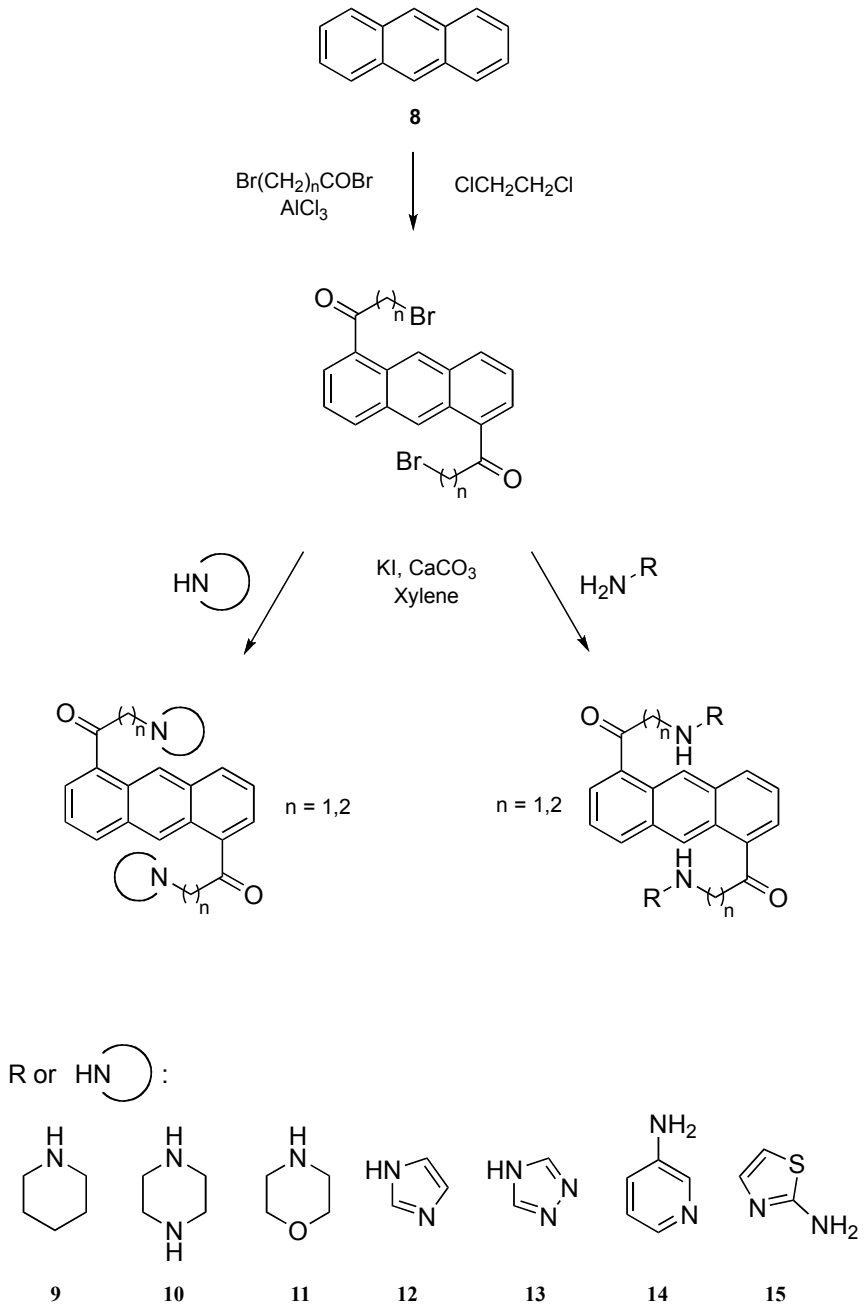
FIG26. Model showing a representation of a possible binding motif.[43]

Concluding this overview on the structural features of the lead compound related to the ones of other ligands and of the G-quadruplex, the model above shows a possible interaction motif between the 1,5 disubstituted anthracene derivative and a structured telomeric sequence of DNA (model of the G-quadruplex DNA from Micco et al.[43]). Docking studies were performed on some of the synthesized compounds to assess the possible binding motif of the molecules to the structured DNA, according to the recent

literature. Obtained models are presented in the following sections and the docking procedure is described in the experimental part.

2. RESULTS AND DISCUSSION

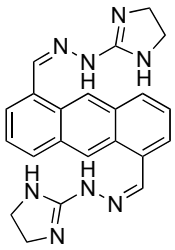
2.1 SCHEME I



9 **10** **11** **12** **13** **14** **15**

Rationale of the synthetic scheme

This synthetic scheme focuses on the synthesis of 1,5-disubstituted analogues of the lead compound **7**.



7

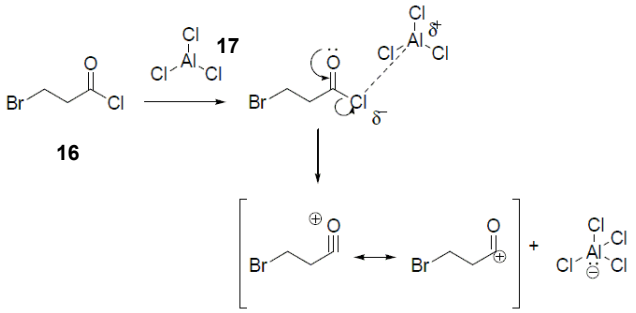
This first synthetic strategy is being developed in order to obtain some 1,5 disubstituted anthracene derivatives to provide additional information in terms of structure-activity relationship. In fact, even though is known to be very effective in stabilizing telomeric G-quadruplexes structure because of its planar π-rich surface, the role or the importance of the imidazoyl containing side chain, and as later will be explained of their conformation, are still poorly understood.

On the other hand also bioisosteric substitution of the hydrazone chains is to be investigated yet. To be more specific in this scheme the keto derivatives are described. In addition to this, some length variations have been tested.

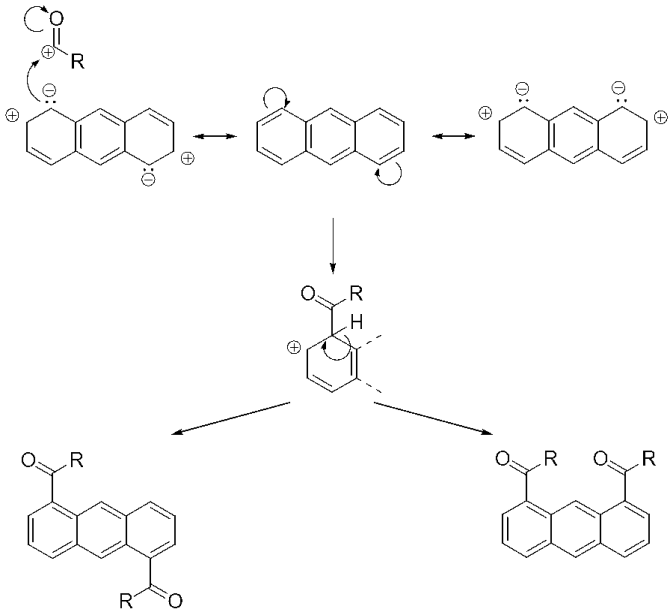
Thanks to this first synthetic scheme some basic derivatives were obtained. All the obtained compounds show a keto group on the side chain with a two or three atoms long spacer. Further modifications and variations in particular of the side chains were introduced in the second synthetic scheme.

Friedel-Craft acylation

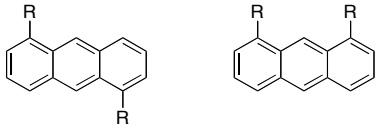
The first step of this reaction scheme is the one that provides a key intermediate: an anthracene with two points for further derivatization in the positions 1 and 5. The reaction consists in an electrophilic aromatic addiction operated by an acyl cation generated as a consequence of the activation of the acyl halide by the stoichiometric *Lewis* acid $AlCl_3$. The reaction was carried out in dichloroethane with yields ranging from 30 to 40 % of recrystalized products.



The *Lewis* acid forms a complex with the acyl chloride. Here the reaction with 3-bromo propionyl chloride is presented, but this step has been carried out also with other reagents: 2-bromo acetyl chloride, acetyl chloride, 3-chloro propionyl chloride.



The resulting carbocation is stabilized by resonance and is a good electrophilic reactive species.[70] Once this is formed, it rapidly reacts with the aromatic system of anthracene. As the scheme above demonstrates, according to this aromatic feature, the negative charge coming from delocalized electrons can be concentrated, because of the resonance, in some positions of the rings more than in others. In fact, the critical and peculiar aspect of this reaction consists in the regioselectivity that can be induced by varying the reaction conditions. Theoretically an anthracene acylation through the procedure reported above should very likely provide a nearly 1:1 mixture of 1,5 and 1,8 disubsituted derivatives.



Despite this what we experimentally experienced is a huge prevalence of the 1,5 isomer (over 90 %) when the reaction is carried out under 0°C and the anthracene is rapidly added to a stirred mixture of $AlCl_3$ and the acyl halide in dichloroethane. This event was confirmed by other studies on this class of reaction previously reported in literature.[71] The desired product can be easily isolated by recrystalization from toluene of the crude.

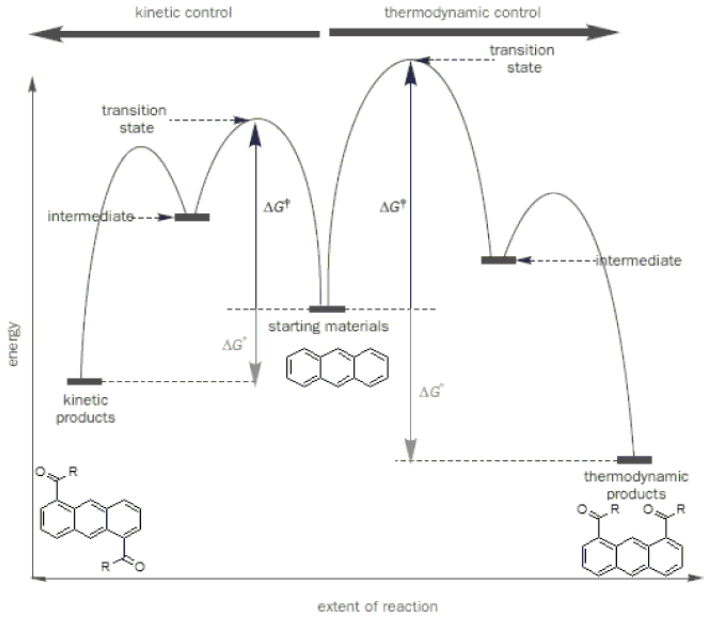
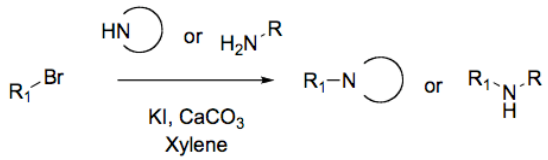


FIG27. Kinetic and thermodynamic drive of the reaction.[72]

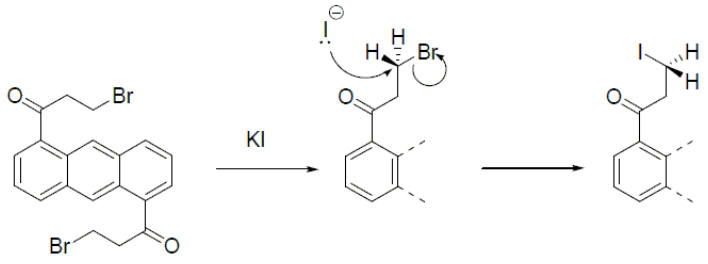
The proposed interpretation of the observed results involves the kinetic/thermodynamic control of the reaction coordinate. The 1,8 disubstituted isomer is proposed to be the thermodynamically induced product, ranking at a lower energy level but produced only after the take over of a considerable activation energy. On the other hand the 1,5 isomer can be considered as the kinetic product, preferentially provided in low-energetic conditions. As a result, this reaction carried out in a cold bath at a controlled temperature allows obtaining the desired isomer in good yields (generally over 30%). A clear demonstration of the fact that 1,5 isomer was isolated was given by the analysis of NMR spectra, where the coupling constants between different signals were compatible with this species (see experimental procedures for the detailed transcription of the spectra).

***Finkelstein* reaction: an efficient nucleophilic substitution**

The intermediates obtained with the reaction described above are of a strategic relevance. After that, once that the scaffold its built, chlorine or bromine atoms in the side chains can be used as linking points to add the desired heterocycle or to enlarge the chain.



According to the fact that the halide is close to a carbonyl group (depending on the acyl chloride/bromide used), its reactivity towards nucleophilic substitutions should be enhanced because of the electron drawing effect and it should represents a good leaving group. Despite this, the early attempts of carrying out the reaction between the anthracene derivative and an amine were unsuccessful. To improve the reactivity, then, the reaction was carried out according to the *Finkelstein* procedure. This consists in the substitution of the halogen in the side chains of the intermediates (Br or Cl) with a more reactive iodine atom.[73] The subsequent nucleophilic substitution of the halogen by the amino derivatives proceeded smoothly in most of the cases. The reaction was carried out in xylene in the presence of KI and $CaCO_3$.



17

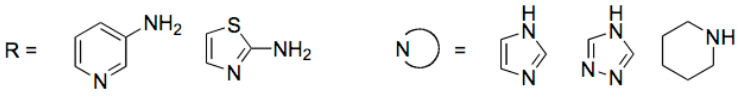
This reaction proceeds through an S_N2 mechanism and comes with the inversion of the configuration of the carbon that exchanges the bindings. The different solubility of the involved halide salts is another help to the driving force of the reaction since it influences the equilibrium position: the resulting KBr is poorly soluble. To be more specific, then, iodine represents a really good leaving group, under these conditions, because of two events: the strength of the C-X bond and the stability on solution of the halide ion, once the C-X bond is broken.

Table 17.13 Halide leaving groups in the S_N1 and S_N2 reactions

Halide (X)	Strength of C–X bond,[1] kJ mol⁻	pK_a of HX
fluorine	118	+3
chlorine	81	–7
bromine	67	–9
iodine	54	–10

FIG28. Halide leaving groups: a ranking.[72]

The strength of the bond can be measured in kJ/mol, while in the table reported above the pK_a of the HX acid gives an indication of the stability of X^- ion alone: the higher the value of the constant, the lower the degree of dissociation of the HX acid. As a result, basing on the fact that both the pK_a value for HI and also the strength of the C-I bond is the lowest among the ones reported for fluorine, chlorine or bromine, I^- represents a better leaving group.



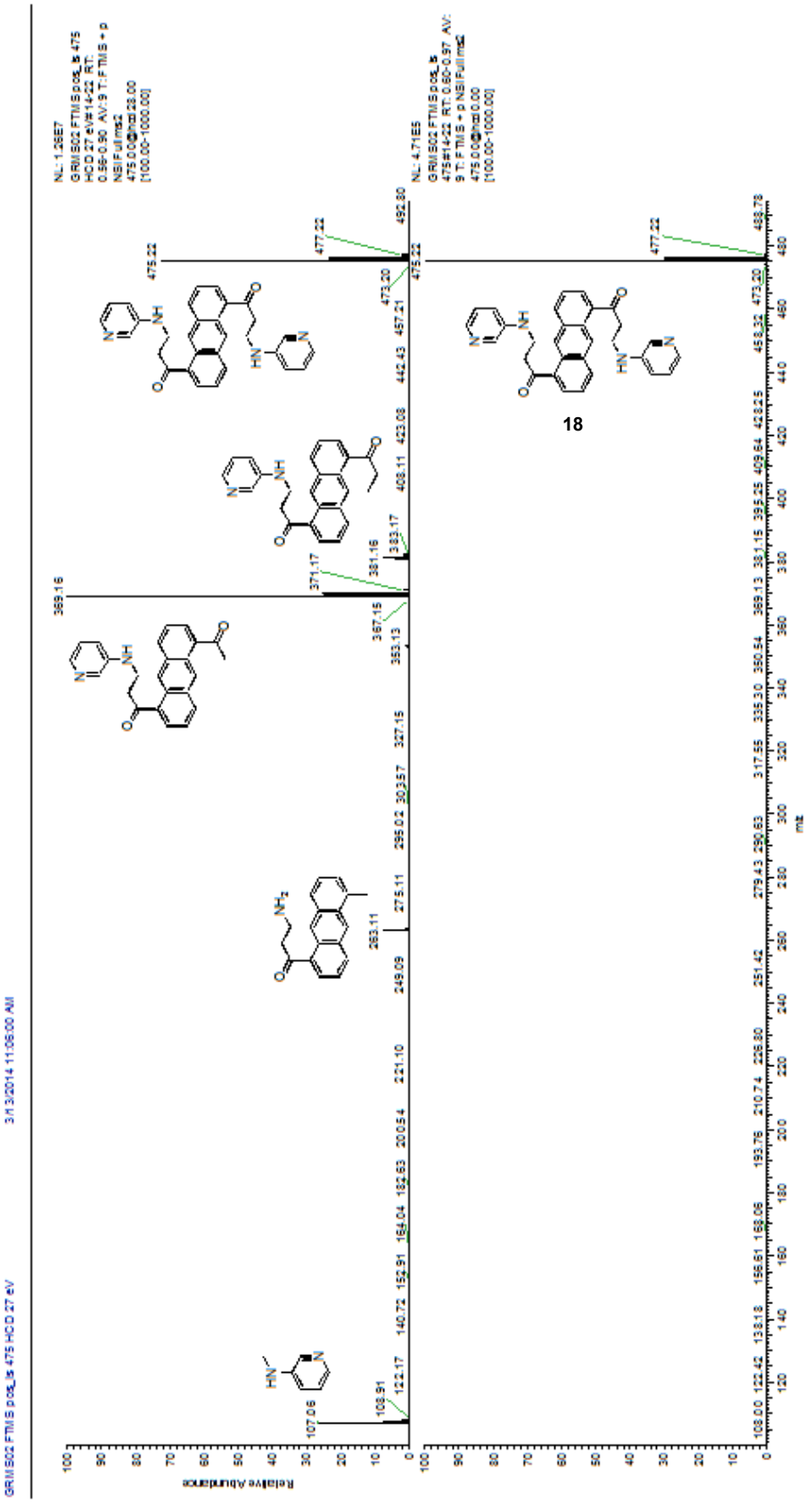
14 **15** **12** **13** **9**

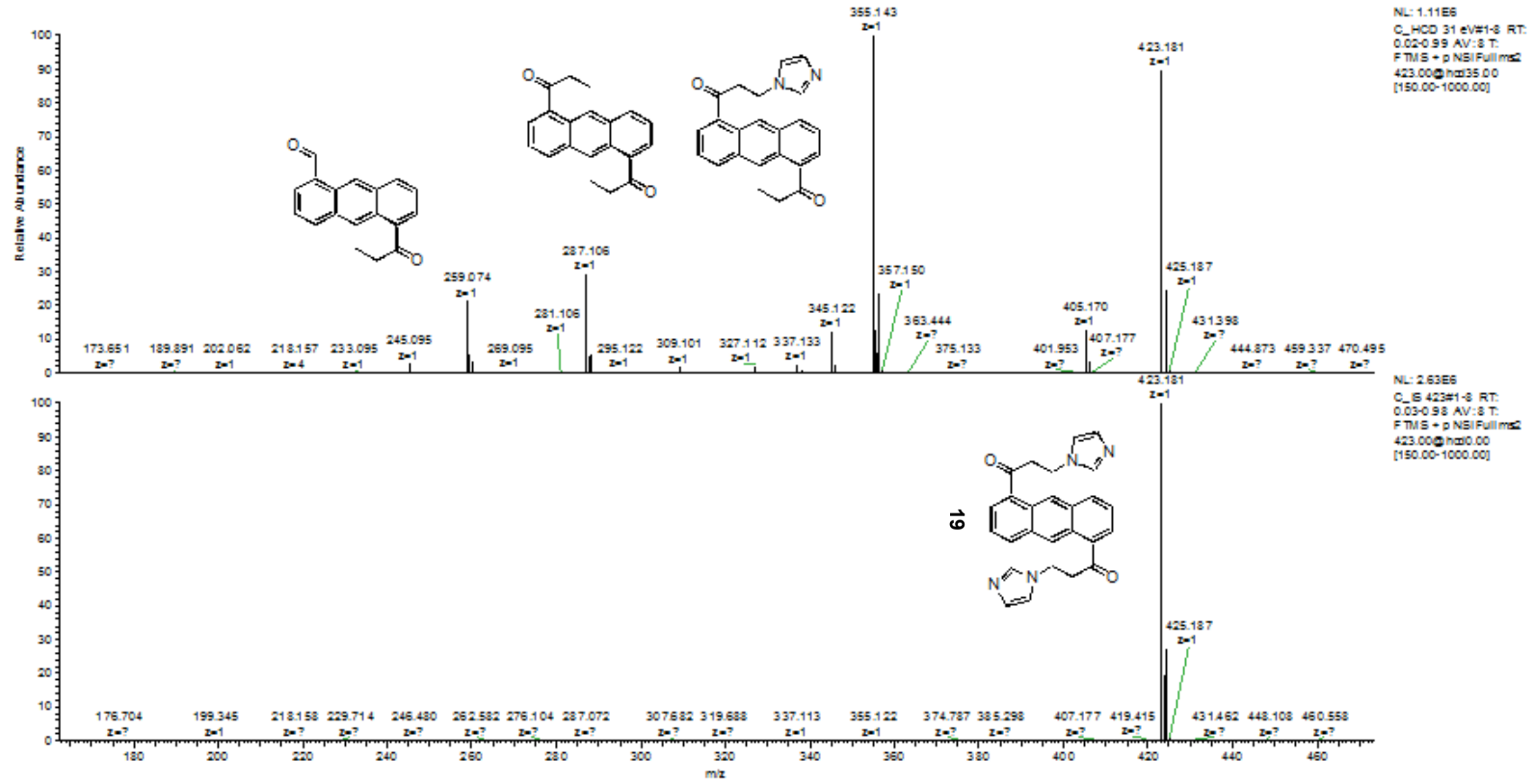
Once that the substitution took place, the resulting intermediate smoothly reacts with the reported amines to give the desired products.

Characterization of the products

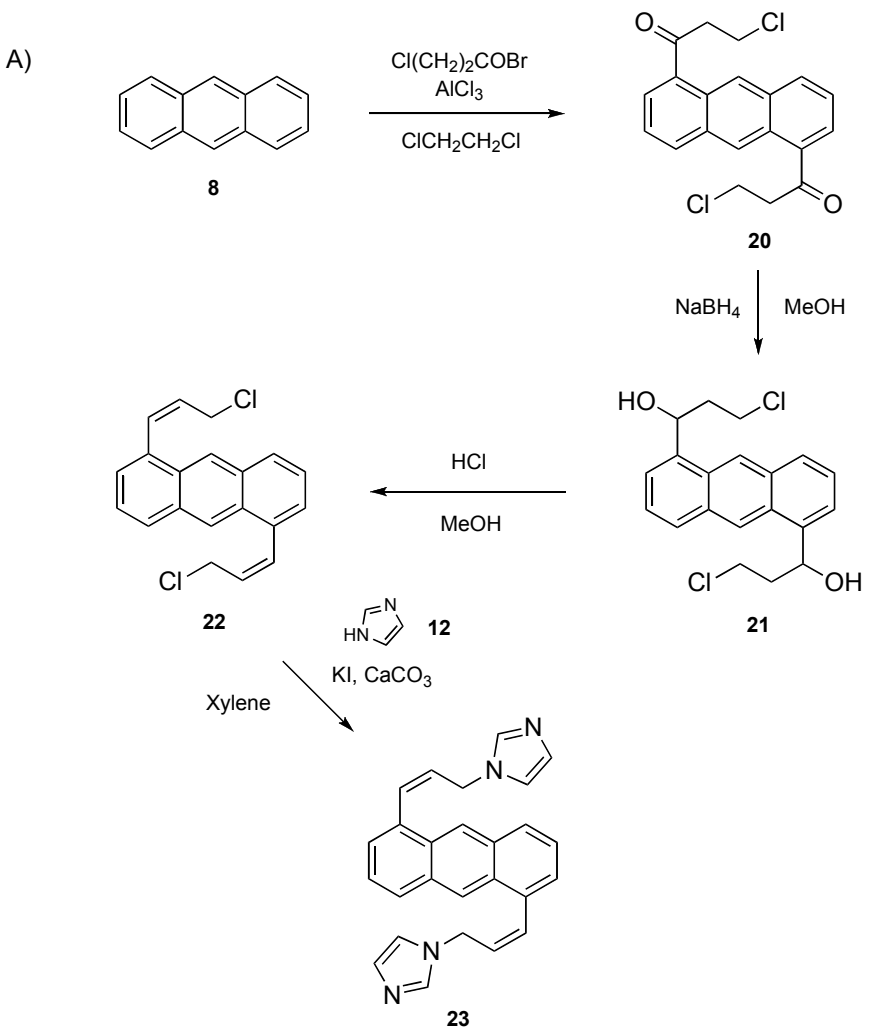
Every compound was deeply investigated through NMR and mass spectrometry experiments (see experimental section for details). In addition to this, more peculiar experiments were carried out in some of the molecules to further investigate some structural properties and confirm once again the structure.

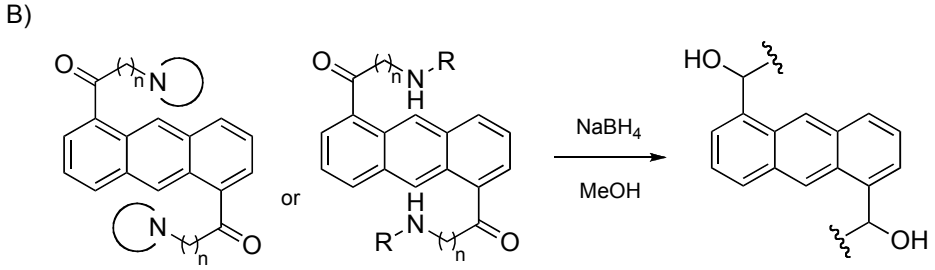
As an example, some of the compounds were analyzed thanks to the facilities available at UAlbany - State University of New York, USA. In particular, the spectra reported in the following page represent a so-called MS-MS tandem mass spectrometry experiment. The LTQ OrbiTrap Velos spectrometer used to record this spectrum allows the isolation of a single peak with its isotopic distribution (lower part of the spectrum) in the full scan mass spectrum. The second part of the experiment is the fragmentation (upper part of the spectrum) was a growing amount of energy is provided to the isolated ion, forcing its decomposition. The fragments resulting from that event could help in the correct interpretation of the structure of the molecule itself. In the spectrum reported below, many fragments were identified confirming the desired structured.


18


M:\Giovanni\041714\C_HCD 31 eV
4/17/2014 6:18:58 PM
NL: 1.11E6
C_HCD 31 eV#1-8 RT:
0.02-0.99 AV: 8 T:
FTMS + p NSI Full ms2
423.00@hcd35.00
[150.00-1000.00]
NL: 2.63E6
C_IS 423#1-8 RT:
0.03-0.98 AV: 8 T:
FTMS + p NSI Full ms2
423.00@hcd0.00
[150.00-1000.00]
Relative Abundance
m/z
19

2.2 SCHEME II

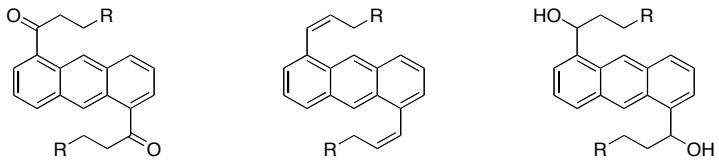

A)
Cl(CH2)2COBr
AlCl3
ClCH2CH2Cl
8
20
NaBH4
MeOH
HCl
MeOH
22
12
21
KI, CaCO3
Xylene
23


B)
or
NaBH4
MeOH

Rationale of the synthetic scheme

In the first synthetic scheme a novel pathway to obtain 1,5 disubstituted anthracene derivatives was enlightened. In addition to this bioisiosteric substitutions of the heterocyclic system were operated. In this scheme further modifications of the side chains were investigated. In particular the keto group was reacted in opportune conditions to obtain derivatives showing hydroxy or C=C double bonds.

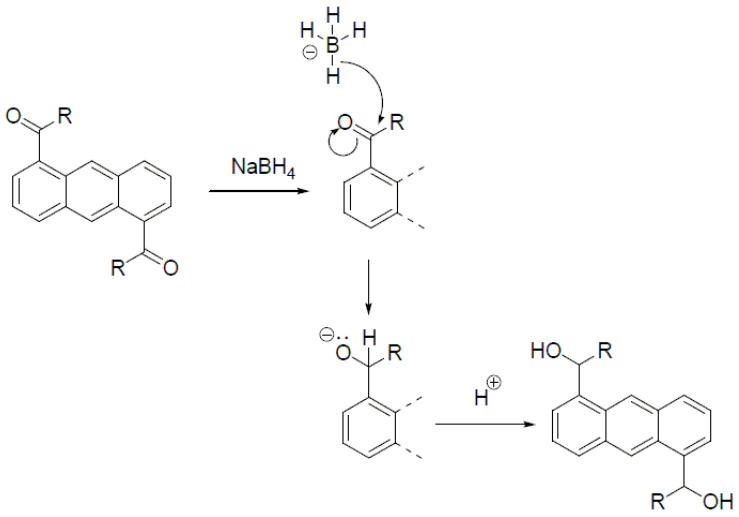
The rationale of these modification was the idea of investigating the influence of different spatial bulkiness due to different carbon hybridizations. To be more specific, two different classes of compounds with an sp^2 carbon atoms in the side chains (keto and alkene derivatives) and a class with only sp^3 carbon atoms in the side chains (hydroxy derivatives) were synthesized to be then compared. Different hybridizations lead to different geometry, spatial orientation, flexibility and rotational freedom.



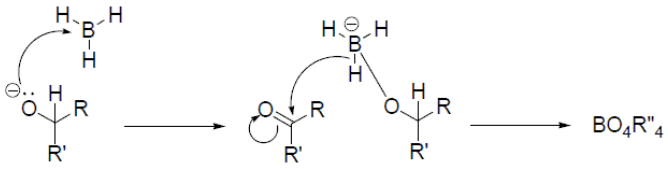
NMR studies were used to assess the correct orientation of the double bond. A particular focus was dedicated to the compounds carrying an imidazole ring, as long as, as will be later explained, were the ones showing more promising activity profiles.

Reduction of a keto group with $NaBH_4$

This reaction, carried out in methanol at room temperature, was designed both to operate a direct reduction of the compounds synthesized in the first scheme to obtain the corresponding hydroxy derivatives and to synthesize an important intermediate for the synthesis of alkene derivatives.



The reduction with $NaBH_4$ proceeds thanks to the transfer of a hydrogen atom with its electronic doublet to the substrate that has to be reduced.[72]

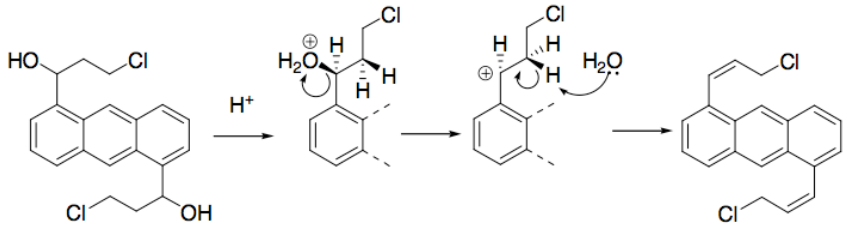


The anion generated by this addition is proposed to stabilize the resulting BH_3 species that is technically able to transfer other hydrogen atoms and electrons. The final compound is isolated after addition of diluted HCl (10%).

This reaction was carried out successfully both using as substrates final compounds from the first scheme and the halogenated intermediates.

From the hydroxy derivative to the alkene

This key step provided the second intermediate for the synthetic route for the obtainment of alkene derivatives.

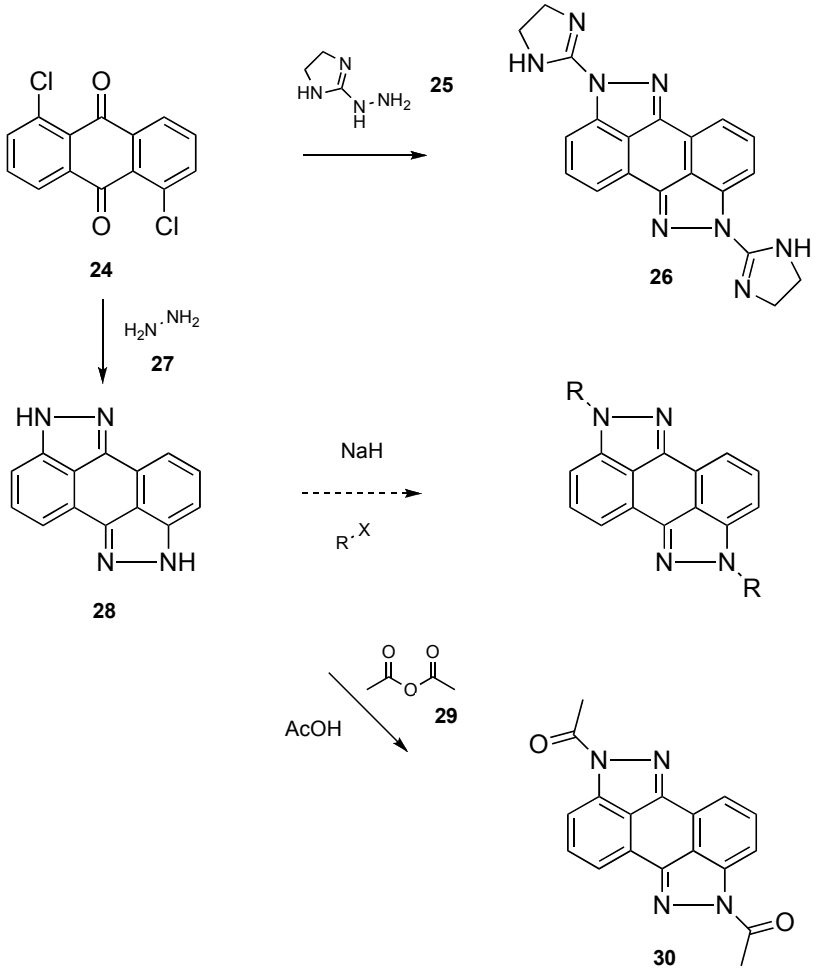


21 **22**

The reaction starts from the alcoholic derivative that is treated with HCl. The hydroxy group, that is the only protonable region of the molecule, acquires a positive charge that transforms it in a good leaving group. Loss of water leads so to the obtainment of the key intermediate with a C=C double bond in both the side chains. This compound can be treated with nucleophilic amines to obtain the final compounds as described in the first synthetic scheme.

An E configuration of the double bond could be expected for this kind of mechanism, but what was observed thanks to NMR analysis was that the obtained compound was in the Z form, as demonstrated bi the 9.1 Hz value of the coupling constant (J) in the proton NMR spectrum.

2.3 SCHEME III



Rationale of the synthetic scheme

Anthrapyrazole derivatives are being synthesized as a strategic family of compounds in G-quadruplex stabilization. The anthrapyrazole nucleus could represent an ideal starting point for the development of novel stabilizers: it shows a more extended π area comparing with anthracene and it maintains H-bond donor/acceptor sites represented by nitrogen atoms. A structural comparison between this scaffold and the lead compound shows how most of their aromatic surface is shared.

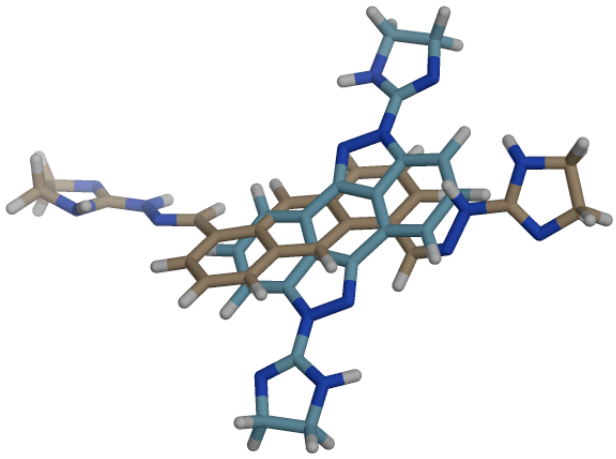
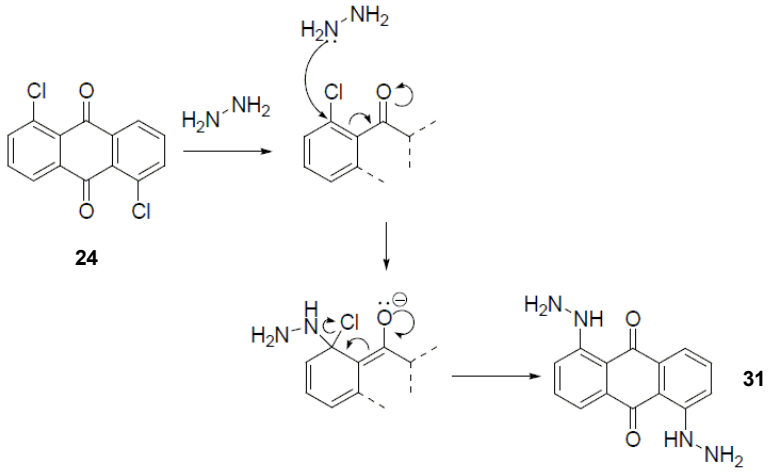


FIG29. Superimposition of anthrapyrazole and anthracene

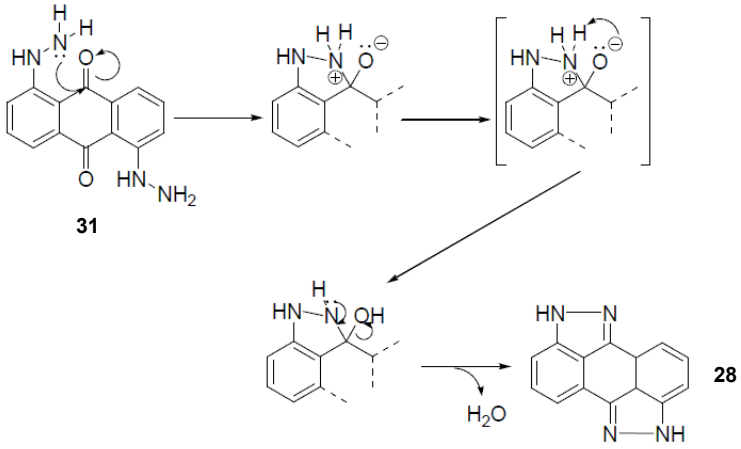
In this synthetic scheme two strategies are involved. In order to obtain N-substituted derivatives of anthrapyrazole, a first attempt was based on the synthesis of the scaffold itself as a first step starting from hydrazine and 1,5-dichloro anthraquinone. The second step, then would have been a further derivatization on the nitrogen atoms through nucleophilic substitution. The second strategy, otherwise, was designed in order to take advantage of the commercially available hydrazo imidazoline. In this way a final compound showing the exact same side chain of the lead compound was obtained in only one step.

Synthesis of the anthrapyrazole scaffold

The synthesis of a plain anthrapyrazole scaffold was operated starting from 1,5-dichloroanthraquinone and hydrazine.



The procedure for the preparation of compounds from this class was previously reported in the patent literature.[74] Several attempts were necessary, anyway, to obtain the desired product tuning the reaction conditions. As the first part of this scheme shows (above), the first step of this reaction is supposed to be the substitution of the chlorine atom. This is confirmed by the fact that in the cited patent the isolation of 1,5 di substituted compounds, without chlorine but still showing the keto group in positions 9 and 10, was reported.[74] The reaction was both carried out in a round bottom flask in dimethyl acetamide or without any solvent in a microwave reactor. The reaction conditions, anyway, are still being optimized as long as the yields are currently still low (see the experimental section for further details).



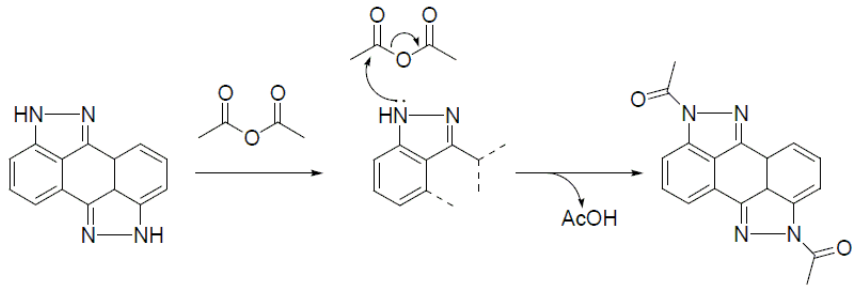
The following step shows an intramolecular reaction, where a new five members ring is formed. In this case the rearrangement of the molecule leads to the loss of water.

Considerations on the conditions for the synthesis of anthrapyrazole

The fact that the first step of the reaction is represented by the substitution of the chlorine atom on the aromatic ring is also confirmed by the fact that the collateral products obtained during the purification process by flash chromatography (mono or di substituted species) were not showing any chlorine isotopic pattern in the MS analysis. The preliminary attempts of obtaining the anthrapyrazole required several purification steps such as automatic column chromatography followed by a preparative TLC. As previously introduced, several attempts to carry out the reaction in different solvents were performed. Dimethylformamide was substituted with dimethylacetamide (a solvent that shows an even higher boiling point) and in addition to that the reaction was also carried out without any solvent and\or in the microwave reactor. In every case, anyway, one or more purification steps were required.

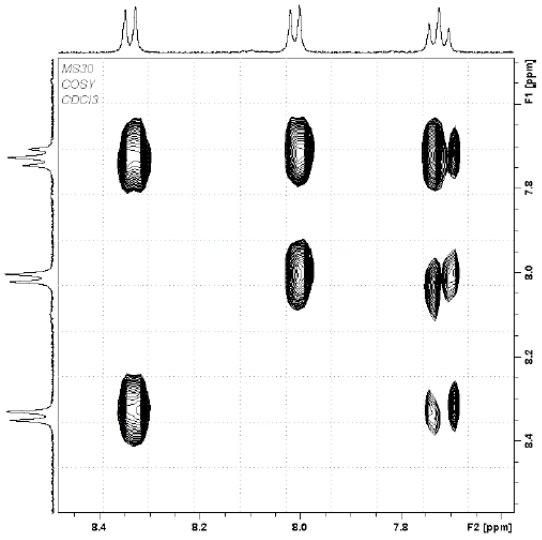
N-derivatization of the scaffold

In the original synthetic strategy the obtained scaffold was supposed to be ready to be modified through substitutions on the nitrogen. Anyway this reaction turned out to be more unwilling to happen than what expected. The experiments using an alkyl halide as substrate for the substitution and different bases (from triethylamine to NaH) were not successful. On the other hand, the reaction of anthrapyrazole with acetic anhydride (with acetic anhydride acting both as a solvent and a reactant) led to the N-substituted compound in good yield.



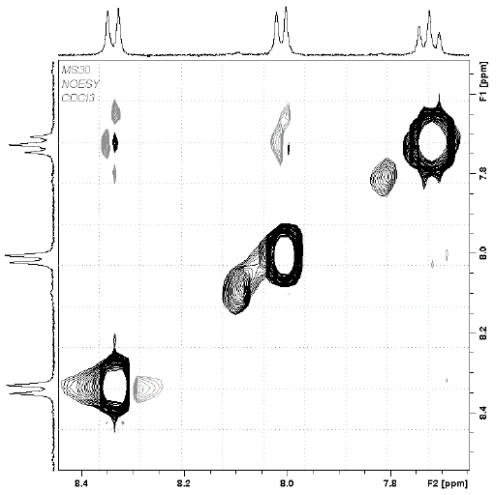
28 **30**

The obtained compound was deeply characterized thanks to the use of NMR spectroscopy.

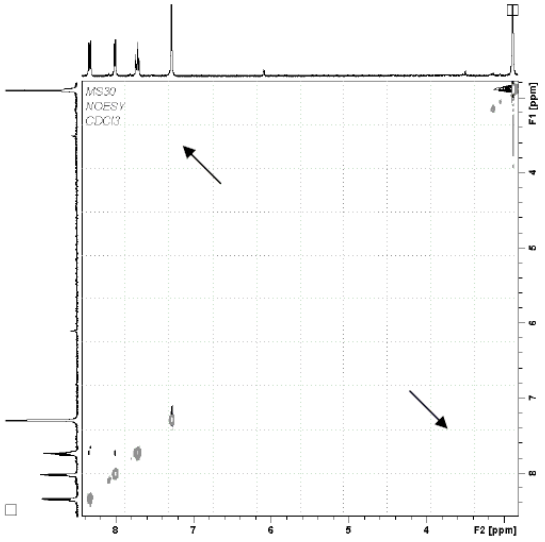


The CoSY experiment is designed to show, on a two dimensional surface, the coupling that occurs between two hydrogen atoms connected by the skeleton of the molecule. Every cross-peak outside the diagonal line of the spectrum can be considered as representative of

a coupling. The spectrum above shows the behavior and the coupling motif of the protons belonging to the aromatic part of the molecule.



The NOESY experiment reported above provides, otherwise, some complimentary information: it enlightens whether or not two protons, even if one far away from the other considering the 2D structure of the molecule are near in the 3D space due to some specific conformations. Even in this case, the cross-peak on a two dimensional surface is the signal that shows that two protons are close one to the other. In the spectrum above the spatial correlation of the aromatic system confirms the information from the previous spectrum.

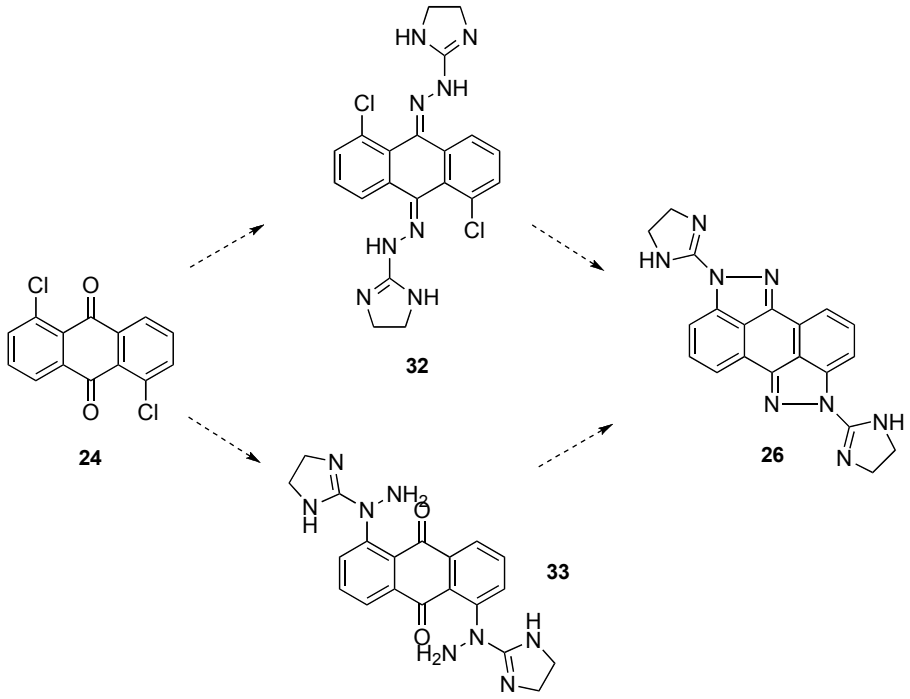


The full scale experiment (above) provides then an even more useful information about the

conformational behavior of the molecule: the aromatic portion does not give any cross peak with the -CH_3 moiety coming from the acetyl group, suggesting that the carbonylic oxygen is the closest atom to the scaffold.

One-step reaction with hydrazo imidazoline

A later strategy involved a direct synthesis of a compound structurally related to the lead using a one-step procedure.



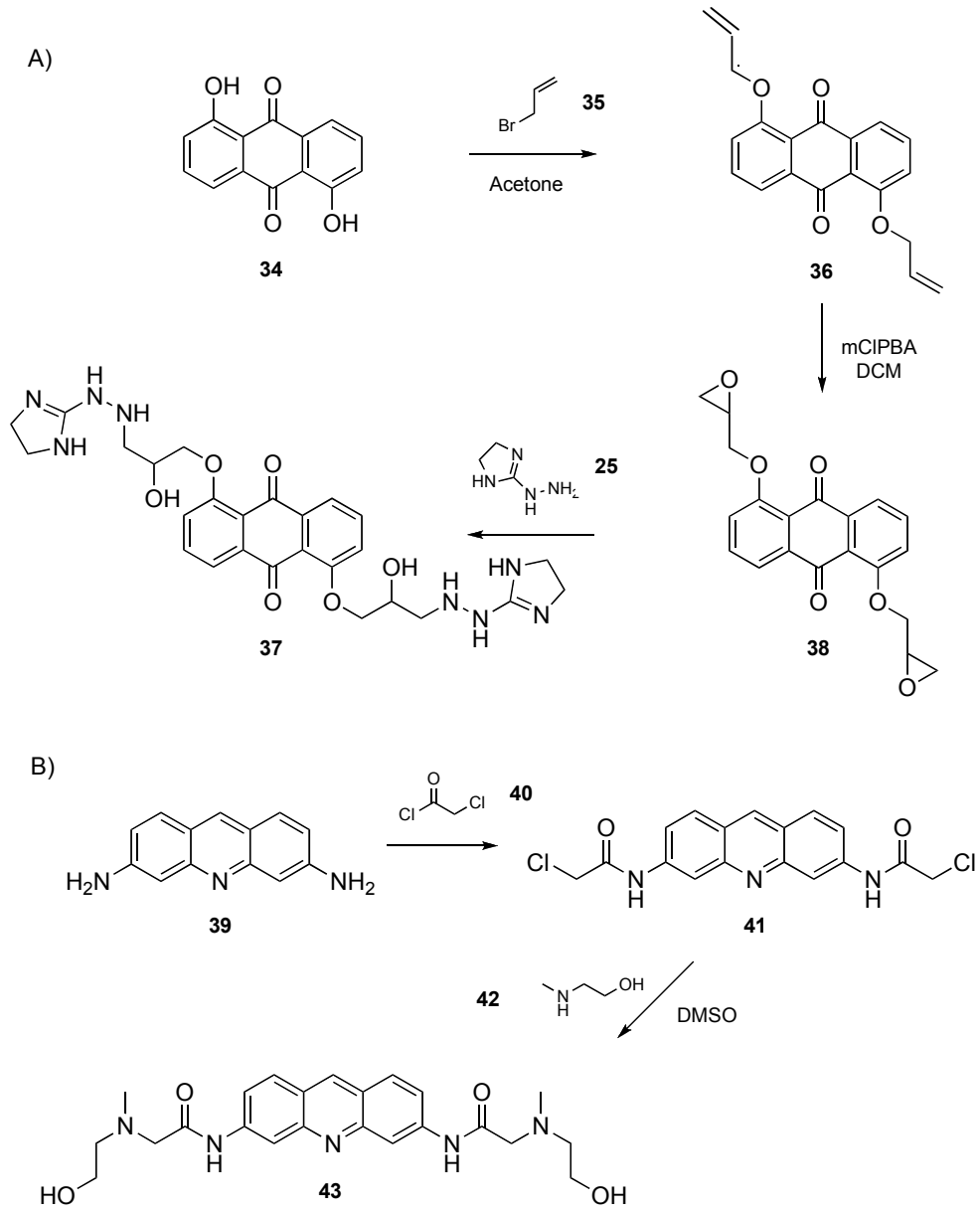
Even in this case, as described above, the mechanism seems to be debatable. In fact, according to the proposed structure here reported, the mechanism seems in open contrast with the one discussed for the reaction with hydrazine, where a first substitution occurs to the chlorine atom on the aromatic ring.

The use of pyridine as a solvent, probably, influences somehow the reactivity of one or both the reactants, driving the reaction firstly to the imine formation and then to the nucleophilic substitution. The attack of the "guanidinic" nitrogen on the C-Cl as a first step appears less probable because of the lower nucleophilic level due to the electronic delocalization. On the other hand, a loss in aromaticity seems unlike to happen and probably the side chains are directed outside the scaffold (as drawn).

The reaction did not lead to the desired compound when conduced in the firstly optimized

condition (DMA/microwave as reported above). Anyway, the structure of the compound will be further investigated by 2D NMR analysis and also side products will be considered with the use of high-resolution mass spectrometry.

2.4 SCHEME IV

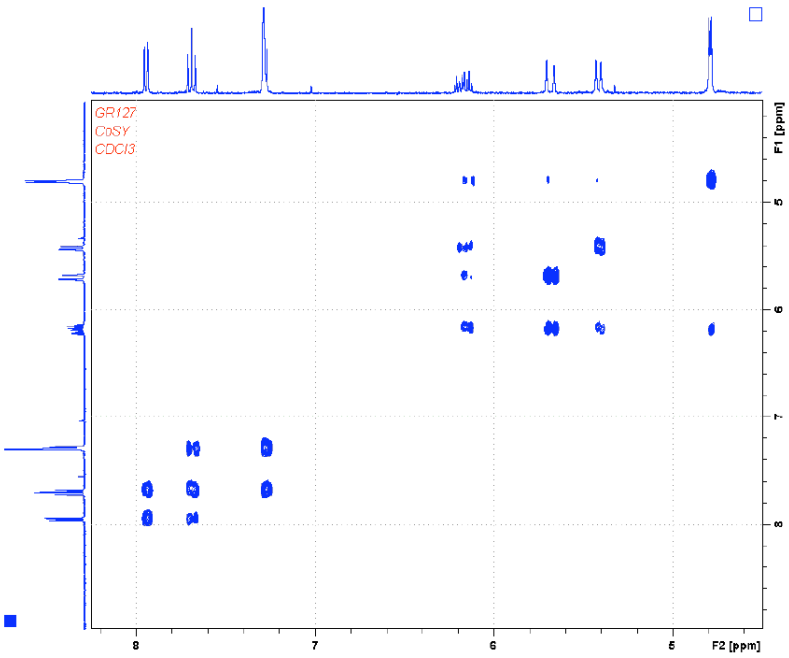

A)
OH
O
O
OH
34
Br
35
Acetone
O
O
O
O
36
mClPBA
DCM
O
O
O
O
38
HN
N
N-NH2
H
25
OH
H
N
N
HN
N
H
O
O
O
OH
N
NH
NH
37
B)
Cl
O
Cl
40
H2N
N
NH2
39
Cl
O
N
H
N
O
N
H
Cl
41
42
N
H
OH
DMSO
HO
N
O
N
H
N
N
H
O
N
OH
43

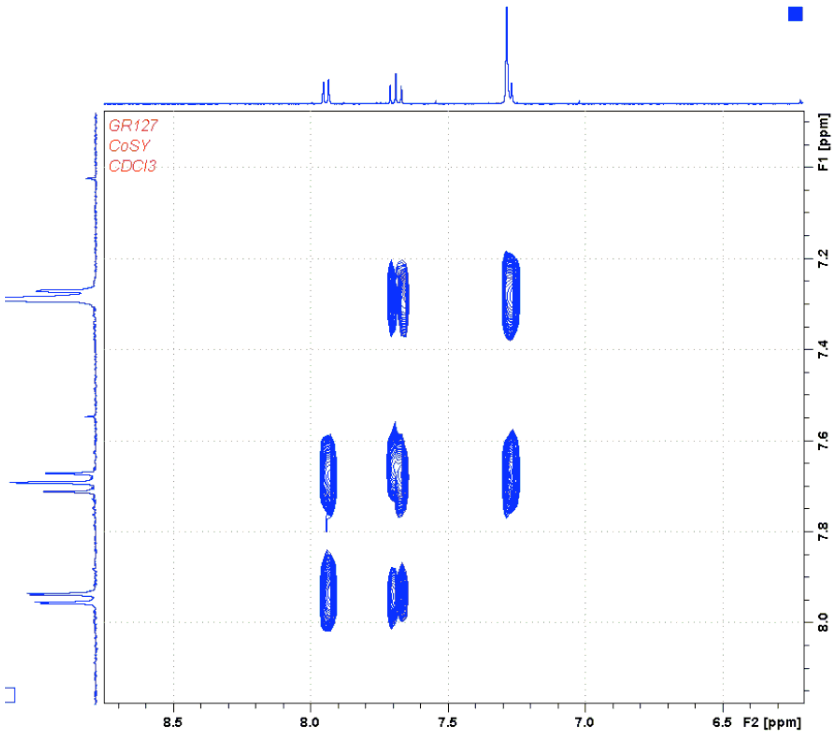
Rationale of the synthetic scheme

To expand the set of synthesized compounds and explore novel synthetic route, two more series of compounds were designed. The relevance of this synthetic scheme goes beyond the compound that were effectively obtained, as long as it opens two more paths for the preparation of potential G-quadruplex stabilizers. In particular anthraquinone and proflavine were chosen as scaffolds for these new derivatives.

Anthraquinone derivatives

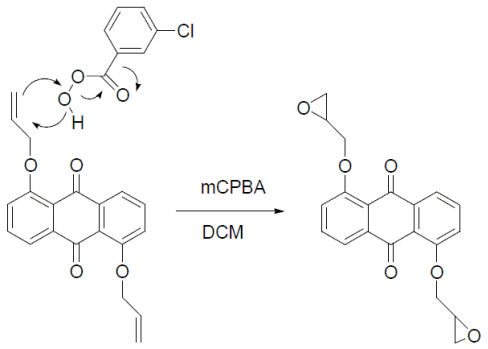
The first step of the synthetic scheme consists in the derivatization of the hydroxy anthraquinone scaffold with a linking chain. In order to obtain a "flexible" and convenient intermediate, suitable to be derivatized with different moieties, the allyl group was chosen as substituent. The reaction was carried out in acetone and allyl bromide was used as a substrate for the nucleophilic substitution. The intermediate obtained was characterized by 2D NMR analysis.





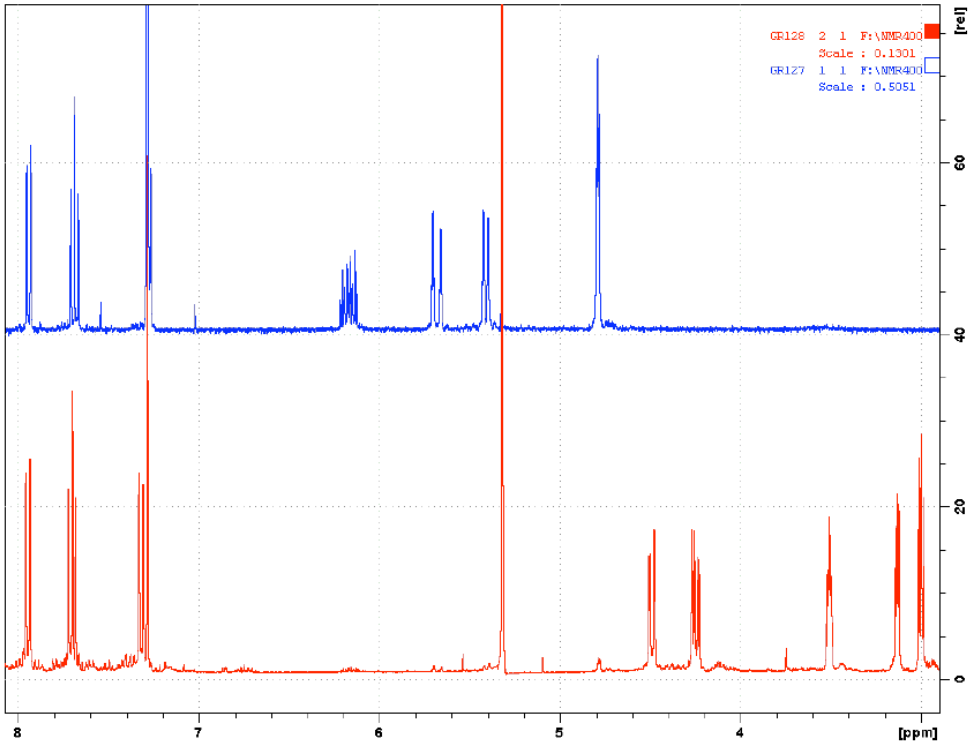
The reported CoSY spectra enlighten the correlation between the protons belonging to the aliphatic portion of the molecule and the ones coming from the aromatic system.

The following step is represented by an epoxidation reaction that was carried out using m-chloroperbenzoic acid.



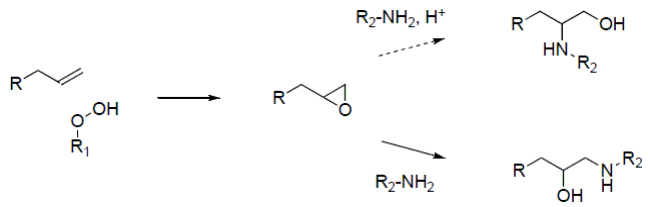
36 **38**

Even in this case the product was studied by NMR and a comparison between the proton spectra of the reactant and the product clearly indicated a full conversion to the desired compound. This reaction gives a double epoxide as a product.



The comparison between the two spectra shows how the signals due to the allylic system completely disappear leaving space to the signals compatible with the epoxyde ring. It is also important to notice that the symmetry of the molecule is maintained: both the functional groups (position 1 and 5 of the anthraquinone scaffold) react the same way.

The last step of this synthetic procedure consists in the derivatization of the scaffold obtained with the previous steps. The procedure developed is suitable for any good nucleophile (e.g. primary or secondary amine) but, coherently with the other derivatives prepared during this research work, the hydrazo imidazolinic side chain was introduced (the same one of the lead compound).



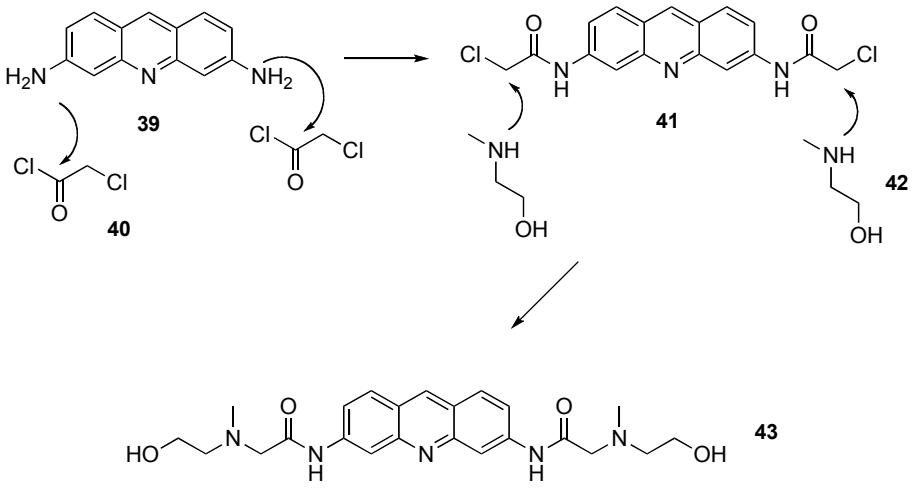
Actually the reactivity of epoxydes is really peculiar and it is possible to drive the reactivity

of the nucleophile involved toward one or the other carbon atom forming the ring, according to the use or the avoidance of an acid during the reaction. Experiments about the conditions of this step are still being carried out.

Proflavine derivatives

Even in this case the relevance of the synthetic scheme consists in the set up of the conditions, described in the experimental section, for the obtainment of an intermediate that is easy to modify and ready for further derivatizations.

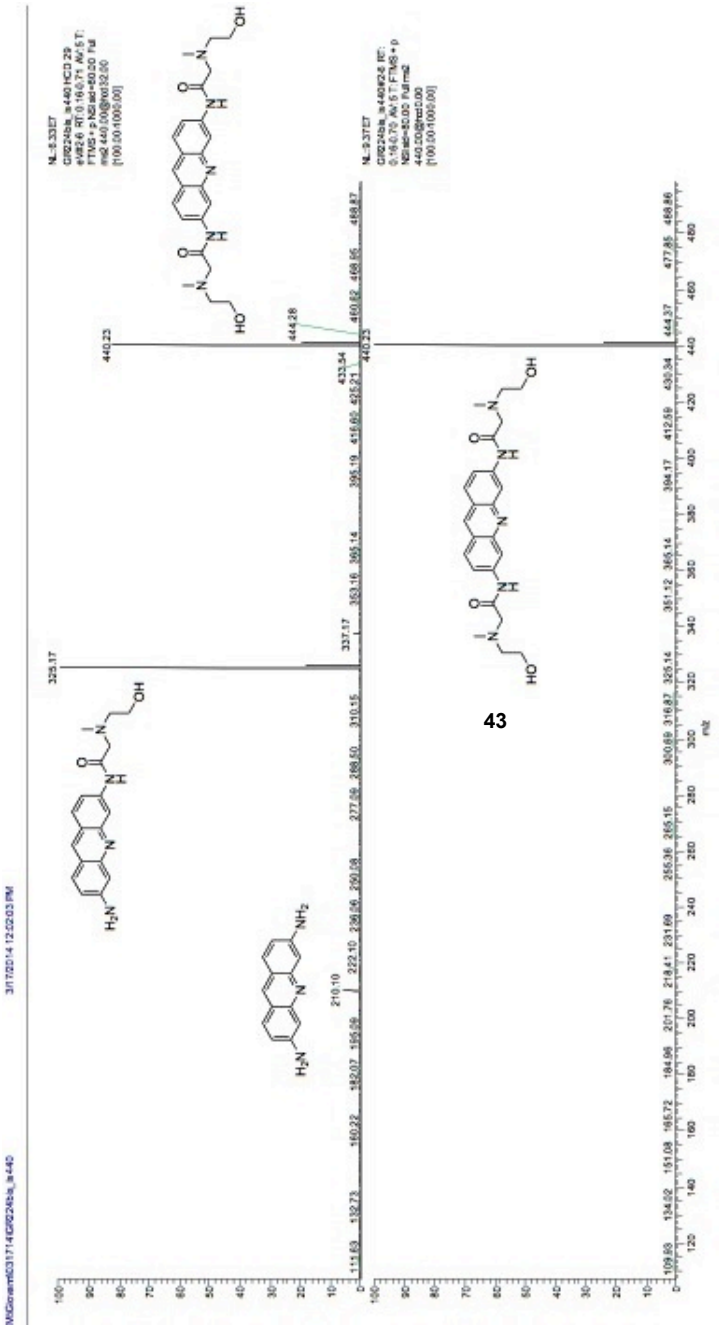
In particular, proflavine was reacted with chloroacetyl chloride, a reactant similar to the ones described in other reaction schemes of these research work with a completely different use. As the mechanism reported below clearly shows, one of the two chlorine atoms of chloroacetyl chloride shows a much higher reactivity toward nucleophilic substitution due to the proximity to the carbonyl group. This reactant, in fact, represents the activated form of chloroacetic acid.



In the case of the here reported compound, the intermediate described above was reacted with a secondary amine in mild conditions to obtain a final compound with protonable and flexible side chains. N-methyl aminoethanol was used but experiments are being carried

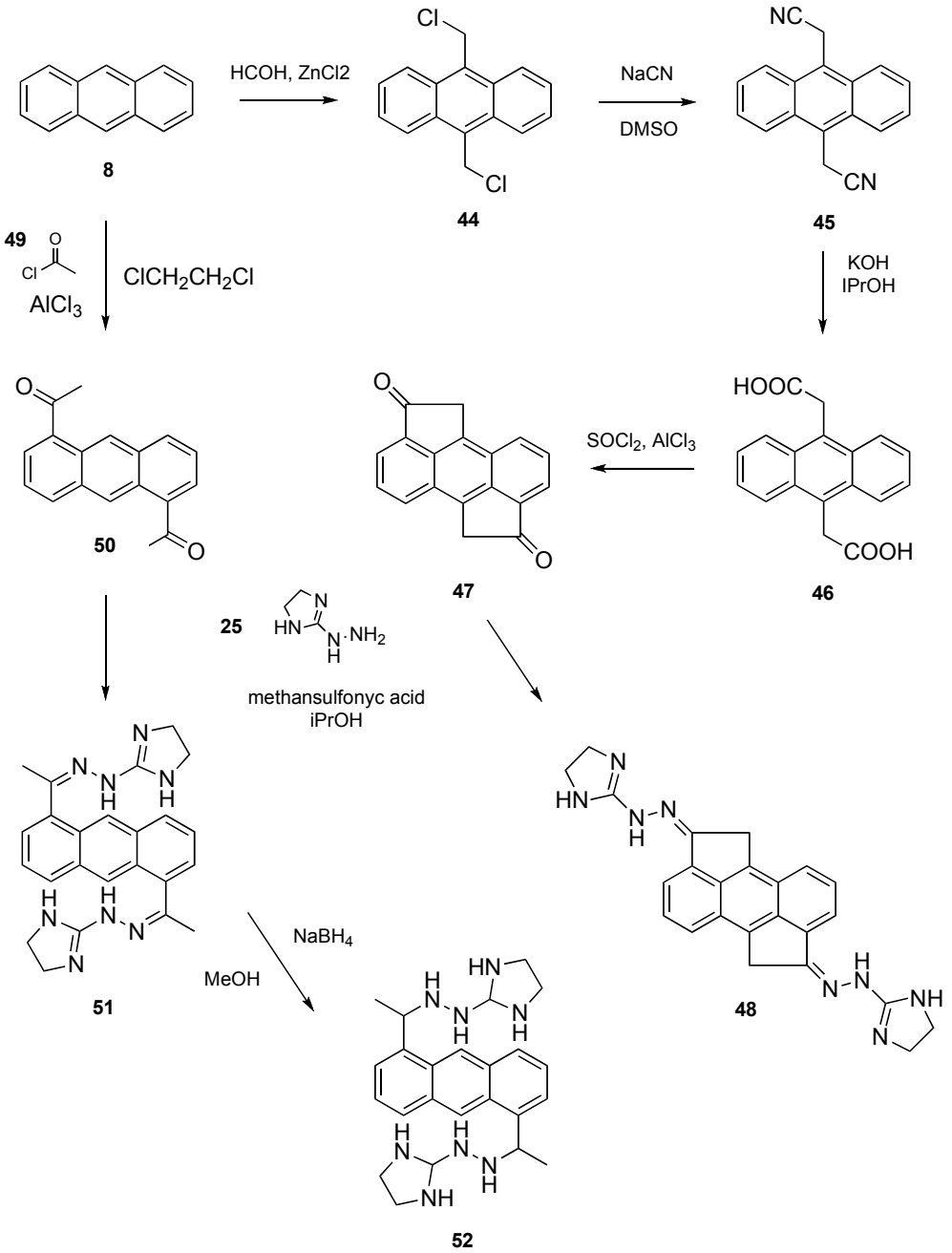
out with different primary and secondary amines. This compound was also used as an intermediate for further derivatization as it will be described in the following synthetic schemes.

Characterization of the compounds



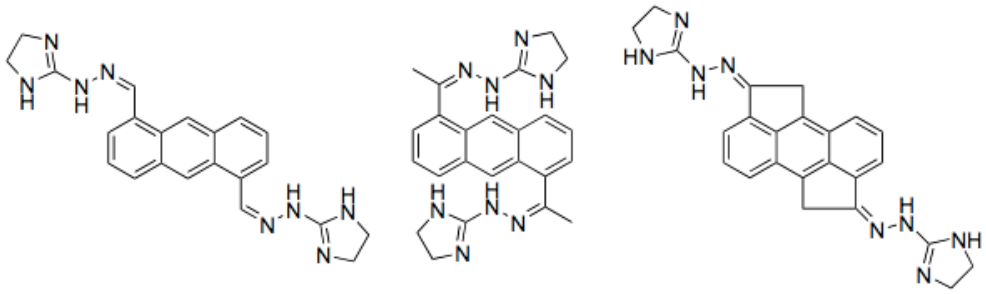
43

2.5 SCHEME V


HCOH, ZnCl2
NaCN
DMSO
KOH
IPrOH
SOCl2, AlCl3
ClCH2CH2Cl
AlCl3
methansulfonyc acid
iPrOH
NaBH4
MeOH
8
44
45
46
47
48
49
50
51
52
25

Rationale of the synthetic scheme

In this synthetic strategy the attention was re-focused on the lead compound giving particular attention to the role of the structural properties of the side chains. To be more specific, constraining groups were introduced to force the spatial orientation of that part of the molecule, very flexible in the lead compound. The introduction in the benzylic position of a methyl or methylene moiety through two novel synthetic schemes led to two more rigid compounds with an opposite constrained orientation of the side chains.



7 **51** **48**

In silico conformational analysis

An *in silico* conformational search was carried out on the three compounds using Avogadro [5,6]. The most stable conformations detected are reported below.[96]

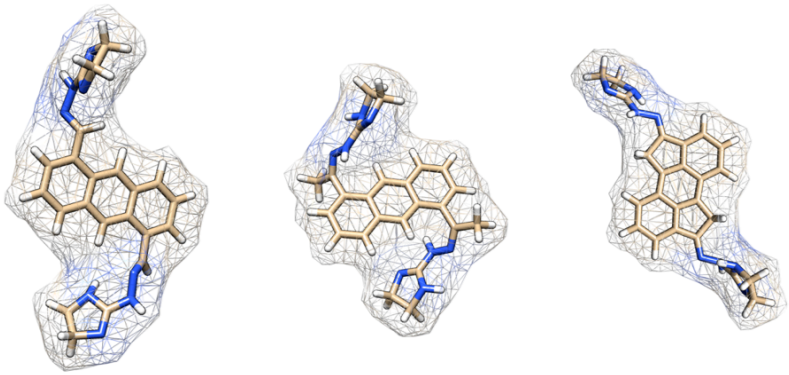


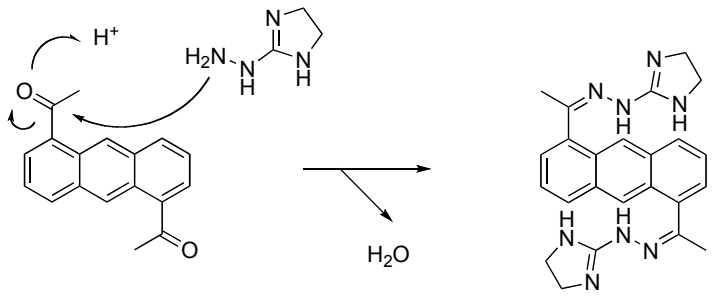
FIG30. Minimized conformation of the cited compounds, MMFF94[96]

According to the hypothesis that the side chain conformation could be responsible for

disposing the aromatic scaffold in a strategic position for an efficient interaction with the quadruplex, to fully dissect the issue of side-chain flexibility, we prepared structurally constrained derivatives of the most effective bisantrene compound exhibiting rigid links and/or lack free rotation with reference to the planar portion and examined their efficiency in stabilizing G-quadruplex. In these compounds the same side chain, connected to positions 1 and 5 of the anthracene nucleus, lays coplanarly to the aromatic system, but only can dispose alternatively inwards (compound **51**) or outwards (compound **48**). In addition to this, the least perturbing constraining residues (methyl or methylene) were used in order to minimize variability and enlighten the role of a mutual adaptation between the nucleic acid and the binder. These predictions were then confirmed by 2D-NMR analysis as reported in the following paragraph.

Constrained derivative 1

The desired compound was directly obtained from 1,5-diacetyl anthracene through imine formation by reacting with 4,5-dihydro-1H-imidazol-2-yl-hydrazine in isopropanol in presence of methansulfonic acid.



50 **51**

The constraining effect was evaluated thanks to a 2D NOESY NMR experiment, showing the proximity between the aromatic proton H_A and the methyl group.

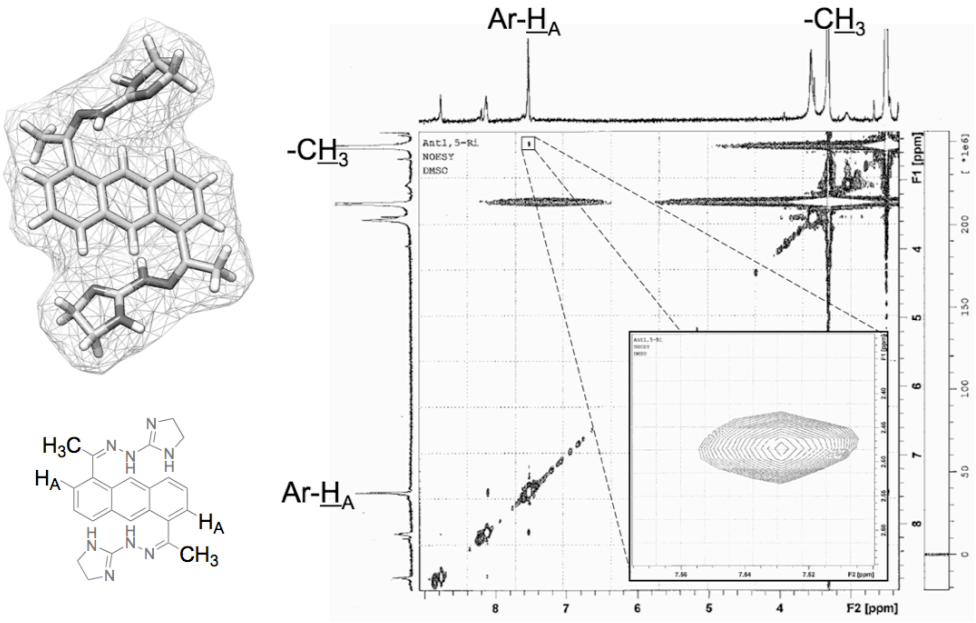


FIG32a. Comparison between predicted (3D model) and observed (NOESY NMR) conformational properties[96]

This spectroscopic piece of information is particularly relevant especially if compared with the results of a series of NOE experiments that were previously conduced on bisantrene-like hydrazones with exactly the same side chain of the 1,5 di substituted lead compound. In that case the NMR experiments confirmed "the possibility of free rotation of the side chain groups with reference to the planar anthracene moiety".[75] A minimized model for compound **51** (force field MMFF94[96]) here proposed in its three dimensional arrangement is consistent with the information given by the NOESY NMR experiment. In the 2D spectrum, as expected, other cross peaks are present confirming the relative proximity between the protons belonging to the aromatic system.

NMR-based structural analysis of derivative **51** was then carried to a further level: a quantitative 2D NOESY (Nuclear Overhauser Effect Spectroscopy) experiment was developed to predict the distance of the protons belonging to the molecule in the prevalent three dimensional conformation in solution. The NOESY NMR experiment provides a visualization as cross-peaks in a 2D surface of the spatial interactions between protons that are close one to the other in the three dimensional arrangement in solution. According to what literature reports, the intensity of the cross peak measured on the map is proportional

to 1/[distance between protons]^{6}.[76] Basing on this, once that a know distance (giving a NOE cross-peak) is set as reference, it is possible to estimate the distance between two protons showing a signal with a detectable intensity in the 2D spectrum using the equation:

$$r_{ij} = r_{ref} \bullet \sqrt[6]{(I_{ref}/I_{ij})}$$

where r are the distances and I the intensities of the NOESY cross-peaks.[76]

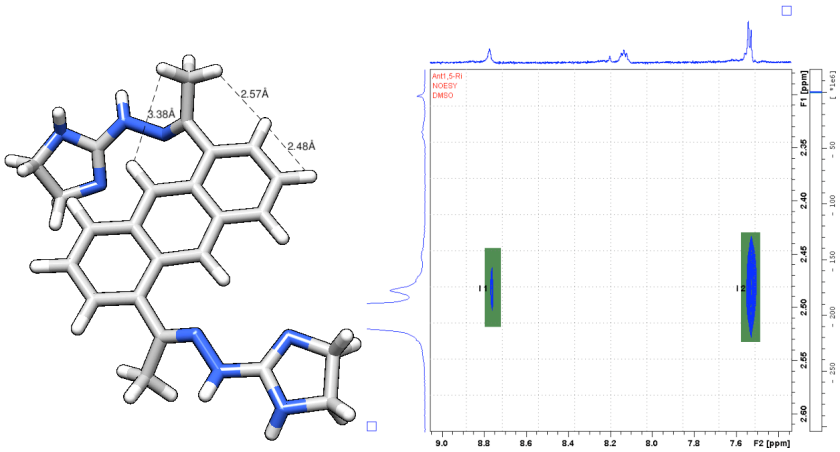
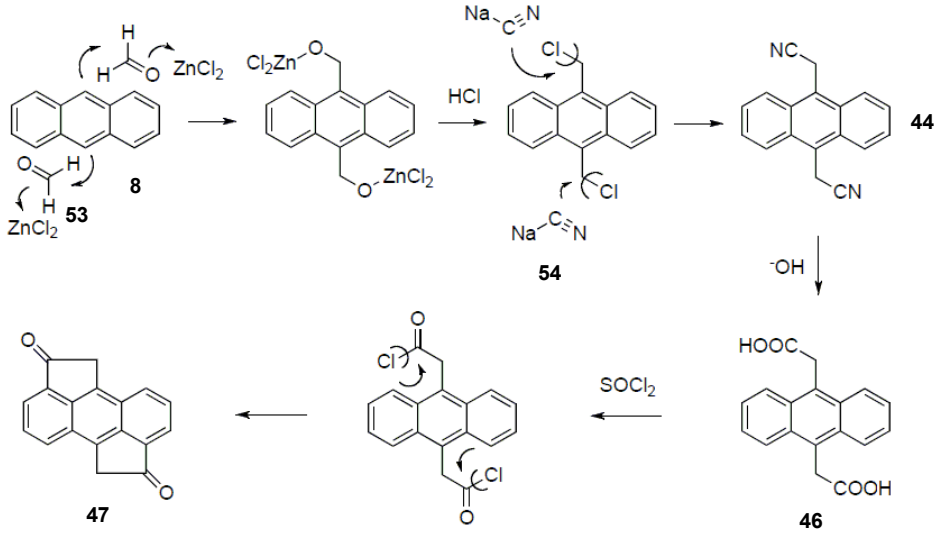


FIG32b. Predicted distances between protons in the minimized model of compound **51**[96] and detailed view of the two cross peaks due to the interaction of the methyl group with the aromatic protons in positions 2 and 9.

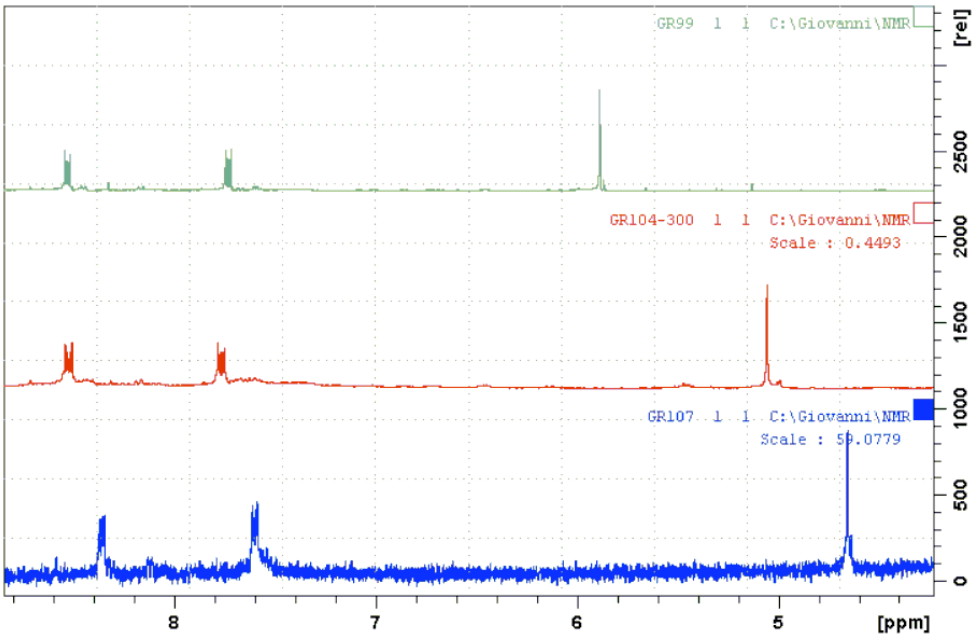
Figure 32b shows the minimized predicted conformation for compound **51** calculated in the previously cited force field. The ArH-ArH distance of 2.48 Å was chosen as reference and, basing on the formula reported above, the calculated distances (2.55 and 3.26 Å) were found to be in excellent agreement with the values measured and shown on the model (2.57 and 3.38 Å respectively).

Constrained derivative 2

For what concerning compound **48**, some of the steps were previously reported by Ryu[77] and Mohebbi.[78]



The chloromethylation of anthracene was performed with $ZnCl_2$ and paraformaldehyde in dioxane in the presence of conc. HCl. The following step leads to the obtainment of a cyano compound by using NaCN in DMSO or in water/dichloromethane with phase-transfer catalysis. The hydrolysis of the cyano residue and the subsequent formation of the propyl esther was conducted wit p-toluensulfonic acid in n-propanol. The hydrolysis of the ester with LiOH in THF/water gave the carboxylic acid, which was then activated with thionyl chloride for the cyclization performed in dichloroethan in the presence of $AlCl_3$. The imine formation with 4,5-dihydro-1H-imidazol-2-yl-hydrazine in isopropanol in presence of methanesulfonic acid gave the final product.



The comparison between the proton NMR spectra above shows the differences, in terms of chemical shifts, between the protons belonging to the different intermediates involved in the synthetic process. In particular, while the aromatic portion does not show great changes, the chemical shift of the benzylic CH2 signal decreases moving from compound **44** through **45** to **46**.

Docking studies

A docking experiment should predict, or at least suggest, the preferential conformation that a ligand assumes while interacting with a macromolecule (the docking site). This computational experiment is growing in appeal over the years as long as it could provide useful suggestions to the synthetic medicinal chemist that, basing on the modeling results and observing the predicted conformations and contact areas, can discuss on how to modify or develope a scaffold to optimize the interaction. While this technique found a wide application in the field of ligand-protein interaction prediction, only few examples of docking to nucleic acids were reported until recently,[79] due to the fact that super molecular

arrangement of DNA (and RNA) are known to be very flexible and transient, while a docking study treats the macromolecule as an at least partially rigid structure.

For this study, a telomeric G-quadruplex sequence (PDB ID: 3T5E) solved in a K+ by X-ray was used as the reference macromolecule for the docking analysis.[80] The docking procedure was designed, developed and carried out in accordance with other recently reported ligand-G-quadruplex interaction studies.[79]

A possible pose of the molecules stacked on a telomeric G-quadruplex was also investigated *in silico* thanks to the software suites cited in the experimental section.

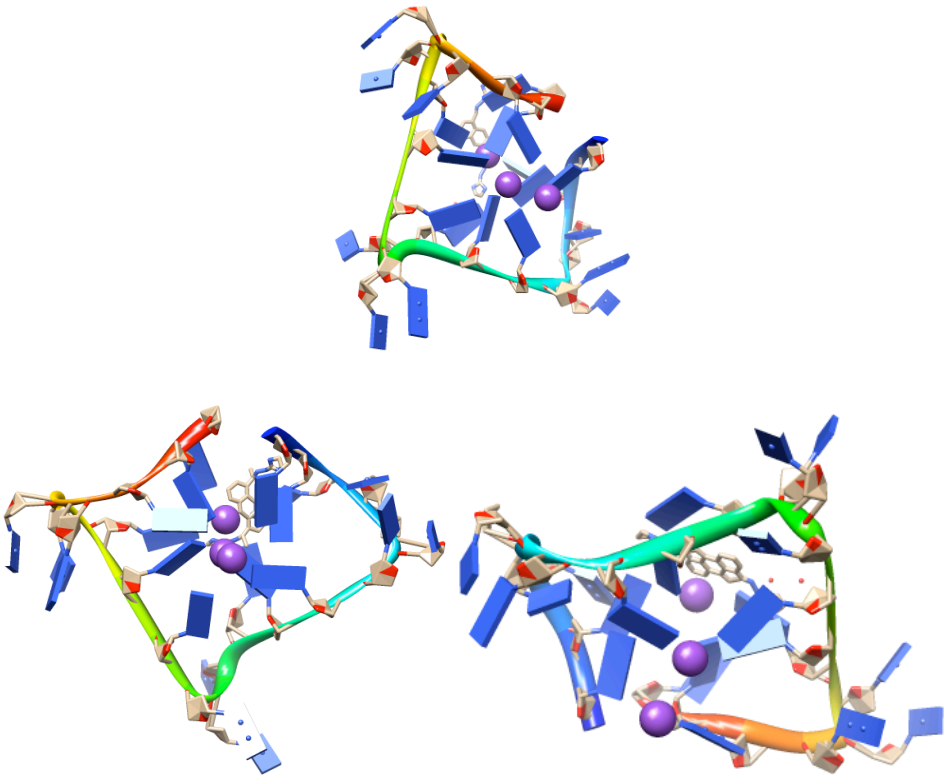


FIG31a. Representation of the interaction between the compounds and a telomeric G-quadruplex

For this study, a telomeric G-quadruplex sequence (PDB ID: 3T5E) solved in a K+ by X-ray was used as the reference macromolecule for the docking analysis.[81] The docking procedure was designed, developed and carried out in accordance with other recently reported ligand-G-quadruplex interaction studies.[79]

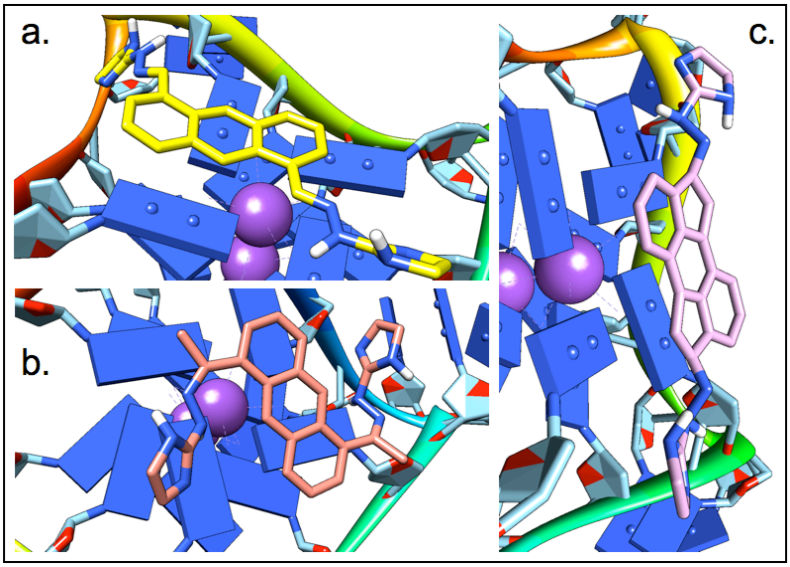


FIG31b. Docking study on the synthesized compounds

Compounds **7** (a), **48** (c) and **51** (b) were docked as previously described to a telomeric G-quadruplex structure. According to these models, the three compounds share the π-π stacking with an external quartet as preferential binding motif. Curiously, compound **51** (b) maintains the same conformation predicted in a previous *in silico* conformational analysis and confirmed by 2D-NOESY NMR in solution.

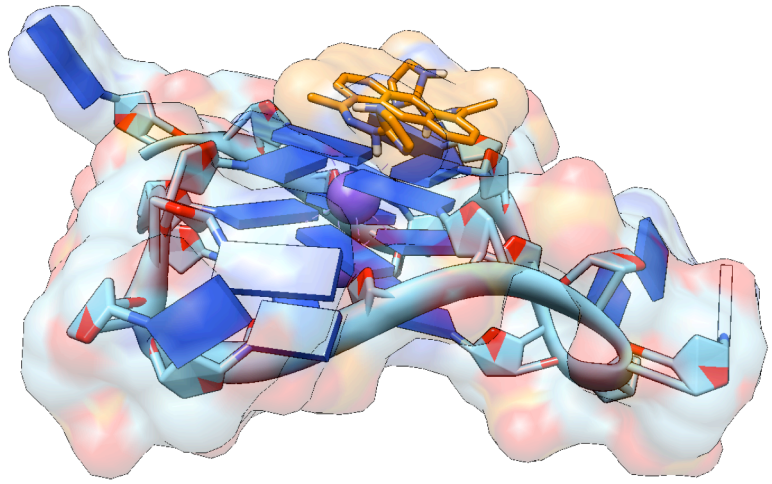
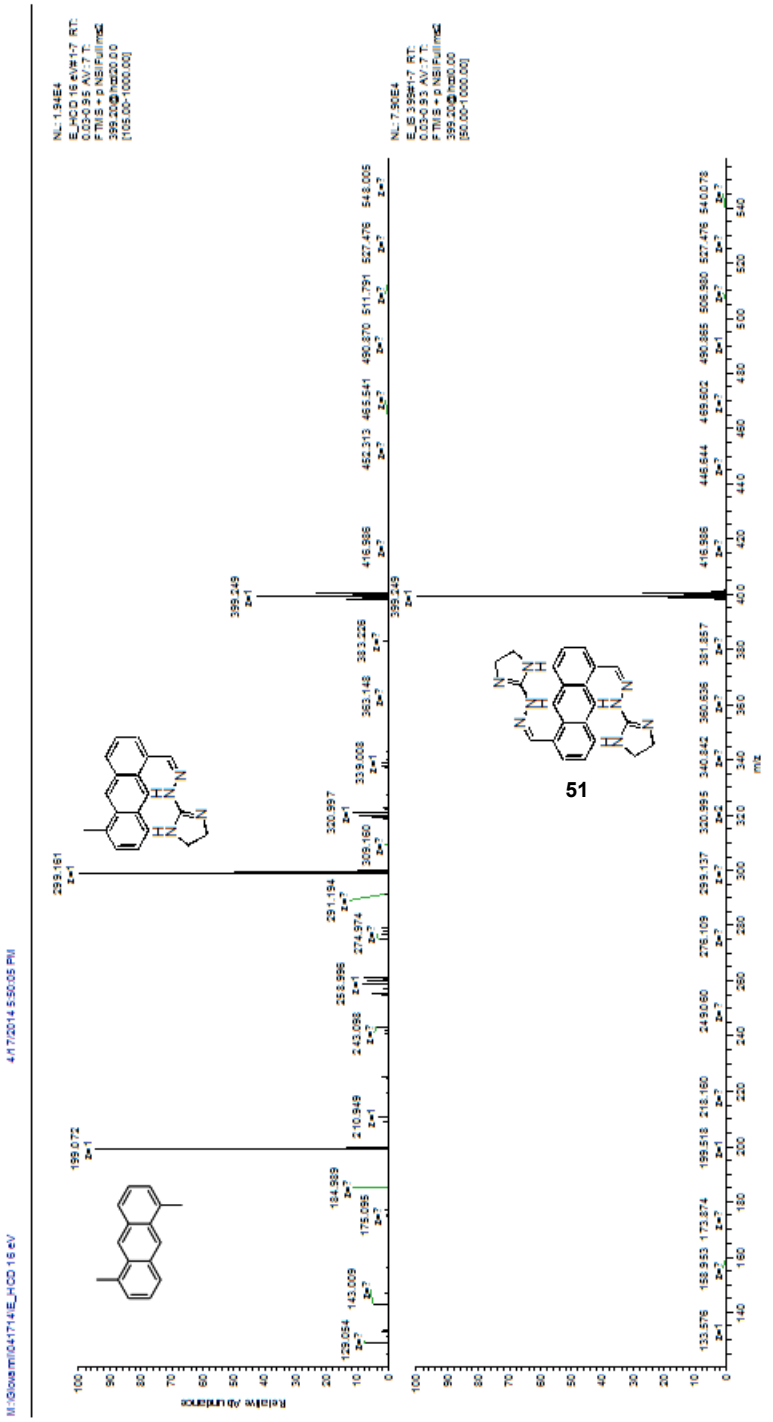
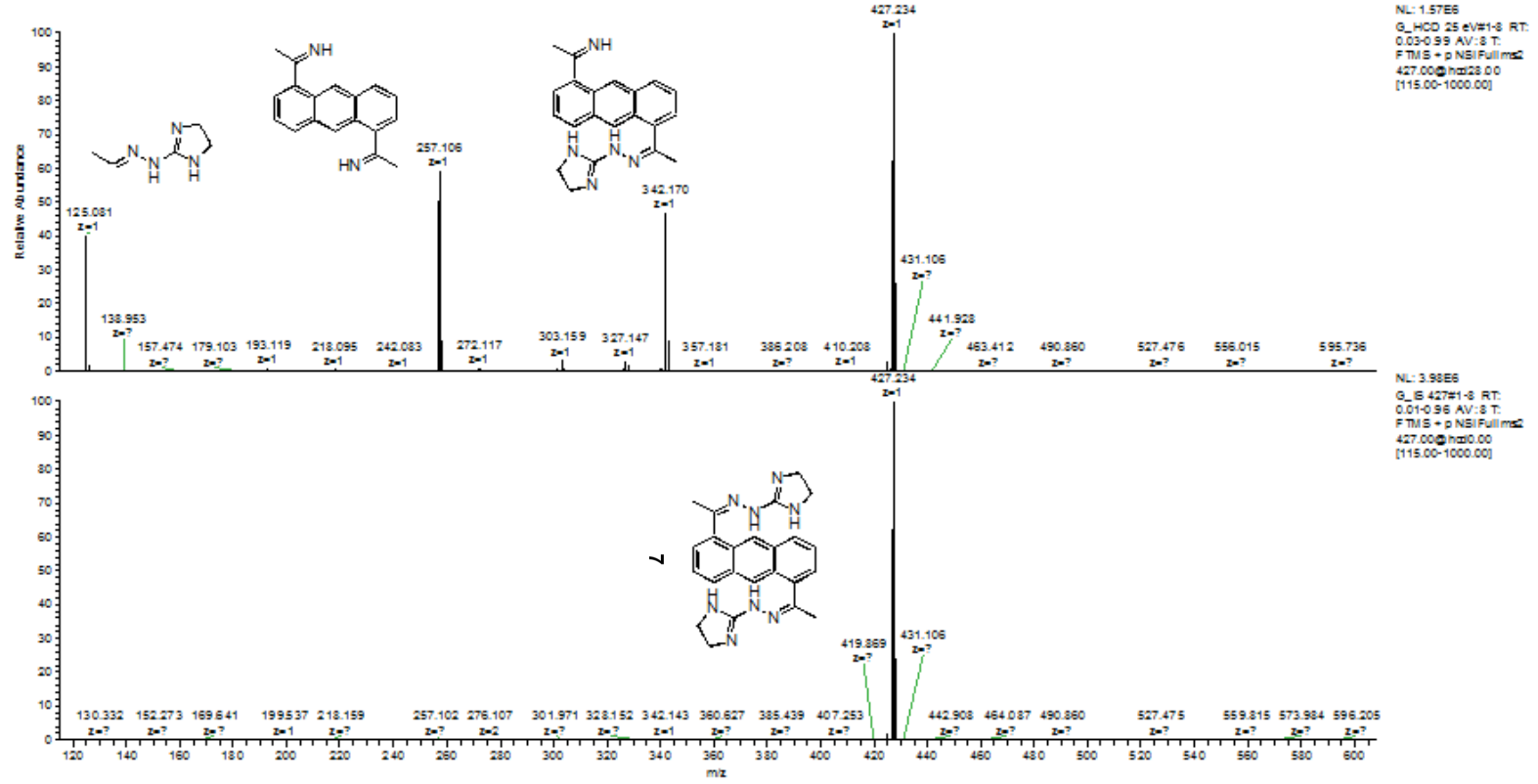


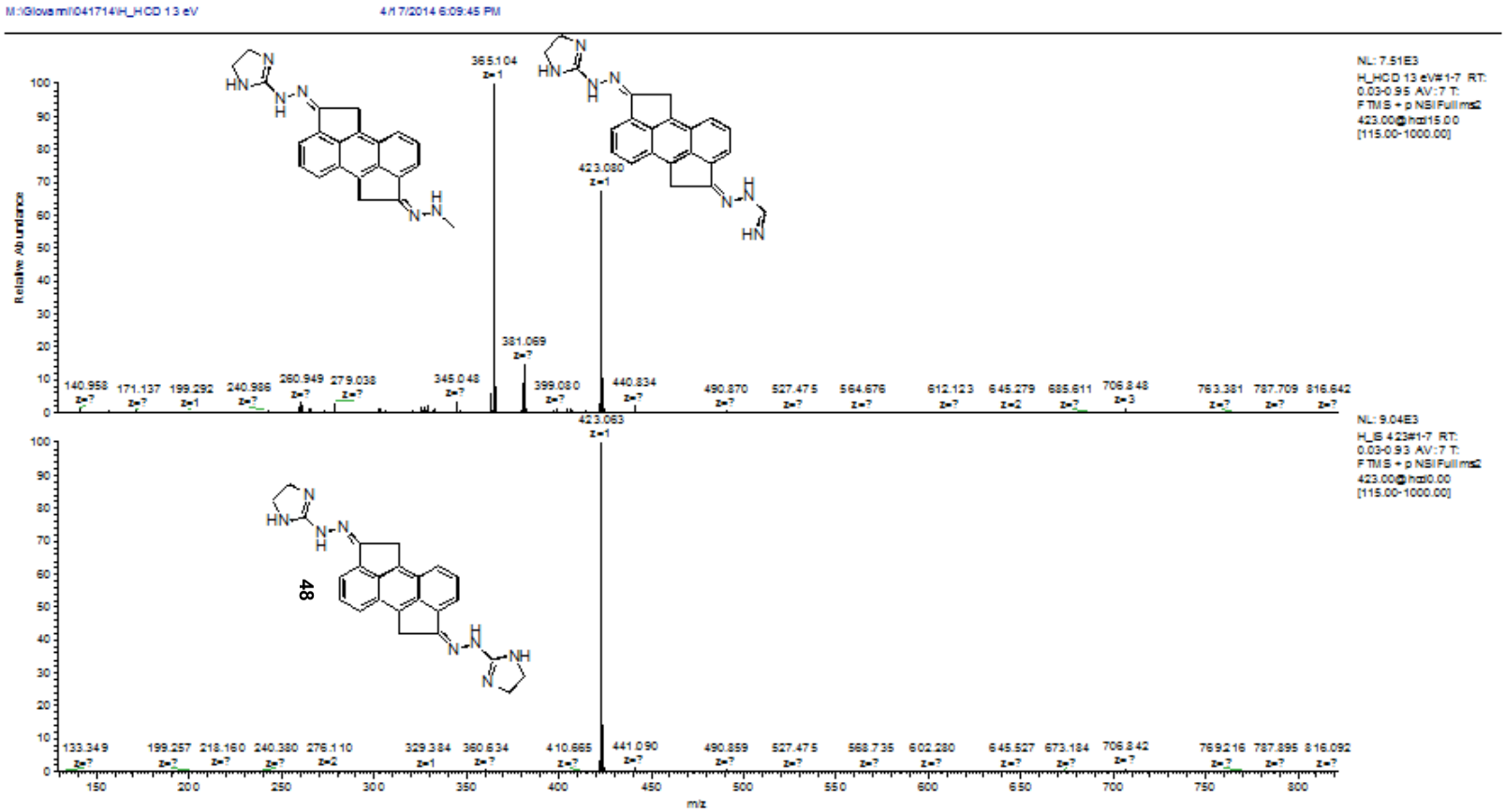
FIG31c. General view of compound **51** docked in the quadruplex.

The docking routine procedure is described in the experimental section.

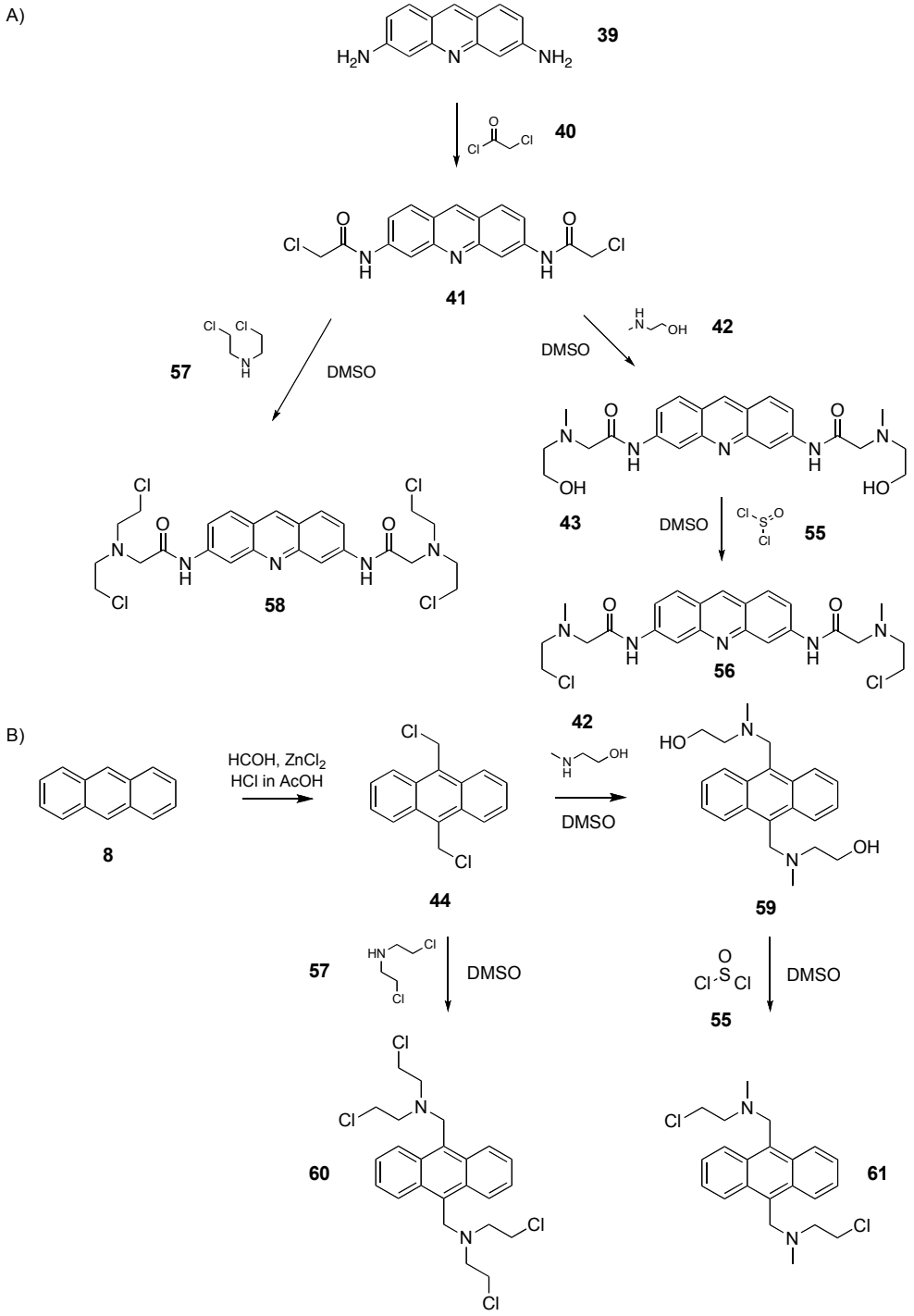
Characterization of the compounds




M:\Glowarm\041714\G_IS 427
4/17/2014 5:56:22 PM
NL: 1.57E6
G_HCD 25 eV#1-8 RT: 0.03-0.99 AV: 8 T: FTMS + p NSI Full ms2 427.00@hcd28.00 [115.00-1000.00]
NL: 3.98E6
G_IS 427#1-8 RT: 0.01-0.96 AV: 8 T: FTMS + p NSI Full ms2 427.00@hcd0.00 [115.00-1000.00]
Relative Abundance
m/z
427.234
257.106
342.170
125.081
431.106
441.928
419.869
7

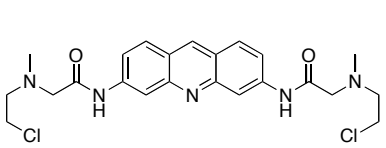

M:\Giovanni\041714\H_HCD 13 eV
4/17/2014 6:09:45 PM
NL: 7.51E3
H_HCD 13 eV#1-7 RT:
0.03-0.95 AV: 7 T:
FTMS + p NSI Full ms2
423.00@hcd15.00
[115.00-1000.00]
365.104
423.080
381.069
423.063
NL: 9.04E3
H_IS 423#1-7 RT:
0.03-0.93 AV: 7 T:
FTMS + p NSI Full ms2
423.00@hcd0.00
[115.00-1000.00]
48
Relative Abundance
m/z

2.6 SCHEME VI

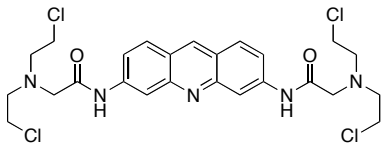

A)
39
40
41
42
DMSO
57
DMSO
43
DMSO
55
58
56
B)
HCOH, ZnCl2
HCl in AcOH
8
44
42
DMSO
59
57
DMSO
DMSO
55
60
61

Rationale of the synthetic scheme

This synthetic scheme was designed in order to obtain some compound with a peculiar feature: the combination of well known scaffolds (anthracene and proflavine) known to give interaction with the quadruplex and of a couple of alkylating function to interact strongly with the bases in a covalent fashion. As an example, the combination of the G-quadruplex stabilizing scaffold and the nitrogen mustard alkylating moiety is a recent trend and was only very recently investigated.[82] A study from Di Antonio et al showed the influence of directioning the alkylating reactivity towards a G-quadruplex structure. It turned out that combining a scaffold coming from a potent G-quadruplex stabilizing agent with the nitrogen mustard moiety led to a selective alkylation of the quadruplex. In this connection, some novel nitrogen mustard derivatives were designed.

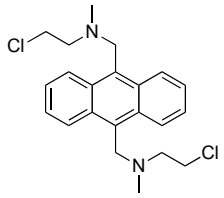


56

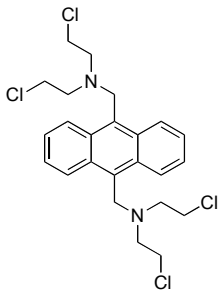


58

In particular, two derivatives show the proflavine scaffold and two are anthracene-based.



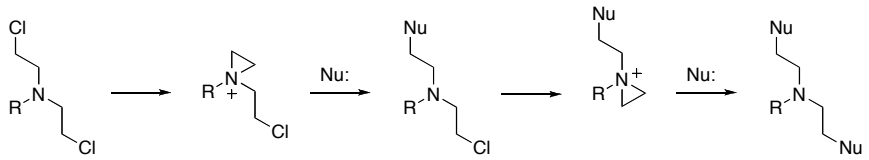
61



60

As can be noticed from the reported structures, for both the classes of compound, monofunctional and bifunctional moieties were synthesized. The compounds with only an alkylating function per side were actually prepared in order to achieve a higher selectivity reducing bias, side-reactivity and the overall reactivity of the molecule.

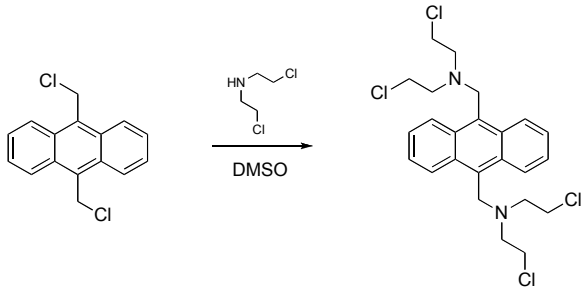
Nitrogen mustards, in fact, show a cross-linking reactivity. Classical bifunctional mustard can be reactive towards two nucleophilic residues going through a cyclic intermediate.



Otherwise, if only a chloroethylamine moiety is present, the reaction stops at the monoalkylation stage.

Reactivity of N,N dichloroethylamine

The reaction between 9,10 dichloromethyl anthracene and dichloroethylamine must be handled with some precautions and in mild conditions in order to obtain the final product and avoid side reaction, and in particular the self-reactivity of dichloroethylamine.



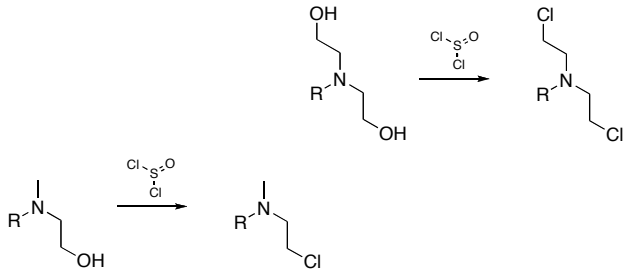
44 **60**

Following a procedure already reported by Yin et al,[83] the nucleophilic substitution was carried out checking temperature and product formation through MS analysis (see experimental section for further details).

Activation of the nitrogen mustard

Most of the reported synthetic procedures for the preparation of nitrogen mustards, when

they do not go through the direct nucleophilic substitution with dichloroethylamine described above,[82] present a hydroxyl intermediate.



The final step of the synthesis is then the substitution of the hydroxyl group with a chlorine atom through a reaction with thionyl chloride.[84]

2.7 EVALUATION OF G-QUADRUPLEXES AND THEIR INTERACTIONS

Experiments carried out at State University of New York

In the following sections are described some experiments aimed to evaluate the formation of G-quadruplexes from guanine rich sequences in solutions and the binding affinity of some synthesized ligands to structured DNA. The gel electrophoresis and mass spectrometry experiments were carried out during the third year of the PhD school in the laboratory of Prof. Fabris at UAlbany, State University of New York, USA.

G-quadruplex forming sequences

To perform the investigation of both the presence of G-quadruplex sequences in solution and of the compound-DNA binding capability two guanine rich DNA sequences were chosen as substrates. To be more specific both these sequence show the peculiar TTAGGG repeat already described to be peculiar of telomeres (see the introduction for further details). In addition to this, a strand as similar as possible to the one used for the fluorescence melting assays was considered.

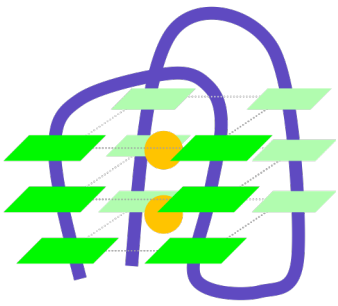


FIG33. Monomeric G-quadruplex arrangement.

A monomeric, self-folding DNA sequence was analyzed (here called GQm): 5' - AGG

GTT AGG GTT AGG GTT AGG GT - 3'. This strand, in an opportune buffer where cations are present should be able to self-fold into a G-quadruplex structure. The expected structure would be formed by three stacked G-quartets with two ions between the layers acting as stabilizers in the middle of the structure. In fact, the sequence shows four repeats of the GGG motif. The other bases are supposed to form the loops of the G-quadruplex structure. As previously introduced, this sequence is the one that was used for the screening of the stabilizing performances of the synthesized compounds through the fluorescence melting assay reported in this research work and in the literature.[4]

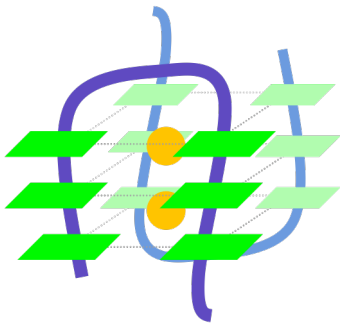


FIG34. Dimeric G-quadruplex arrangement.

To widen the range of the investigation, another, shorter guanine rich sequence was considered (here called GQd): 5' - TAG GGT TAG GGT - 3'. As long as this sequence shows only two repeats of the GGG motif, two of these strands should combine in solution to form a G-quadruplex made by three stacked quartets. Also this sequence shows the peculiar TTAGGG telomeric repeat and it has already been reported to be a suitable substrate for mass spectrometry investigation of G-quadruplexes.[90,92]

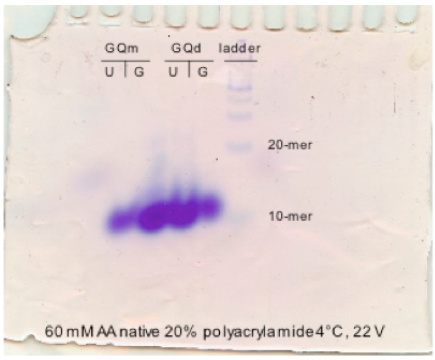
In addition to this, these strands were chosen also because, besides reproducing the telomeric repeat, many structures of drug-DNA complexes were solved by NMR or X-ray analysis of these exact sequences.[85]

2.8 GEL ELECTROPHORESIS

Gel electrophoresis to enlighten G-quadruplexes

This part of the research work was carried out in the laboratory of Prof. Fabris, UAlbany, State University of New York, USA. The capability of guanine rich sequences of forming G-quadruplex structures in solution was evaluated firstly through electrophoresis analysis. Some procedures for the investigation of G-quadruplexes through this technique have already been reported.[4] The rationale on which this analysis bases is represented by the fact that an electrophoresis gel should separate DNA sequences according to their molecular weight (a direct reflection of the strain length in terms of number of bases) but also to their shape. In this connection, a DNA sequence arranged in the G-quadruplex structure would be running more in a gel than what expected considering only its molecular weight.

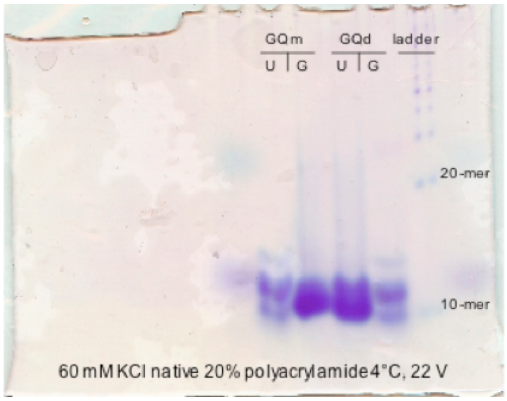
As long as other experiments reported in this research work involved G-quadruplex forming sequences incubated in ammonium acetate (see experimental section for details), a first attempt was performed using ammonium acetate in the running buffer in order to harmonize all the experiments. A native (non denaturing) 20% polyacrylamide gel was run under the described conditions.



The sequence used are GQm (5' - AGG GTT AGG GTT AGG GTT AGG GT - 3') and GQd (5' - TAG GGT TAG GGT - 3'). The gel showed how all the sequences seemed to run

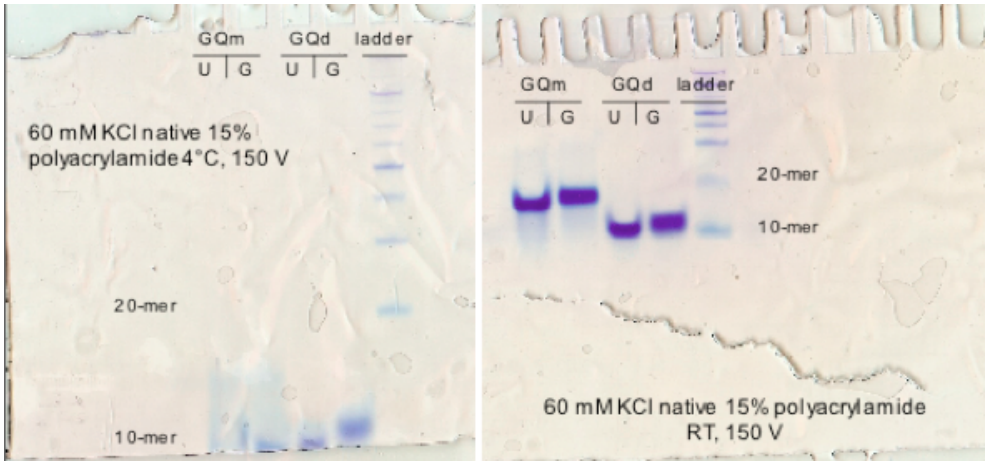
the same way, even if the molecular weight is different. A reference ladder was used as a control. In addition to that it is evident how the pre-denatured samples, to which 14 M urea was added before the run (U) were running the same as non-denatured (G). This could suggest a quick refolding o the structure.

Given the non optimal quality of the gel obtained with ammonium acetate as running buffer, the same experiment was repeated with potassium chloride as running buffer.



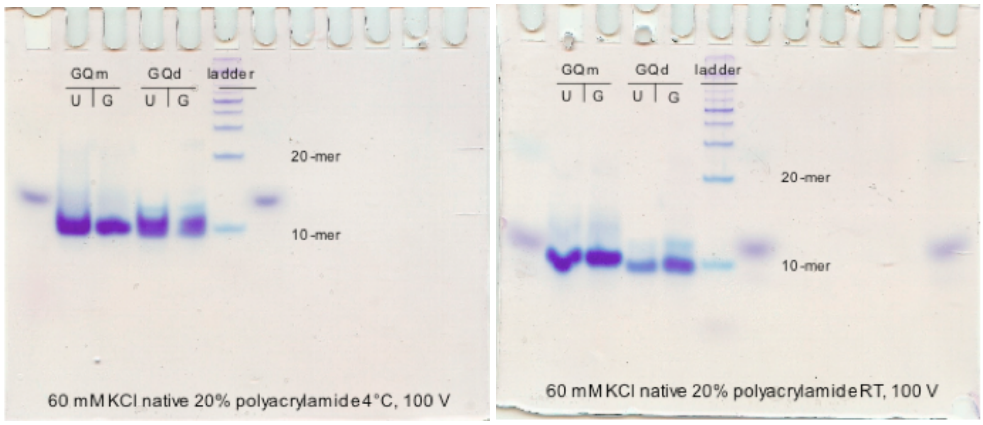
Even in this case, anyway, the same effect was observed.

The following attempt, then, was to compare the behavior of the two sequences, GQm and GQd, in two different temperature conditions. Using potassium chloride as running buffer, the two gels reported below were obtained at 4°C (left) and room temperature (right) with a 15% polyacrylamide gel in native conditions.



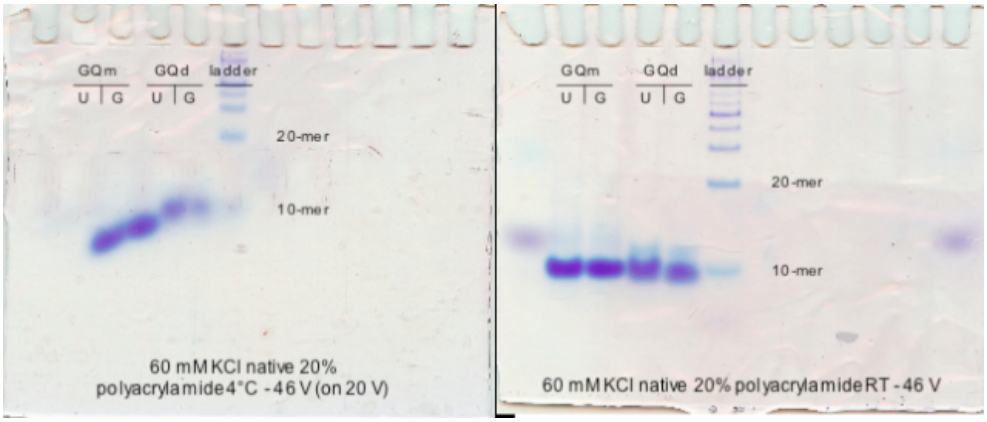
The gel at room temperature seems to differentiate the two sequences. Anyway, even this

experiment does not distinguish between the denatured (U) and non denatured (G) sample. The same experimental conditions were kept for a later experiment, involving a 20% denaturing gel.



Even in this case the experiment was repeated to give a comparison between the gel at 4°C (left) and room temperature (right). Surprisingly, the separation between the two bands (GQm and GQd) was lost again.

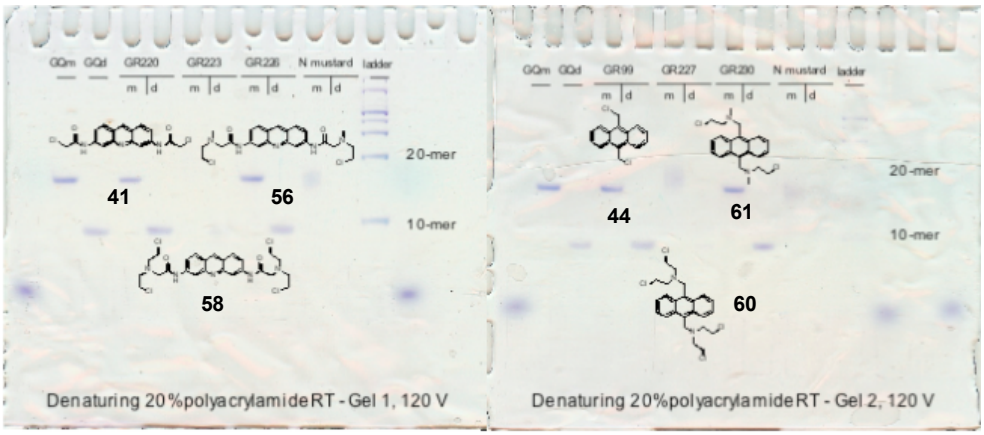
To increase the resolution of the two bands the same gels were run at a lower voltage, but the result was similar to the one obtained before.



Screening of the synthesized compounds

The synthesized nitrogen mustard derivatives were screened on both the sequences. In this

case, a denaturing gel was used. In fact, even if the samples are denatured, a linkage between the oligonucleotides should anyway be enlighten by the gel electrophoresis, as long as nitrogen mustards lead to a covalent cross-link of the DNA sequence, with a corresponding increase of the overall molecular weight.



None of the synthesized compounds, anyway, neither the ones with the proflavine or anthracene scaffold showed the capability of covalently binding DNA.

2.9 ESI-MS OF G-QUADRUPLEX FORMING SEQUENCES

Individuation of a G-quadruplex through MS analysis

This part of the research work was carried out in the laboratory of Prof. Fabris, UAlbany, State University of New York, USA. The presence of the G-quadruplex arrangement in solution was evaluated through both mass spectrometry and gel electrophoresis. As will be further described in the experimental section, the experiments were carried out after incubation of the samples in ammonium acetate to promote the G-quadruplex formation as long as cations are proven to stabilize this structure by disposing in the central channel. Ammonium was chosen instead of potassium (or sodium) because it is reported to be more "MS-friendly", and anyway at the same time is generally accepted as a good model ion for the investigation of quadruplexes through MS analysis.[90]

The ESI-MS experiments on GQm and GQd sequences were carried out after three weeks of incubation of the cleaned-up oligonucleotide sample at 4°C in 150 mM ammonium acetate on a Thermo OrbiTrap Velos LTQ. Samples were diluted do 5 μM in 150 mM ammonium acetate before MS analysis.

ESI mass analysis showed that ammonium adducts to the DNA strands were present, with a good abundance of the + $2NH_4^+$ ion. Considering the structure of the analyzed sequences, a G-quadruplex coming from the self-arrangement of the monomeric strand or from a pairing of two short chains (the dimeric sequence) should be structured as three planar G-quartets stacked one over the other (there are three G in every repeat, the other bases are supposed to form the loops; see the introduction for further details on the structure of G-quadruplexes in general and FIG33-34 for a structural representation). Basing on this, a G-quadruplex made by three G-quartets should be structured around two ammonium ions (the motif should be one ion between two quartet). + $1NH_4^+$ and + $0NH_4^+$ ions, always detected in the

mass analysis, can be assumed as non-G-quadruplex structured DNA sequences coming from non specific interactions between the ammonium ion and the nucleic acid or from the loss of the cation during the mass experiment.

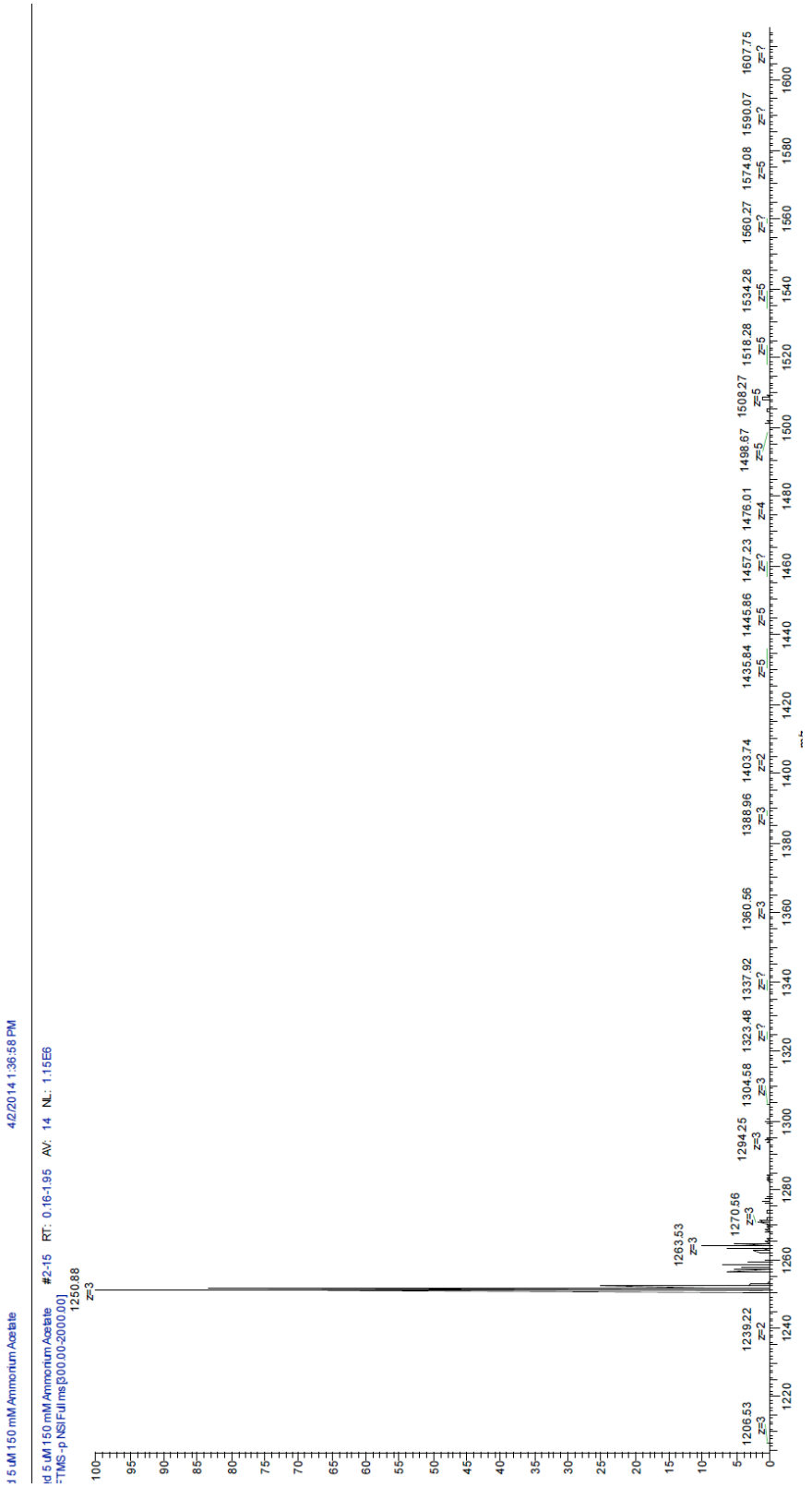


FIG35. ESI-MS full scan spectrum of GQd sequence (5 μM in 150 mm ammonium acetate).

The spectrum above shows a full scan of the ESI-MS experiment on the short GQd dimeric sequence. It is evident how, in this condition, most of the nucleic acid detected by the instrument is in the single strand conformation. The main peak (1251, (m-3)/3) corresponds to the exact mass of the naked sequence (3755.6614 Da) divided by the charge state.

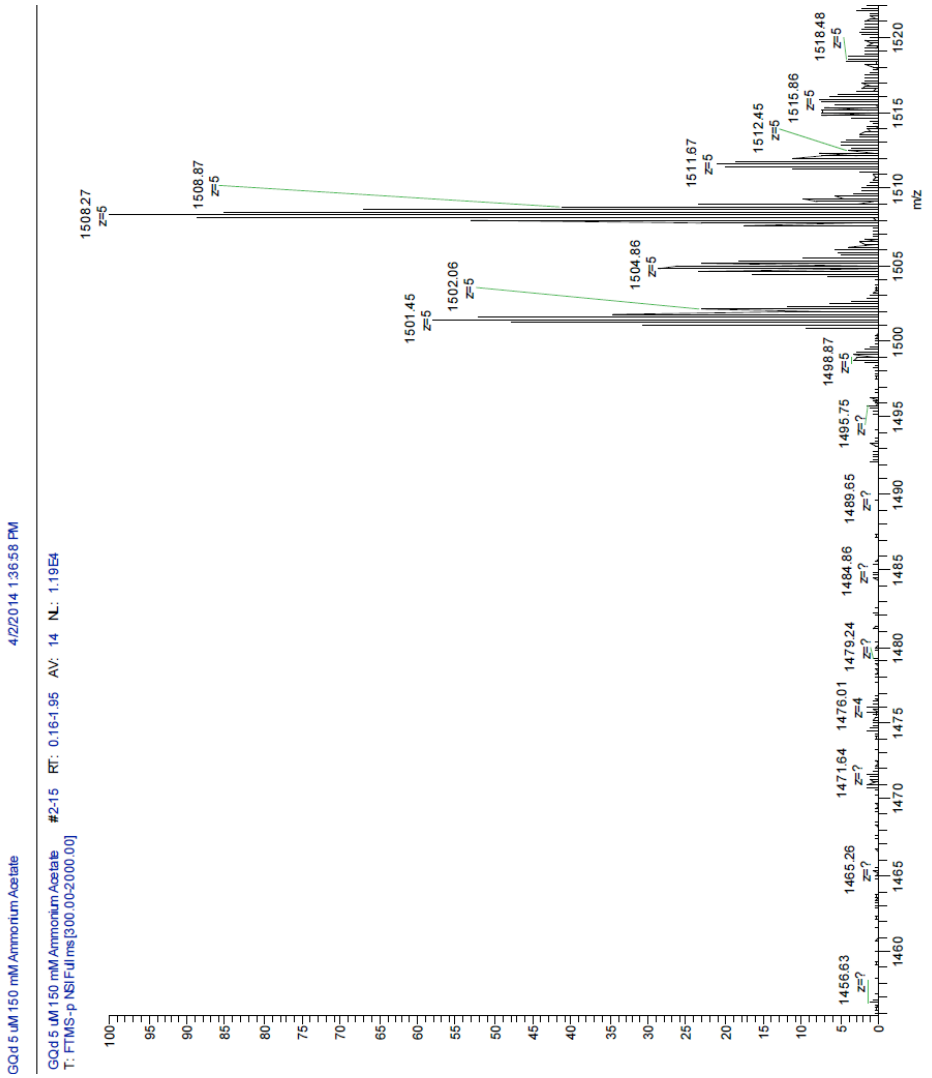


FIG36. Detailed view of the 1400-1550 m/z region.

Only a little amount of quadruplex is present: a detailed view of the 1440-1550 m/z region of the spectrum reveals a peak at 1508 that stands for the combination of two GQd sequences and two ammonium ions, the expected asset for the quadruplex. It is also relevant to notice that the $+2NH_4^+$ peak is the most intense of the group, more than the naked and the $+1NH_4^+$ peak (1501 and 1505 respectively). The same experiment was

carried out on the longer GQm monomeric sequence, under the same conditions.

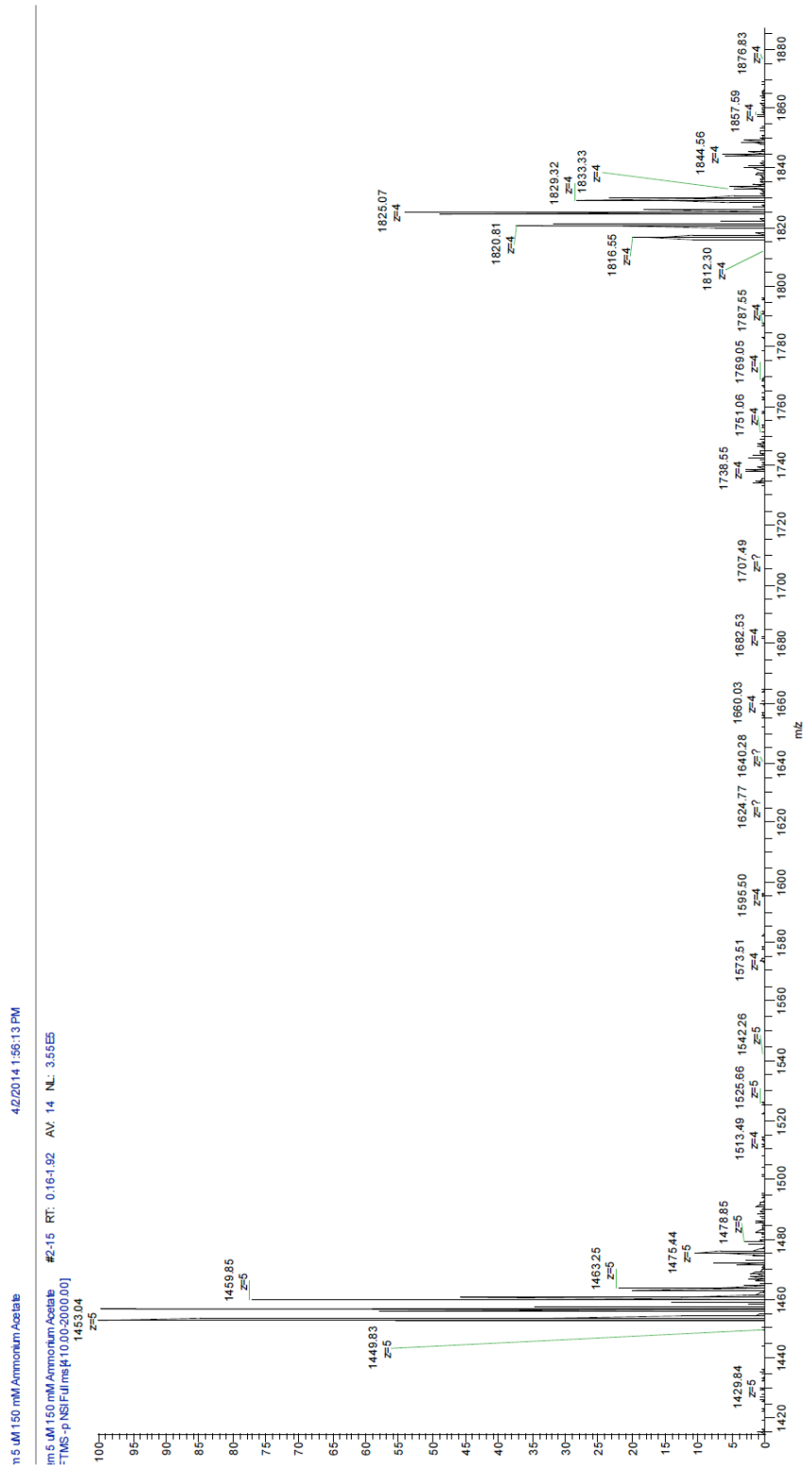


FIG37. ESI-MS full scan spectrum of GQm sequence (5 μM in 150 mm ammonium acetate)

In the spectrum above naked oligonucleotide, mono and di ammonium adducts are clearly detectable, suggesting a higher tendency to form complexes and, probably, lower specificity. It is also true, on the other hand, that considering the -4 charge state the $+2NH_4^+$ peak (1825), that can be assumed as the one representative for the quadruplex structure, is the most intense of the group. The overall presence of the $+2NH_4^+$ peaks, also in the -4 charge state, is much more intense than the one observed with the GQd sequence.

Binding assay

Once that the presence of structured DNA in solution was confirmed, a binding assay involving some of the synthesized compounds were involved. First of all, the capability of the lead compound of binding the quadruplex was evaluated. According to the results of the experiments showed before, the ESI-MS analysis was performed on the GQm sequence using a 5:1 compound to DNA ratio with a final concentration of 5 µM in oligonucleotide.

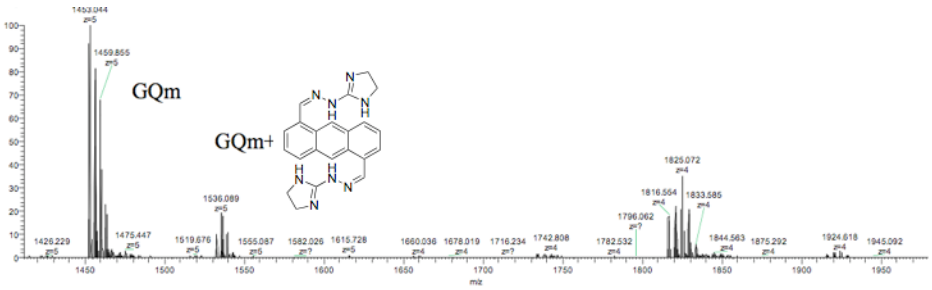
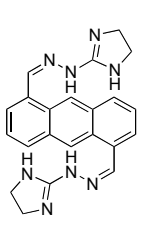
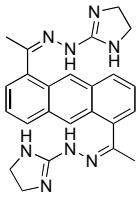


FIG38. Binding experiment of compound **7**

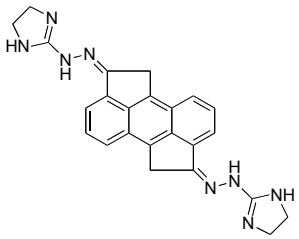
The spectrum above demonstrates the presence of the complex between the lead compound and the GQm sequence, with a relative intensity of more than 20%.



7



51



48

Particular attention was given to the series of constrained compounds structurally related to the lead. A competitive binding study was performed running at once the group of three compounds

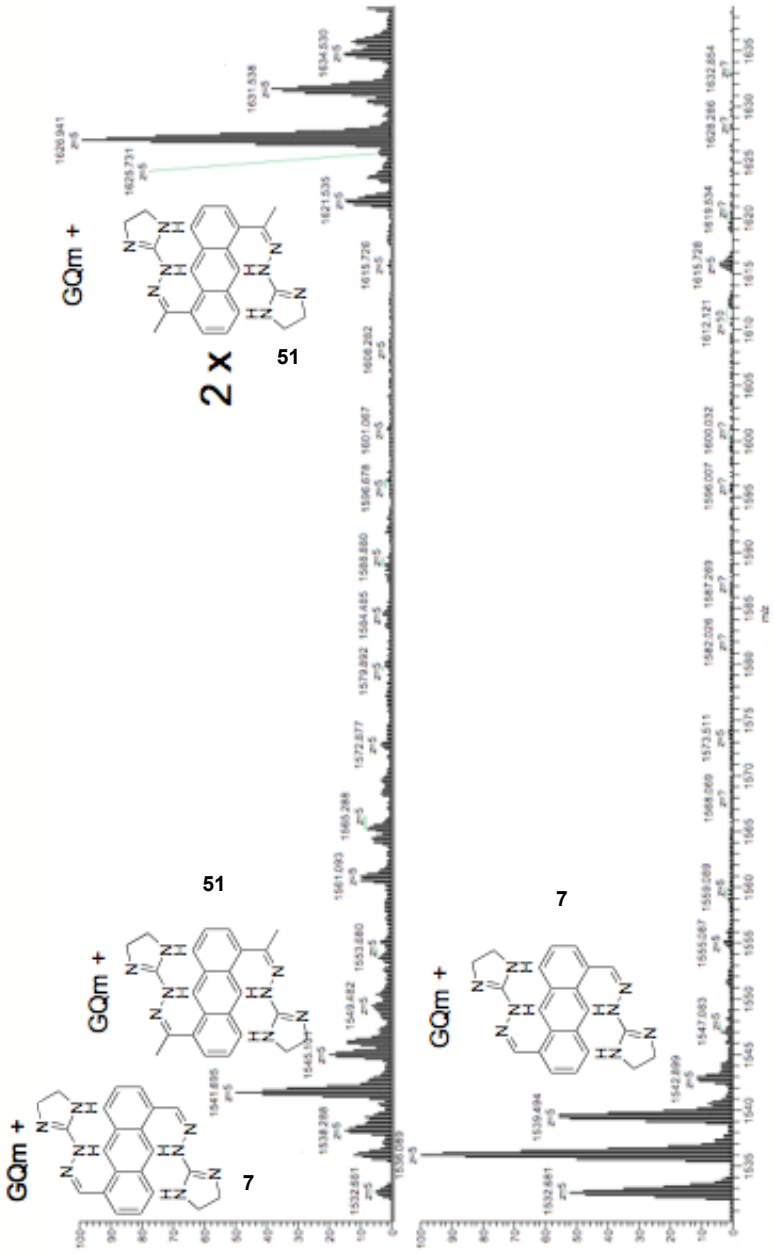


FIG39. Competitive binding experiment of compounds **7**, **51**, **48** compared to the binding experiment of **7** alone

The combined spectrum above compares the spectrum of the lead alone **7** (spectrum below) and the result of the competitive binding study. It appears clearly that compound **51**, the slightly more constrained derivative, has a greater tendency to bind the oligonucleotide. In

addition to this it is also noticeable how this compounds leads to the formation of the 2:1 complex with DNA. The completely constrained derivative didn't show any relevant interaction with the nucleic acid.

Quantitative estimation of binding constant and affinity

The tendency of the synthesized compounds of forming a complex or, in general, binding the G-quadruplex structured nucleic acid is generally expressed in terms of equilibrium association constant (K)[86,] and binding affinity (BA)[79]. Briefly, as long as relative intensities (I) in a mass spectrum are assumed to be proportional to the concentrations of the injected solution

$$I_{(DNA)}/[DNA] = I_{(DNA\text{-stabilizer complex})}/[DNA\text{-stabilizer complex}]$$

and the concentrations of G-quadruplex, of the stabilizer and of the complex at the equilibrium are readily calculated

$$[DNA] = C_0 \bullet I_{(DNA)}/(I_{(DNA)} + I_{(DNA\text{-stabilizer complex})})$$

$$[DNA\text{-stabilizer complex}] = C_0 \bullet I_{(DNA\text{-stabilizer complex})}/(I_{(DNA)} + I_{(DNA\text{-stabilizer complex})})$$

$$[stabilizer] = C_0 - [DNA\text{-stabilizer complex}]$$

the equilibrium association constant value can be estimated as follows:

$$K = [DNA\text{-stabilizer complex}] / [DNA] \bullet [stabilizer]$$

This calculation gives an idea of the relative stability of the complex formed between the folded macromolecule and the low molecular weight binder.

On the other hand, another way to express the tendency of a small molecule of binding G-quadruplex structures investigated by mass spectrometry is the binding affinity (BA).[79]

$$BA = I_{(DNA\text{-stabilizer complex})} / (I_{(DNA\text{-stabilizer complex})} + I_{(DNA)})$$

This formula can be also applied to cases where the binding ratio is higher then 1:1,[79] as the

one reported for compound **51** in the spectrum above:

$$BA = (I_{(DNA\text{-}stab.\ complex\ 1:1)} + I_{(DNA\text{-}stab.\ complex\ 2:1)}) / (I_{(DNA\text{-}stab.\ complex\ 1:1)} + I_{(DNA\text{-}stab.\ complex\ 2:1)} + I_{(DNA)})$$

For what concerning the lead compound **7** a K of 3.1 x 10^{-3} M^{-1} was estimated. As previously reported, anyway, even if this compound gave the best result among the obtained compounds according to the fluorescence melting assay, compound **51** shows a binding affinity more than 1.5 times higher. This, anyway, does not correlate necessarily with the capability of being a good stabilizer: this experiment only estimates the tendency of the synthesized molecule to bind the structured DNA.

Tuning of experimental conditions

To further investigate this event another experiment was carried out: the concentration of the constrained compound was doubled in the contest of a similar competitive binding study.

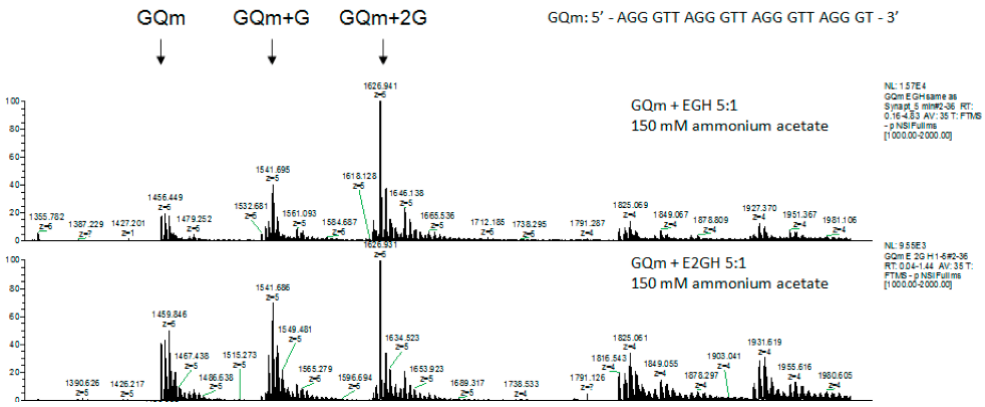


FIG40. Conditions of the experiments were finely tuned

Curiously no evident difference in the spectrum was observed and ratio remains 2:1 as a maximum. The couple of spectra reported above show the original experiment (spectrum above) compared to the one with the doubled concentration of only compound **51** (below).

Getting closer: potassium complexing G-quadruplexes

A further challenge was to carry on some experiments to investigate the possibility of using potassium as G-quadruplex promoting ion even during MS analysis. The use of potassium goes in the direction of obtaining a harmonized model to study G-quadruplex, as long as it was demonstrated how different ions could promote different quadruplex topologies.[1] According to this, the future aim would be to set up the conditions for a multi-technique investigation of the compound-DNA interaction.

To achieve this, the GQm and GQd sequences were stored in a 150 mM solution of potassium acetate following the procedure reported in the experimental section. After suitable dilutions the MS experiment was performed. Preliminary results are shown in the spectrum below (ESI-MS).

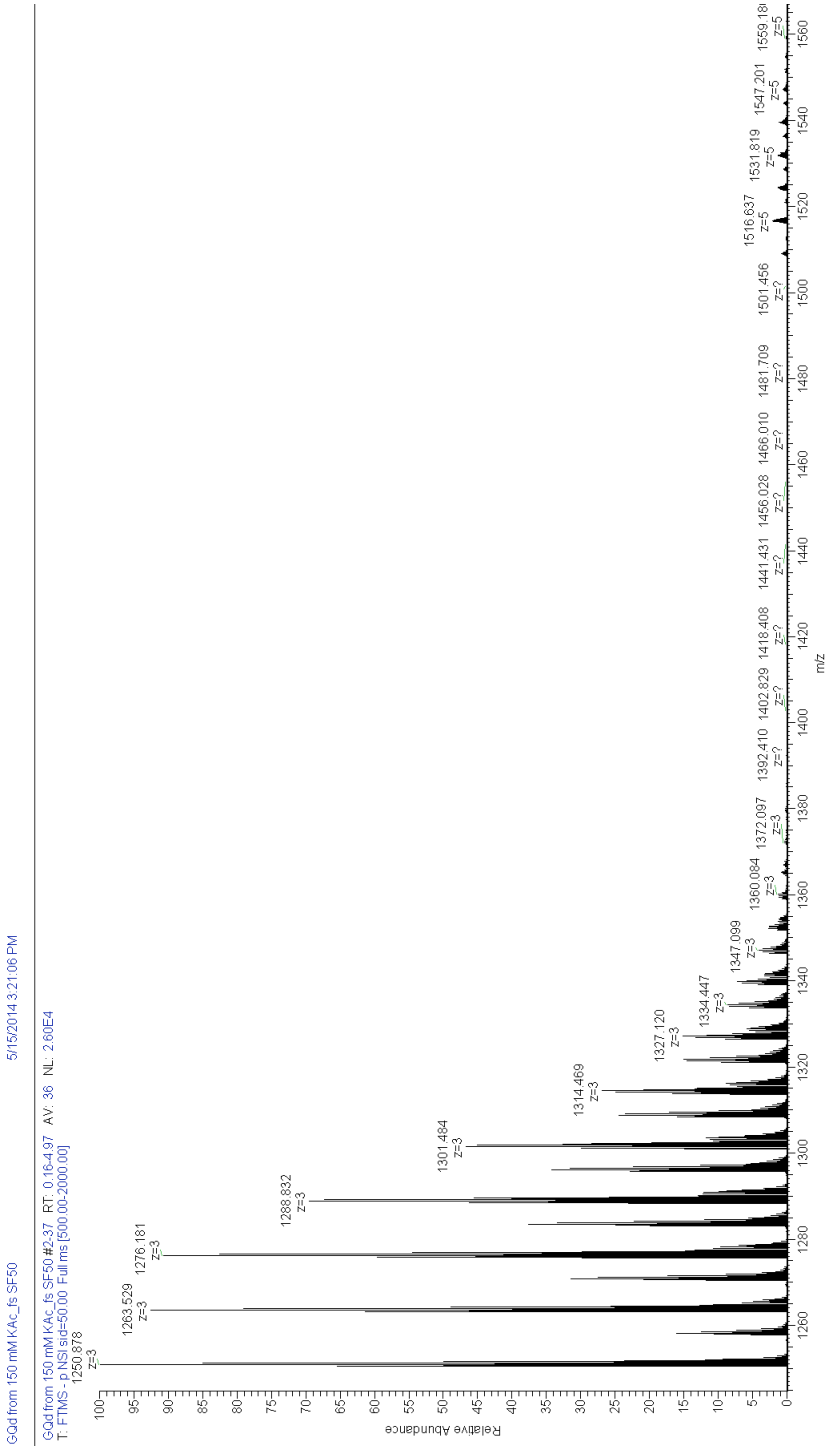


FIG41. Full scan experiment of GQd sequence

In this case the GQd sequence was used. Even if the main series of peak is represented by the GQd peak at 1250.878 with a sequence of K^+ adducts (most realistically a sequence of non specific interactions). Anyway, a little amount of G-quadruplex was observed.

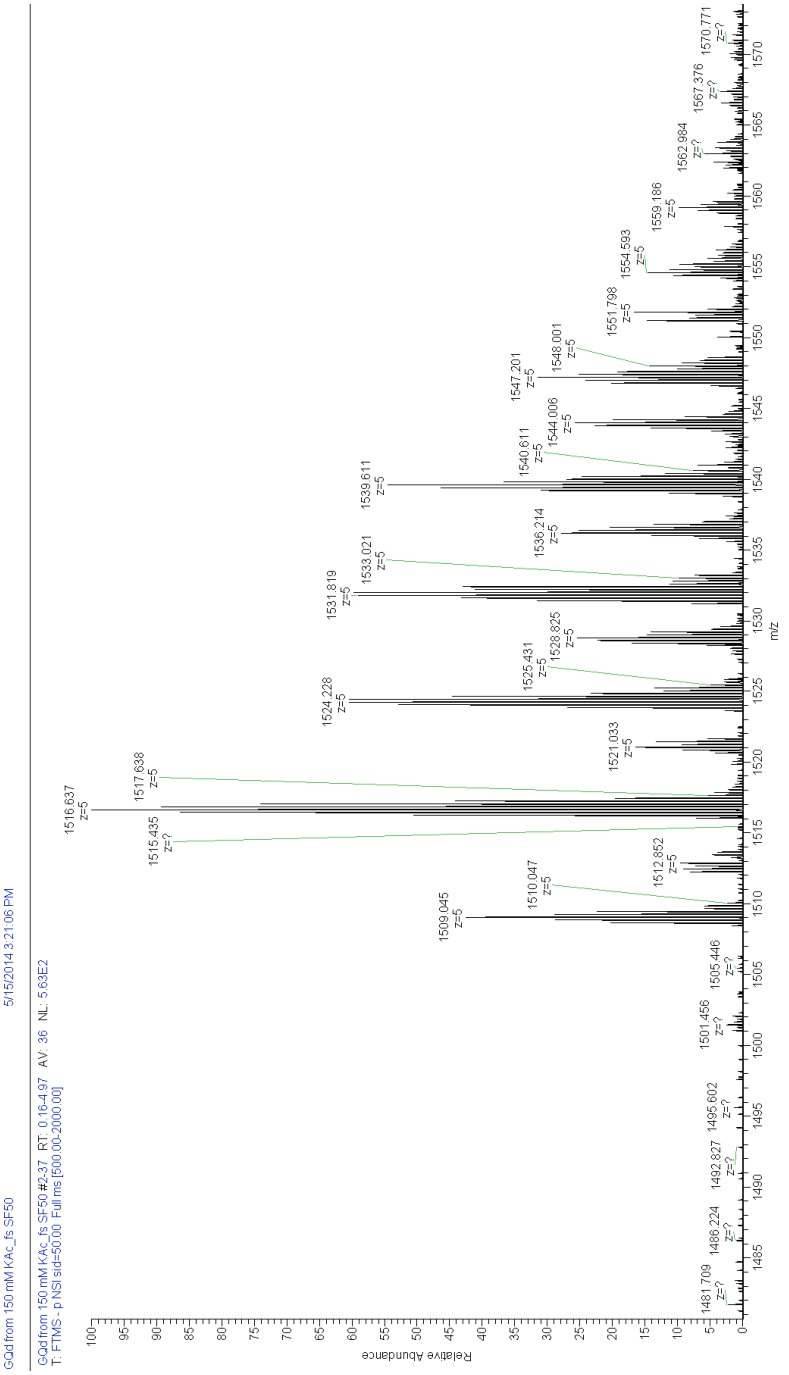


FIG42. Detailed view of a portion of the spectrum

In particular the peak at 1516.637, the higher in relative intensity for its charge state, corresponds to the complex GQd + GQd + $2K^+$, compatible with the quadruplex structure that, as described above, this short dimeric sequence should form.

2.10 ION MOBILITY MASS SPECTROMETRY

An introduction to IMMS

This part of the research work was carried out in the laboratory of Prof. Fabris, UAlbany, State University of New York, USA. Ion mobility mass spectrometry (IMS or IMMS) can be described among these innovative branches of MS evolutions aimed to the investigation of structural and dynamical features of a compound or a complex of biological interest. According to this technique ions can be separated not only by their mass to charge ratio but also by their size and shape.[87]

While the ion mobility phenomenon is known since the beginning of the 20th century, the coupling with mass determination was only later achieved.[88] The "size" of the analyzed species is estimated according to its collision cross section (CCS). In fact, in a general setup of an ion mobility mass spectrometer an ion mobility cell is placed before the actual mass analyzer. Gaseous ions are sprayed into this cell and accelerated by an electric field. The peculiar feature of the cell is the presence of a buffer gas that leads to a certain number of collisions between the analyzed molecule and the gas itself: an higher number of collisions goes together with an higher collision cross section. As a consequence a loss of energy occurs after each collision and the ion takes a longer time to reach the mass analyzer (drift effect). As a result, a simultaneous CCS analysis and m/z separation can be performed. Ideally, the resulting three-dimensional spectrum obtained considers mass, drift time and relative intensity.[89]

The instrument that was used for the measures reported in this work (Waters Synapt G2 HDMS) is based on the electronspray ionization (ESI) or nanoflow ESI and traveling wave ion mobility (TWIM) principle. The ion mobility cell shows subsequent stacked electrode rings interested by voltage pulses that create a wave that an ion can ride (more slowly if the

size is bigger). A time of flight (TOF) device provides then mass analysis.[89]

IMMS and G-quadruplex

Even if mass spectrometry investigation of drug-nucleic acid interactions is now well accepted as a screening technique, ion mobility experiments still represent a developing branch. G-quadruplexes, in particular, represent an outstanding substrate for this kind of analysis according to their peculiar structure, so different in size and shape from the single strand DNA. Very elegant investigations of the DNA arrangement itself[90] and of drug-G-quadruplex interaction[91,92,93] were carried out in the recent years These works reported binding experiments carried out common ESI mass spectrometry and also showed some examples of an approach to ion mobility study. As for this study, the experiments were carried out on telomeric DNA sequences. In this research work a monomer and a dimeric G-quadruplex forming sequence, already studied and accepted and also comparable to the one used for the fluorescence melting experiments, were chosen as substrates.

Discussion of the results

According to the measures carried out on the ion mobility mass spectrometer a peculiar phenomenon was enlightened. In fact, while the cross section of the nucleic acid combined with one ammonium ion increases when compared to the one of the nucleic acid alone, suggesting as expected a mass and size increase, when another ammonium ion is bound to the nucleic acid the cross section decreases when compared to the one of + $1NH_4^+$. This event enlightens that even if the mass increased, the overall size of the complex could be smaller, suggesting the fact that the DNA is actually arranged in a more compact structure like the G-quadruplex. This event was recorded with both the nucleic acid alone in ammonium acetate and with the two compounds that showed a good binding affinity

among the screened ones, confirming this hypothesis. The table below shows graphically what previously explained: how the cross section measured in square angstrom (y axis) changes upon addition of ammonium ions (x axis) to the nucleic acid.

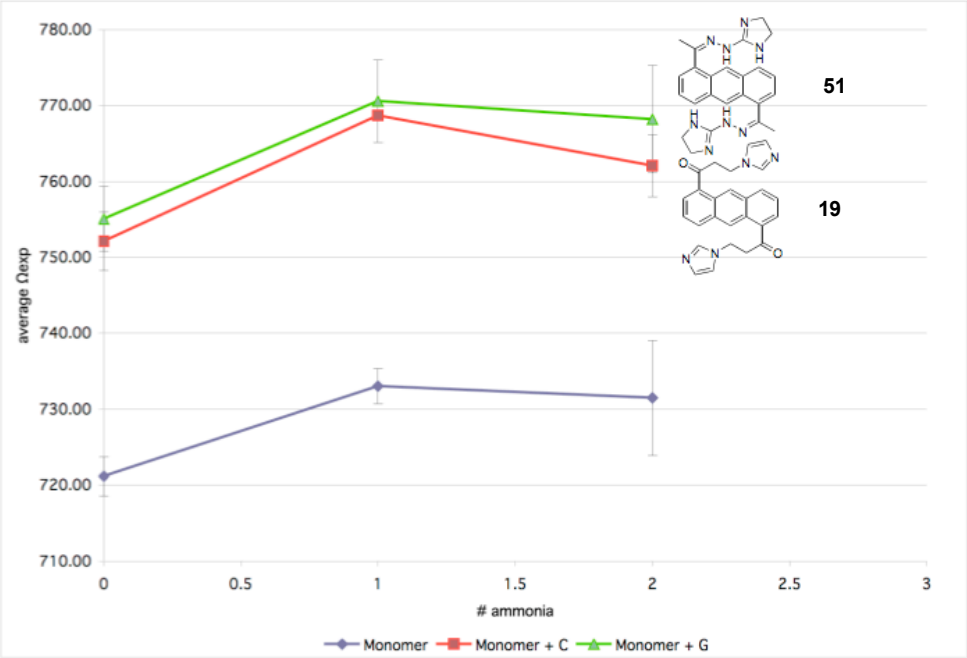


FIG43. Ion mobility profile

The most remarkable data is that the complex between compound **19** (called C in the table), the nucleic acid and two NH_4^+ shows a noticeable slope in the reduction of the cross section when compared to the $+1NH_4^+$ complex, much more intense then the one observed with compound **51** (called G in the table). This interesting piece of information could be the result of a specific binding of the molecule to the quadruplex, resulting in an overall stabilization of the structure. In fact, a reduction of the cross section could be coming from a decrease of the "breathing" movements of the molecule (vibrations, rapid partial unfoldings and refoldings) caused by the stabilization of the G-quadruplex tight arrangement itself. On the other hand, with the other compound the slope of the curve seems comparable to the one of the nucleic acid alone, suggesting a non-specific binding that does not improve the stability of the quadruplex and does not reduce the "breathing" of the complex.

2.11 FLUORESCENCE MELTING

Description of the experiment

The fluorescence melting assay evaluates the stability of a complex super molecular structure according to the denaturation temperature of the structure itself. The efficacy in stabilization of the G-quadruplex by the compounds can be quantified from the variation of the melting temperature (Tm). Tm can be defined as the temperature showing 50 % of structured DNA and 50 % of denaturated DNA. At room temperature (structured DNA) 5' and 3' endings, respectively marked with the cromophore FAM (6-carboxy fluorescein) and the quencher Dabcyl, are close one to the other. Fluorescence observed in this condition is very low as long as Dabcyl has an absorbance maximum at the length of emission of FAM. On the other hand a destructured DNA shows a remarkable increase in fluorescence. Tm is measured at growing concentration of the compound and the ΔTm can be obtained.

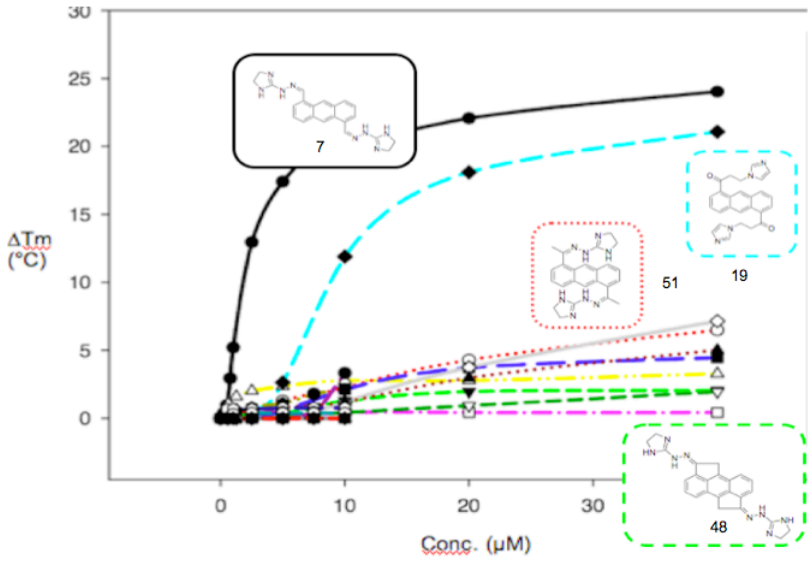


FIG44. Fluorescence melting assay

According to the curves in FIG44, the lead compound remains the best stabilizer of the

telomeric G-quadruplex sequence (TTAGGG repeat) screened. On the other hand, compound **19** reaches very close values.

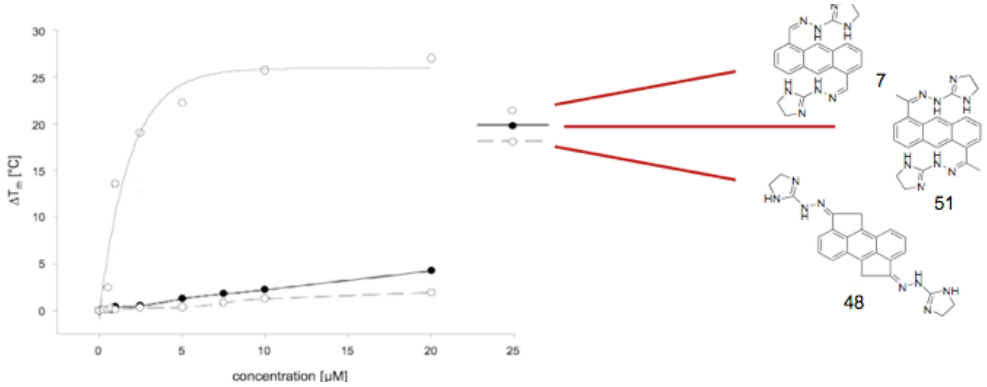


FIG45. Detailed comparison of the fluorescence melting assay for compounds 7, 48, 51

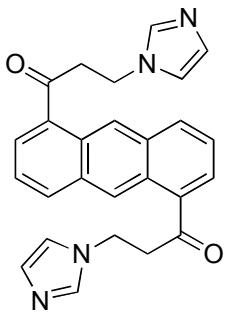
Going back to the constrained derivatives of the lead compound, **48** and **51**, fluorescence melting enlightens the relevance of the role of the flexibility of the side chains in a G-quadruplex stabilizer: compound **51** shows a worst stabilizing capability when compared to **7**, but **48**, the completely rigid molecule, is the worst stabilizer. As previously introduced, the binding affinity measured through MS analysis does not necessarily correspond to the ability of stabilizing structured DNA: even if **48** has a greater binding affinity to DNA and binds also in the 2:1 ratio (probably due to non specific or unefficient/energetically unuseful interactions), **7** remains the best stabilizer. Interestingly, anyway, compound **19** shows promising results both according to mass spectrometry and fluorescence melting experiments.

CONCLUSIONS AND PERSPECTIVES

A wide set of potential G-quadruplex stabilizers was prepared and characterized during this research project and, through the exploration of different chemical classes, some compounds with a promising stabilizing activity were identified.

In addition to the fluorescence melting experiment, during the third year innovative biological evaluation techniques were set up and optimized (ESI-binding MS, tandem MS, ion mobility MS, gel electrophoresis), expanding the field of application of the synthesized molecules from the medicinal chemistry to the world of chemical probing. A docking routine procedure was also optimized for the investigation of the possible binding motifs, opening the path for *in silico* preliminary screenings of designed molecules and for a more detailed study of the mechanism of action/binding fashion of the molecules.

A wide set of compounds with various chemical and structural features was obtained. Structural and conformational properties of the synthesized compounds were deeply investigated and their relevance in obtaining a good stabilization effect was spot out.



19

Moreover a promising molecule, compound **19**, was enlightened as a potential stabilizer by both fluorescence melting and ion mobility MS, confirming the reliability of these complementary techniques in the set up of a screening process. In addition to this, the role of different ions in influencing the behavior and topology of G-quadruplexes and their capability of being targeted by the synthesized molecules is now being investigated tuning

the experimental and sample preparation conditions.

In this connection, the value of this research work goes beyond the encouraging results reached both in the synthetic part and in the optimized evaluation processes; indeed, we developed a multi-technique model-approach for the whole process of the investigation of G-quadruplex and its ligands that involves every step from the *in silico* design to a multi-technology evaluation, going through the synthesis and the enhanced characterization of novel probes\stabilizers. This valuable information will be of a great help in designing novel derivatives and it represents a promising starting point for the further convergent development of these complementary screening techniques.

3. EXPERIMENTAL PROCEDURES

3.1 ABBREVIATIONS

°C	celsius degree
$CDCl_3$	deuterated chloroform
$CHCl_3$	chloroform
δ	delta, ppm
d	doublet
dd	double doublet
DCM	dichloromethane
DMSO	dimethylsulfoxide
D_2O	deuterated water
EtOAc	ethyl acetate
EtOH	ethanol
Et_2O	diethyl ether
g	gram/grams
h	hour/hours
HPLC	High Performance Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
Hz	Hertz
J	coupling constant
L	liter/liters
m	multiplet or milli

M	molarity
MeOD	deuterated methanol
MeOH	methanol
mg	milligram/milligrams
MHz	MegaHertz
min	minutes
mL	milliliters
mmol	millimole/millimoles
mol	mole/moles
MS	mass spectrometry
MW	molecular weight
µ	micro
NMR	Nuclear Magnetic Resonance
ppm	parts per million
rt	room temperature
s	singlet
SAR	structure-activity relationship
t	triplet
THF	tetrahydrofuran
TLC	Thin Layer Chromatography

3.2 MATERIALS

Reactants and reagents

Sigma-Aldrich and Fluka substances were used without any further purification.

Solvents

Normapur Prolabo, Sigma-Aldrich, Riedler-De Haen, VWR and Baker solvents were used.

Oligonucleotides

Oligonucleotides were purchased from IDT - Integrated DNA Technologies (USA).

Deuterated solvents

Deuterated solvents were purchased from Sigma-Aldrich.

TLC Plates

Glass supported Thin Layer Chromatography plates Silica Gel 60 F254 by Merck were used.

Flash chromatography cartridges

10 g and 25 g SNAP Cartridges were purchased by Biotage.

DNA purification

The clean-up procedure operated on DNA oligonucleotides was performed using Merck 3000 MWCO tubes.

Gel electrophoresis

The preparation of native gels for the electrophoresis experiments was carried out according to the procedures reported later in this section and using AccuGel 19:1 40% w/v acrylamide/bis-acrylamide by Natural Diagnostics as starting material. For denaturing gels Ultrapure Sequagel UreaGel System Concentrated, System Diluent and System Buffer by National Diagnostics were used. Thermo O'Range 10 bp was used as ready to use ladder as a molecular weight indicator during the experiments. TBE (Tris/Borate/EDTA) buffer and its dilutions/modifications were prepared according to the reported procedures. Staining of the gels was performed with Sigma Stains All.

3.3 INSTRUMENTATION

Nuclear Magnetic Resonance (NMR)

NMR spectra were recorded on a Bruker AMX 300 MHz and on a Bruker AVANCE III 400 MHz NMR.

High Resolution Mass Spectrometry (HRMS)

An ESI-TOF Mariner instrument from Applied Biosystems was used to perform MS analysis.

Tandem Mass Spectrometry (MS-MS)

MS spectra showing fragmentation MS-MS experiments were recorded on a Thermo LTQ OrbiTrap Velos.

Ion Mobility Mass Spectrometry (IMMS)

Ion mobility data sets were acquired on a Waters Synapt G2 HDMS.

Automatic flash chromatography

An Isolera One chromatographic apparatus from Biotage was used for preparative flash chromatography.

UV-Vis Spectrophotometer

A Thermo NanoDrop 2000c spectrophotometer was used to quantify DNA samples and

compounds.

High Performance Liquid Chromatography (HPLC)

A Varian ProStar 210 combined with a Zorbax Eclipse XDB-C8 by Agilent Technologies column was used for HPLC analysis. An appropriate ratio of water (A) and acetonitrile (B) was used as mobile phase with an overall flow rate of 1 mL min^{-1}; the general method for the analyses is here reported: 0 minutes (90% A-10% B), 15 minutes (10% A-90% B), 20 minutes (10% A-90% B), 21 minutes (90% A-10% B), 25 minutes (90% A-10% B). The purity of all compounds was ≥ 95%, unless otherwise stated.

Gel electrophoresis equipment

E.C. Apparatus Corporation electrophoresis apparatus was used together with Bio-Rad equipment.

3.4 COMPUTATIONAL ANALYSIS AND ARTWORKS

All the artworks reported in this manuscript are original and produced specifically for this research work. The .pdb files of proteins and nucleic acids were downloaded from the RCSB Protein Data Bank.[94] If the artwork represents an unmodified X-ray or NMR solved structure, the reference number leads to the article that is connected to the .pdb file of interest.[94] The workflow for their visualization or for the structure editing and analysis was designed in order to take advantage of the usage of freeware, multiplatform software, running the experiments on a Unix G4 machine.

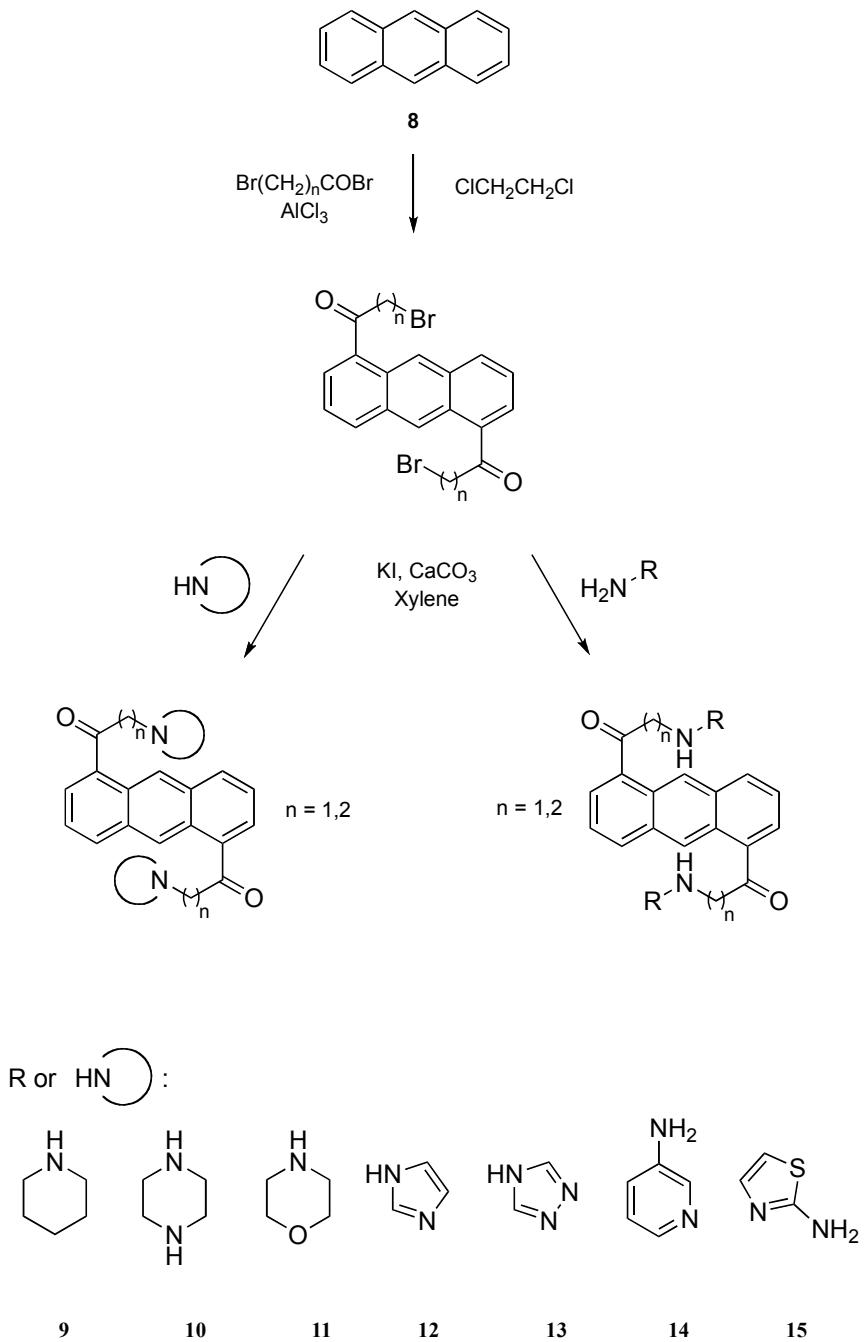
The structures of the ligands of interest were designed using the JSDraw online tool[95] and exported to the .mol format. The Avogadro software[96] was used to process these files and, in particular, for a preliminary visualization of the three dimensional structure of ligands, nucleic acid and their complexes. In addition to this, manual or automated sculpting and energy minimization scripts (MMFF94 and UFF force fields) from this software were applied to obtain a rough model representing the interaction motif. In some preliminary model analysis, PyMOL was used.[97] Figures were then obtained exploring the resulting .mol or .pdb files using UCSF Chimera.[98] This software was also used for the structure analysis/measurements of both the designed ligands and the .pdb files from RCSB Protein Data Bank.[94]

Structure to name conversion was operated using ChemDoodle Web Components online tools.[99]

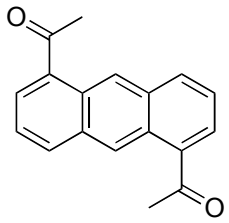
For what concerning docking studies, AutoDockTools 1.5.6 was used to establish the Autogrid points and to visualize the docked poses resulting from the Autodock4 analysis.[100] Both ligand and G-quadruplex 3D models were optimized for the experiment removing the co-crystallized ligand, adding polar hydrogens to the structures and computing partial charges. The docking grid maps were spaced at 0.500 Å, and the center of the G-

quadruplex structure was set as the grid center. One hundred docking runs were performed. This procedure was optimized and carried out in accordance to other recently reported docking studies of small molecules on G-quadruplexes.[79]

3.5 SCHEME I: PROCEDURES



Anthracene-1,5-diacetyl (50)



A three-neck round-bottom flask was charged with $AlCl_3$ (23 g, 168 mmol) in 150 mL of dichloroethane. The stirred suspension was cooled in an ice bath and acetyl chloride (11.8 mL, 168 mmol) was added. Anthracene (10 g, 56 mmol) was added in portions avoiding the temperature to go over 0°C. After the addition of the reactants, the red mixture was allowed to stir for 4 hours at room temperature. The proceeding of the reaction was monitored through TLC (dichloromethane, ethyl acetate 98:2). The red solid collected from filtration was poured in a mixture of 37% HCl (40 mL) and ice (400 g) and let under vigorous stirring for additional 2 hours. A yellow solid was then collected by filtration and recrystallized from acetic acid and then from toluene. Yield: 30%.

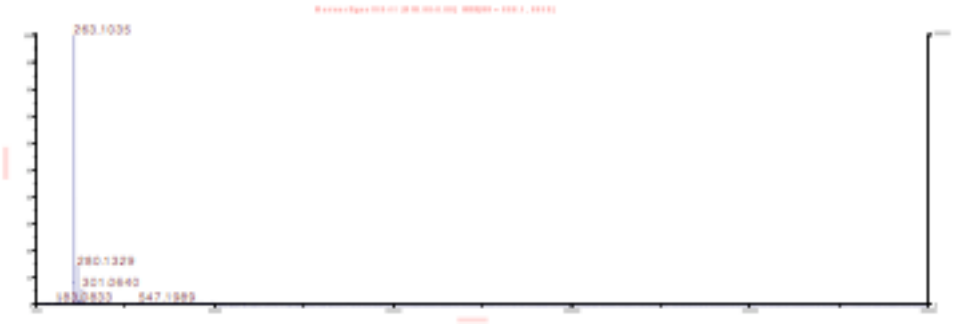
$C_{18}H_{14}O_2$

MW: 262.30

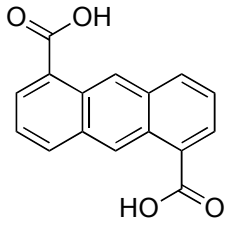
1H-NMR: δH (400 MHz, CDCl3) 9.60 (2H, s, ArH), 8.29 (2H, dd, J 8.4 Hz, J 0.8 Hz, ArH), 8.12 (2H, dd, J 7.0 Hz, J 1.1 Hz, ArH), 7.56 (2H, dd, J 8.4 Hz, J 7.0 Hz, ArH), 2.85 (6H, s, CH3)

13C-NMR {1H}: δC (100 MHz, CDCl3) 199.6, 136.5, 133.2, 132.8, 131.3, 129.6, 126.6, 121.5, 29.9

HRMS (ESI): exp. 263.1035 (M+1), calc. 263.0994



Anthracene-1,5-dicarboxylic acid (63)



A three-neck round-bottom flask was charged with **50** (1 g, 3.812 mmol) in 200 mL of ethanol and the mixture was heated to reflux. A freshly prepared solution of NaOBr, obtained by mixing 33 mL of NaOH 30% and 3.50 mL of bromine in an ice bath, was then dropped in the flask during 1 hour. This procedure was repeated twice and then the reaction was allowed to stir at refluxing temperature overnight. The reaction was quenched by the addition of a solution of 7 g of $Na_2S_2O_5$ in 200 mL of water. After the addition of 10% HCl a solid formed, which was collected by filtration. Another solid was collected from the concentration of the filtered solution. An extraction with ethanol from the solid in a soxhelet led to a liquid that, after evaporation, gave the product. Yield: 66%.

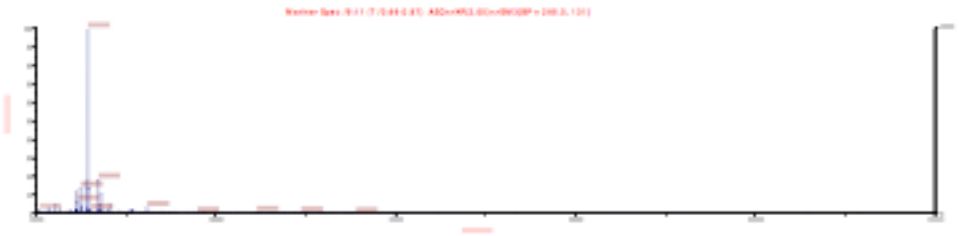
$C_{16}H_{10}O_4$

MW 266.25

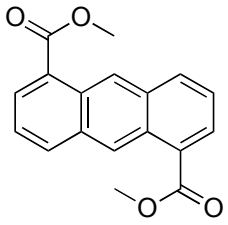
1H-NMR: δH (400 MHz, MeOD) 9.61 (2H, m, ArH), 8.27 (4H, m, ArH), 7.59 (2H, m, ArH)

13C-NMR {1H}: δC (100 MHz, MeOD) 169.8, 136.5, 132.9, 129.4, 128.2, 127.5, 125.3, 121.5

HRMS (ESI): exp. 265.0525 (M-1), calc. 265.0600



Dimethyl anthracene-1,5-dicarboxylate (64)



A round-bottom flask was charged with 600 mg of **63** and 25 mL of ethanol. After the dissolution of the intermediate, allowed by the warming of the mixture, freshly prepared NaOMe (from MeOH and Na) was added (365 mg, 6.761 mmol). Dimethylsulfate (640 μL, 6.761 mmol) was added after that the reaction mixture was cooled to room temperature. The obtained solution was heated to reflux and allowed to stir overnight. The residue obtained after the evaporation of the solvent was washed with water and 205 mg of the desired compound were obtained. Yield: 31%.

$C_{18}H_{14}O_4$

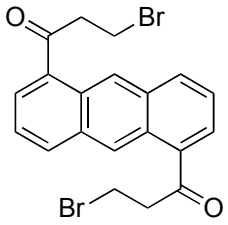
MW 294.30

1H-NMR: δH (400 MHz, MeOD) 9.60 (2H, m, ArH), 8.33 (4H, m, ArH), 7.61 (2H, m, ArH), 4.62 (3H, s, CH3), 4.07 (3H, s, CH3)

13C-NMR {1H}: δC (100 MHz, MeOD) 168.9, 136.4, 132.8, 130.2, 128.3, 127.1, 125.8, 121.5, 53.3

HRMS (ESI): exp. 295.1006 (M+1), calc. 295.0900

3-Bromo-1-{5-(3-bromopropionyl)-1-anthryl}-1-propanone (65)



A three-neck round-bottom flask was charged with $AlCl_3$ (23.0 g, 168.3 mmol) in 150 mL of dichloroethane. The stirred suspension was cooled in an ice bath and 3-Br propionyl chloride (17.0 mL, 168.3 mmol) was added. Anthracene (10.0 g, 56.1 mmol) was added in portions avoiding the temperature to go over 0°C. After the addition of the reactants, the red mixture was allowed to stir for 4 hours at room temperature. The proceeding of the reaction was monitored through TLC (dichloromethane, ethyl acetate 98:2). The dark yellow solid collected from filtration was poured in a mixture of 37% HCl (40.0 mL) and ice (400.0 g) and let under vigorous stirring for additional 3 hours. A yellow solid was then collected by filtration, wich was refluxed in 100 mL of acetic acid for 2 hours. The non dissolved solid was then dissolved in chloroform and the organic solution was washed with basic water (K_2CO_3). The organic phase was then evaporated to give the product. Yield: 33%.

$C_{20}H_{16}Br_2O_2$

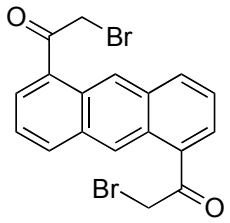
MW 448.15

1H-NMR: δH (300 MHz, CDCl3) 9.46 (2H, s, ArH), 8.22 (2H, d J 9.6 Hz, ArH) , 8.00 (2H, dd J 6.7 Hz, J 1.1 Hz, ArH), 7.50 (2H, dd J 8.4 Hz, J 1.6 Hz, ArH) 3.8-3.7 (4H, m, CH2), 3.7-3.6 (4H, m, CH2)

13C-NMR {1H}: δC (75 MHz, CDCl3) 199.21, 135.5, 132.7, 132.5, 131.3, 126.6, 125.9, 121.5, 30.9, 30.2

HRMS (ESI): exp. 446.9012 (M+1), calc. 446.9524

2-Bromo-1-{5-(2-bromoacetyl)-1-anthryl}-1-ethanone (66)



A three-neck round-bottom flask was charged with $AlCl_3$ (23.0 g, 168.3 mmol) in 150 mL of dichloroethane. The stirred suspension was cooled in an ice bath and 3-Br acetyl chloride (14.7 mL, 168.3 mmol) was added. Anthracene (10.0 g, 56.1 mmol) was added in portions avoiding the temperature to go over 0°C. After the addition of the reactants, the red mixture was allowed to stir for 4 hours at room temperature. The proceeding of the reaction was monitored through TLC (dichloromethane, ethyl acetate 98:2). The dark yellow solid collected from filtration was poured in a mixture of 37% HCl (40.0 mL) and ice (400.0 g) and let under vigorous stirring for additional 3 hours. A yellow solid was then collected by filtration; which was refluxed in 100 mL of acetic acid for 2 hours. The non dissolved solid was then dissolved in chloroform and the organic solution was washed with basic water (K_2CO_3). The organic phase was then evaporated to give the product. Yield: 12%.

$C_{18}H_{12}Br_2O_2$

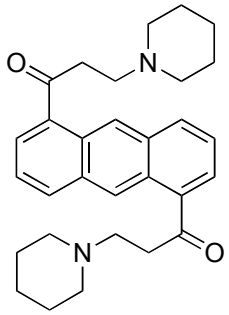
MW 420.09

1H-NMR: δH(400 MHz, CDCl3) 9.53 (2H, s, ArH), 8.35 (2H, d J 8.7 Hz,) , 8.00 (2H, dd J 7.0 Hz, J 1.0 Hz, ArH), 7.50 (2H, t, J 8.7 Hz, ArH) 4.70 (4H, s, CH2)

13C-NMR {1H}: δC (100 MHz, CDCl3) 191.2, 136.4, 133.2, 131.2, 129.5, 126.6, 125.9, 121.5, 32.3

HRMS (ESI): exp. 419.0021 (M+1), calc. 418.9212

3-Piperidino-1-{5-(3-piperidinopropionyl)-1-anthryl}-1-propanone (67)



A round-bottom flask was charged with **65** (50 mg, 0.112 mmol) and 30 mL of xylene. $CaCO_3$ (33 mg, 0.335 mmol), KI (55 mg, 0.335 mmol) and piperidine (33 μL, 0.335 mmol) were then added. The reaction mixture was allowed to stir at reflux for 7 hours and the proceedings were monitored through TLC (chloroform, methanol 6:1). The obtained suspension was then filtered. The solid collected was dried in oven for 2 hours and then dissolved in 20 mL of chloroform, filtering off the insoluble salts. The filtrate was washed with basic water (K_2CO_3) and the organic phase evaporated to dryness giving the desired compound as a brown solid. Yield: 71%.

The hydrochloride salt of the compound can be easily obtained: to a solution of the compound in ethanol, HCl 37% was slowly added to acid pH. The mixture was stirred at room temperature for 2 hours and then the solvent was evaporated.

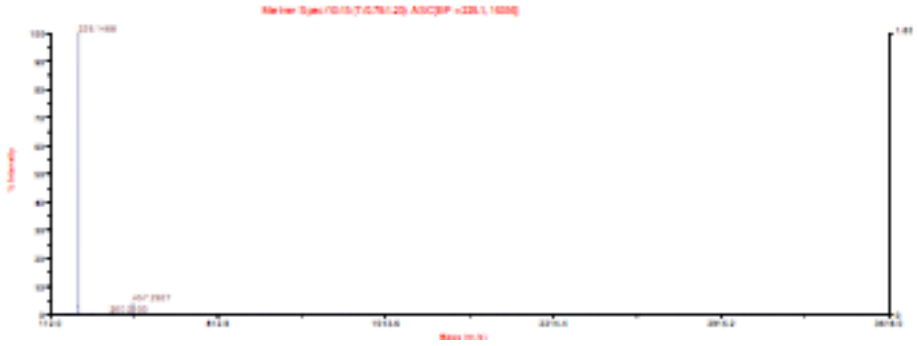
$C_{30}H_{36}N_2O_2$

MW 456.62

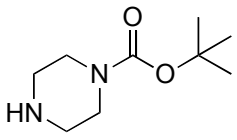
1H-NMR: δH (400 MHz, D2O) 9.19 (2H, s, ArH), 8.23 (2H, d J 8.6 Hz, ArH) , 8.15 (2H, d J 6.8 Hz, ArH), 7.55 (2H, m, ArH), 7.50 (2H, dd J 7.6 Hz, J 6.4 Hz, ArH), 3.69 (4H, t J 6.4 Hz, CH2), 3.6-3.5 (8H, m, CH2), 2.99 (4H, t J 11.8 Hz, CH2), 1.8-1.7 (8H, m, CH2)

13C-NMR {1H}: δC (100 MHz, D2O) 199.2, 136.5, 133.3, 132.6, 131.3, 126.6, 125.9, 121.2, 51.4, 49.2, 35.2, 26.7, 24.0

HRMS (ESI): exp. 229.1468 ((M+2)/2), calc. 229.1461; exp. 457.2927 (M+1), calc. 457.2850



Tert-butyl piperazine-1-carboxylate (N-Boc piperazine) (68)



In a round-bottom flask piperazine (12 g, 0.139 mol) was dissolved in 500 mL of dichloromethane. The obtained solution was stirred and cooled in an ice bath. In another flask di-tert-butyl dicarbonate (3.8 g, 0.017 mol) was dissolved in 100 mL of dichloromethane. This solution was slowly added drop by drop to the piperazine solution in 2.5 h and the reaction mixture was then allowed to stir at room temperature overnight. The white solid obtained was removed by filtration and the filtered solution was evaporated to give a white residue that was suspended in dichloromethane. This suspension was washed with water and the organic phase was evaporated to give the product. Yield: 83 %.

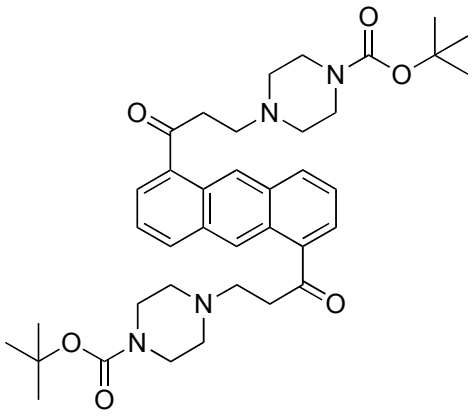
$C_9H_{18}N_2O_2$

MW 186.25

1H-NMR: δH (400 MHz, Acetone) 3.30 (4H, m, CH2), 2.70 (4H, m, CH2), 1.44 (9H, s, CH3)

13C-NMR {1H}: δC (100 MHz, Acetone) 160.0, 81.4, 45.8, 44.9, 28.2

tert-Butyl 4-(3-oxo-3-{5-[3-(4-tert-butoxycarbonyl-1-piperazinyl)propionyl]-1-anthryl}propyl)-1-piperazinecarboxylate (69)



A round-bottom flask was charged with **65** (50 mg, 0.112 mmol) and 30 mL of xylene. $CaCO_3$ (33 mg, 0.335 mmol), KI (55 mg, 0.335 mmol) and N-Boc piperazine (62 mg, 0.335 mmol) were then added. The reaction mixture was allowed to stir at reflux for 7 hours and the proceedings were monitored through TLC (chloroform, methanol 6:1). The obtained suspension was then filtered. The solid collected was dried in oven for 2 hours and then dissolved in 20 mL of chloroform, filtering off the insoluble salts. The filtrate was washed with basic water (K_2CO_3) and the organic phase evaporated to dryness giving the desired compound as a brown solid. Yield: 52%.

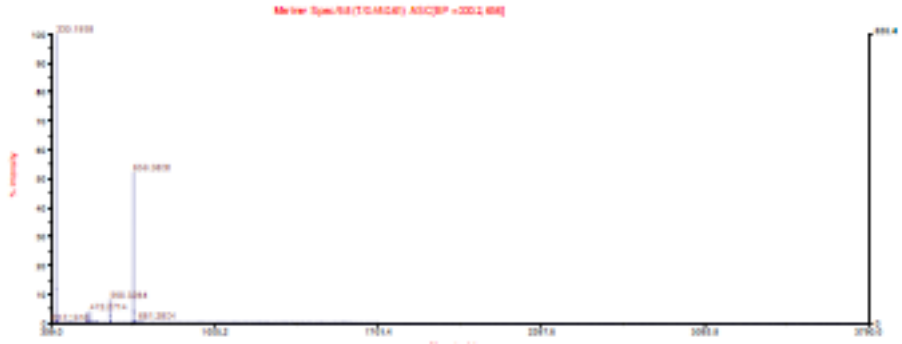
$C_{38}H_{50}N_4O_6$

MW 658.83

1H-NMR: δH (300 MHz, MeOD) 8.12 (2H, s, ArH), 8.02 (2H, dd J 8.9 Hz, J 8.4 Hz, ArH), 7.98 (2H, d J 8.9 Hz, ArH), 7.56 (2H, d J 8.4 Hz, ArH), 2.90 (4H, t, CH2), 2.78 (4H, t, CH2), 2.69 (8H, m, CH2), 2.60 (8H, m, CH2), 2.22 (18H, s, 6 x CH3)

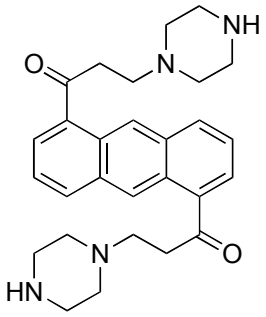
13C-NMR {1H}: δC (75 MHz, MeOD) 200.2, 154.8, 136.6, 133.0, 132.6, 131.3, 126.9, 125.4, 121.8, 81.3, 52.9, 49.6, 43.6, 35.2, 28.9

HRMS (ESI): exp. 330.1958 ((M+2)/2), calc. 330.1938; exp. 659.3808 (M+1), calc. 659.3803



3-(1-Piperazinyl)-1-{5-[3-(1-piperazinyl)propionyl]-1-anthryl}-1-propanone

(70)



The N-Boc protected compound **69** (38 mg, 0.050 mmol) was dissolved in 5 mL of hot dichloromethane. After cooling to room temperature 5 mL of trifluoroacetic acid were added and the mixture was allowed to stir at room temperature for 1 hour. The solvent was then removed under reduced pressure. The residue was suspended in toluene and the reaction mixture was dried again to give the compound. Yield: 91%.

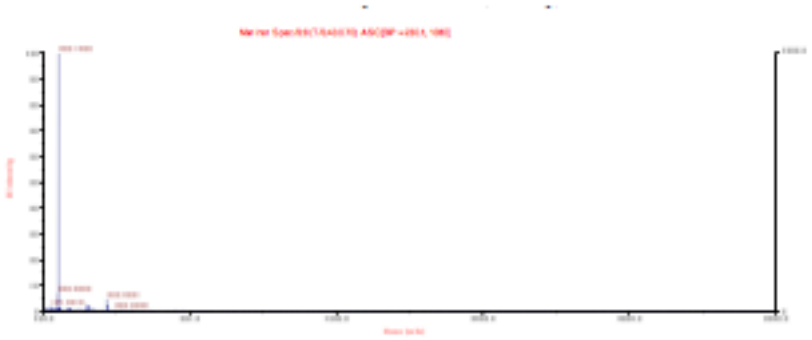
$C_{28}H_{34}N_4O_2$

MW 458.60

1H-NMR: δH (300 MHz, MeOD) 8.03 (2H, s, ArH), 8.00 (2H, dd J 9.1 Hz, J 8.3 Hz, ArH), 7.98 (2H, d J 8.3 Hz, ArH), 7.90 (2H, d J 9.1 Hz, ArH), 2.91 (4H, t, CH2), 2.72 (4H, t, CH2), 2.68 (8H, m, CH2), 2.52 (8H, m, CH2)

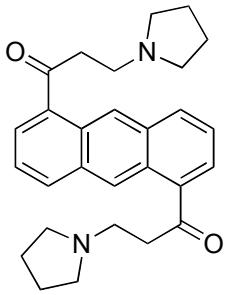
13C-NMR {1H}: δC (75 MHz, MeOD) 199.2, 136.5, 131.2, 130.2, 129.4, 126.6, 125.4, 120.2, 54.5, 50.4, 44.8, 35.2

HRMS (ESI): exp. 230.1458 ((M+2)/2), calc. 230.1414; exp. 459.2861 (M+1), calc. 459.2755



3-(1-Pyrrolidinyl)-1-{5-[3-(1-pyrrolidinyl)propionyl]-1-anthryl}-1-propanone

(71)



A round-bottom flask was charged with **65** (50 mg, 0.112 mmol) and 30 mL of xylene. $CaCO_3$ (33 mg, 0.335 mmol), KI (55 mg, 0.335 mmol) and pyrrolidine (24 mg, 0.335 mmol) were then added. The reaction mixture was allowed to stir at reflux for 7 hours and the proceedings were monitored through TLC (chloroform, methanol 6:1). The obtained suspension was then filtered. The solid collected was dried in oven for 2 hours and then dissolved in 20 mL of chloroform, filtering off the insoluble salts. The filtrate was washed with basic water (K_2CO_3) and the organic phase evaporated to dryness giving the desired compound as a brown solid. Yield: 79%.

The hydrochloride salt of the compound can be easily obtained: to a solution of the compound in ethanol, HCl 37% was slowly added to acid pH. The mixture was stirred at room temperature for 2 hours and then the solvent was evaporated.

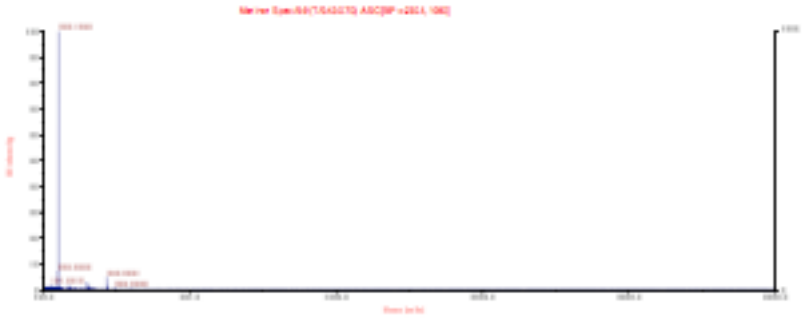
$C_{28}H_{32}N_2O_2$

MW 428.57

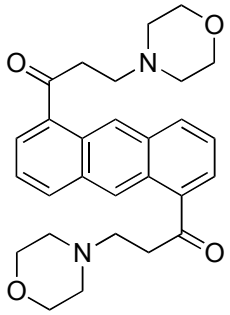
1H-NMR: δH (400 MHz, CDCl3) 9.32 (2H, s, ArH), 8.13 (2H, d, J 8.43 Hz, ArH), 7.94 (2H, dd, J 6.9 Hz, J 0.8 Hz, ArH), 7.44 (2H, t, J 8.50 Hz, ArH), 3.30 (4H, m, CH2), 3.0-2.9 (4H, m, CH2), 2.52 (8H, m, CH2), 1.72 (8H, m, CH2)

13C-NMR {1H}: δC (75 MHz, CDCl3) 200.1, 136.7, 132.8, 131.0, 126.9, 125.4, 120.9, 53.4, 49.1, 35.2, 22.5

HRMS (ESI): exp. 215.1394 ((M+2)/2), calc. 215.1305



3-Morpholino-1-{5-(3-morpholinopropionyl)-1-anthryl}-1-propanone (72)



A round-bottom flask was charged with **65** (50 mg, 0.112 mmol) and 30 mL of xylene. $CaCO_3$ (33 mg, 0.335 mmol), KI (55 mg, 0.335 mmol) and morpholine (29 µL, 0.335 mmol) were then added. The reaction mixture was allowed to stir at reflux for 7 hours and the proceedings were monitored through TLC (chloroform, methanol 6:1). The obtained suspension was then filtered. The filtrate was evaporated and the residue was dissolved in chloroform. This suspension was washed with basic water (K_2CO_3). The organic phase was evaporated to give the product as brown solid. Yield: 64%.

The hydrochloride salt of the compound can be easily obtained: to a solution of the compound in ethanol, HCl 37% was slowly added to acid pH. The mixture was stirred at room temperature for 2 hours and then the solvent was evaporated.

$C_{28}H_{32}N_2O_4$

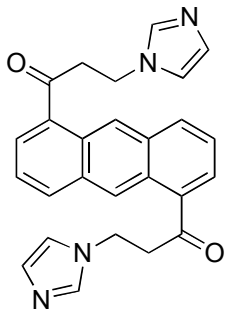
MW 460.57

1H-NMR: δH(400 MHz, D2O) 9.27 (2H, s, ArH), 8.28 (2H, d, J 8.16 Hz, ArH), 8.19 (2H, d, J 7.08 Hz, ArH), 7.6-7.5 (2H, m, ArH), 3.9-3.8 (4H, m, CH2), 3.70 (4H, m, CH2), 3.67 (8H, m, CH2), 3.2-3.1 (4H, m, CH2)

13C-NMR {1H}: δC (75 MHz, D2O) 200.0, 136.5, 132.9, 131.0, 126.6, 125.5, 121.5, 66.4, 55.0, 49.4, 35.2

HRMS (ESI): exp. 231.1274 ((M+2)/2) calc. 231.1254

3-(1H-Imidazol-1-yl)-1-{5-[3-(1H-imidazol-1-yl)propionyl]-1-anthryl}-1-propanone (19)



A round-bottom flask was charged with **65** (50 mg, 0.112 mmol) and 30 mL of xylene. $CaCO_3$ (33 mg, 0.335 mmol), KI (55 mg, 0.335 mmol) and imidazole (24 mg, 0.335 mmol) were then added. The reaction mixture was allowed to stir at reflux for 7 hours and the proceedings were monitored through TLC (chloroform, methanol 6:1). The obtained suspension was then filtered. The filtrate was evaporated and the residue was dissolved in chloroform. This suspension was washed with basic water (K_2CO_3). The organic phase was evaporated to give the product as yellow solid. The solid collected from the filtration described above was dried in oven for 2 hours and then dissolved in 20 mL of chloroform, filtering off the insoluble salts. The filtrate was washed with basic water (K_2CO_3) and the organic phase evaporated to dryness giving the desired compound as a yellow solid, that was merged with the one obtained from the filtrate. Yield: 87%.

The hydrochloride salt of the compound can be easily obtained: to a solution of the compound in ethanol, HCl 37% was slowly added to acid pH. The mixture was stirred at room temperature for 2 hours and then the solvent was evaporated.

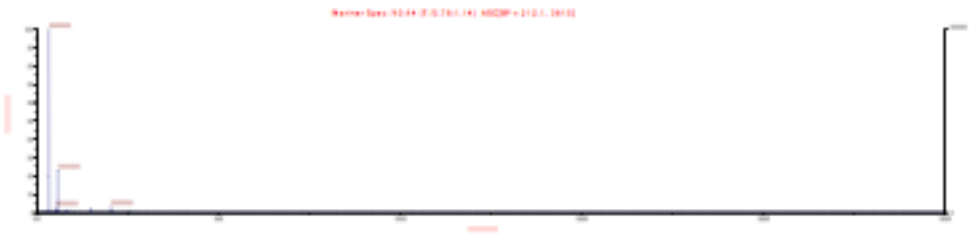
$C_{26}H_{22}N_4O_2$

MW 422.48

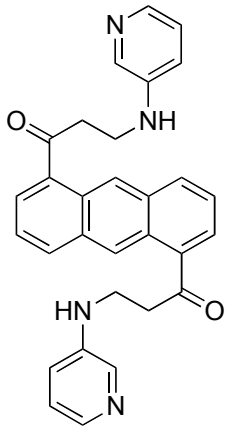
1H-NMR: δH (400 MHz, CDCl3) 9.42 (2H, s, ArH), 8.29 (2H, d, J 8.66 Hz, ArH), 7.99 (2H, dd, J 7.5 Hz, J 1.1 Hz, ArH), 7.68 (2H, s, ArH), 7.65 (2H, "t" J 8.36 Hz), 7.08 (2H, m, ArH), 7.06 (2H, m, ArH), 4.56 (4H, t, J 6.0 Hz, CH2), 3.64 (4H, t, J 6.0 Hz, CH2)

13C-NMR {1H}: δC (100 MHz, D2O) 200.87, 135.28, 135.19, 131.84, 131.53, 130.96, 125.89, 124.94, 123.99, 122.09, 119.70, 44.21, 40.01

HRMS (ESI): exp. 212.0958 ((M+2)/2) calc. 212.0944; exp. 423.1853 (M+1) calc. 423.1816



3-(3-Pyridylamino)-1-{5-[3-(3-pyridylamino)propionyl]-1-anthryl}-1-propanone (18)



A round-bottom flask was charged with **65** (50 mg, 0.112 mmol) and 30 mL of xylene. $CaCO_3$ (33 mg, 0.335 mmol), KI (55 mg, 0.335 mmol) and 3-amino pyridine (31 mg, 0.335 mmol) were then added. The reaction mixture was allowed to stir at reflux for 7 hours and the proceedings were monitored through TLC (chloroform, methanol 6:1). The obtained suspension was then filtered. The solid collected was dried in oven for 2 hours and then dissolved in 20 mL of chloroform, filtering off the insoluble salts. The filtrate was washed with basic water (K_2CO_3) and the organic phase evaporated to dryness giving the desired compound as a brown solid. Yield: 23%.

$C_{30}H_{26}N_4O_2$

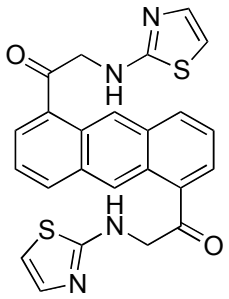
MW 474.55

1H-NMR: δH (400 MHz, CDCl3) 8.27 (2H, m, ArH), 8.10 (2H, m, ArH), 8,00 (2H, m, ArH), 7.54 (2H, m, ArH), 7.45 (2H, m, ArH), 7.21 (2H, m, ArH), 7.13 (2H, m, ArH), 6.99 (2H, m, ArH), 3.72 (4H, m, CH2), 3.50 (4H, m, CH2)

13C-NMR {1H}: δC (75 MHz, DMSO) 202.79, 145.17, 137.28, 136.73, 135.73, 134.33, 132.76, 124.07, 123.94, 120.03, 117.81

HRMS (ESI): exp. 238.1105 ((M+2)/2), calc. 238.1028; exp. 475.2239 (M+1), calc. 475.2056

2-(1,3-Thiazol-2-ylamino)-1-{5-[2-(1,3-thiazol-2-ylamino)acetyl]-1-anthryl}-1-ethanone (73)



A round-bottom flask was charged with **66** (50 mg, 0.119 mmol) and 30 mL of xylene. $CaCO_3$ (33 mg, 0.337 mmol), KI (55 mg, 0.335 mmol) and 2-amino thiazole (36 mg, 0.337 mmol) were then added. The reaction mixture was allowed to stir at reflux for 7 hours and the proceedings were monitored through TLC (chloroform, methanol 6:1). The obtained suspension was then filtered. The solid collected was dried in oven for 2 hours and then dissolved in 20 mL of chloroform, filtering off the insoluble salts. The filtrate was washed with basic water (K_2CO_3) and the organic phase evaporated to dryness. The solid was further purified through preparative TLC and yellow crystals were obtained. Yield: 20%.

$C_{24}H_{18}N_4O_2S_2$

MW 458.56

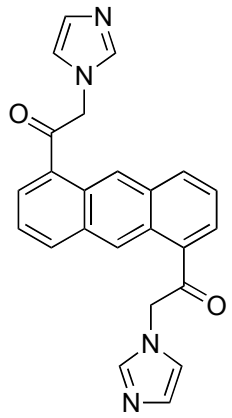
1H-NMR: δH (400 MHz, CDCl3) 9.24 (2H, s, ArH), 8.05 (2H, d, J 9.1 Hz, ArH), 7.78 (2H,dd, J 7.0 Hz J 1.1 Hz, ArH), 7.60 (2H, d, J 4.4 Hz, ArH), 7.52 (2H, m, ArH), 6.96 (2H, d, J 4.4 Hz ArH), 3.67 (4H, m, CH2)

13C-NMR {1H}: δC (75 MHz, CDCl3) 197.2, 160.8, 137.5, 136.5, 132.9, 131.3, 129.7, 126.6, 125.8, 121.4, 113.1, 50.3

HRMS (ESI): exp. 230.0536 ((M+2)/2), calc. 230.0456

2-(1H-Imidazol-1-yl)-1-{5-[2-(1H-imidazol-1-yl)acetyl]-1-anthryl}-1-ethanone

(74)



A round-bottom flask was charged with **66** (50 mg, 0.119 mmol) and 30 mL of xylene. $CaCO_3$ (33 mg, 0.337 mmol), KI (55 mg, 0.335 mmol) and imidazole (24 mg, 0.337 mmol) were then added. The reaction mixture was allowed to stir at reflux for 7 hours and the proceedings were monitored through TLC (chloroform, methanol 6:1). The obtained suspension was then filtered. The solid collected was dried in oven for 2 hours and then dissolved in 20 mL of chloroform, filtering off the insoluble salts. The filtrate was washed with basic water (K_2CO_3) and the organic phase evaporated to dryness. The residue was washed with a diluted ethanolic solution of HCl and then with aqueous NaOH. The final compound was obtained as a yellow solid. Yield: 53%.

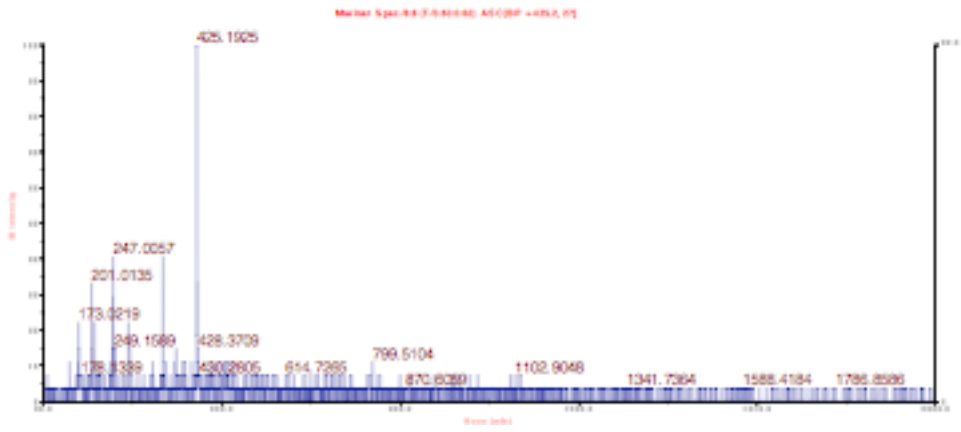
$C_{24}H_{19}N_4O_2$

MW 394.43

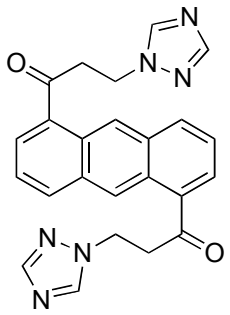
1H-NMR: δH (400 MHz, DMSO) 9.56 (2H, s, NHC=N), 8.94 (2H, s, ArH), 8.51 (4H, d, ArH), 7.80 (2H, dd, J 8.4 Hz J 7,2 Hz, ArH), 7.70 (4H, d, NCH=CH), 6.14 (4H, s, CH2)

13C-NMR {1H}: δC (75 MHz, DMSO) 196.71, 185.54, 135.38, 132.99, 131.69, 131.05, 127.48, 126.32, 125.32, 121.34, 54.85, 31.03

HRMS (ESI): exp. 198.0840 ((M+2)/2), calc. 198,.0715; exp. 395.1717 (M+1), calc. 395.1430



3-(1H-1,2,4-Triazol-1-yl)-1-{5-[3-(1H-1,2,4-triazol-1-yl)propionyl]-1-anthryl}-1-propanone (75)



A round-bottom flask was charged with **65** (75 mg, 0.167 mmol) and 30 mL of xylene. $CaCO_3$ (50 mg, 0.502 mmol), KI (83 mg, 0.502 mmol) and triazole (35 mg, 0.502 mmol) were then added. The reaction mixture was allowed to stir at reflux for 7 hours and the proceedings were monitored through TLC (chloroform, methanol 6:1). The obtained suspension was then filtered. Once cooled in an ice bath, the filtrate gave yellow crystals of the final compound. Yield: 24%.

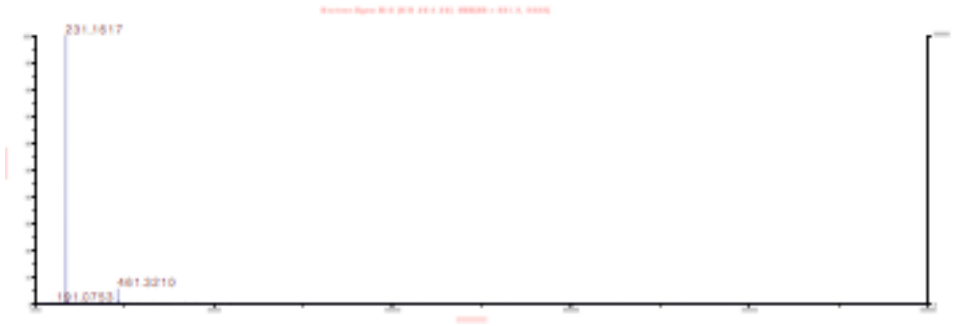
$C_{24}H_{20}N_6O_2$

MW 424.45

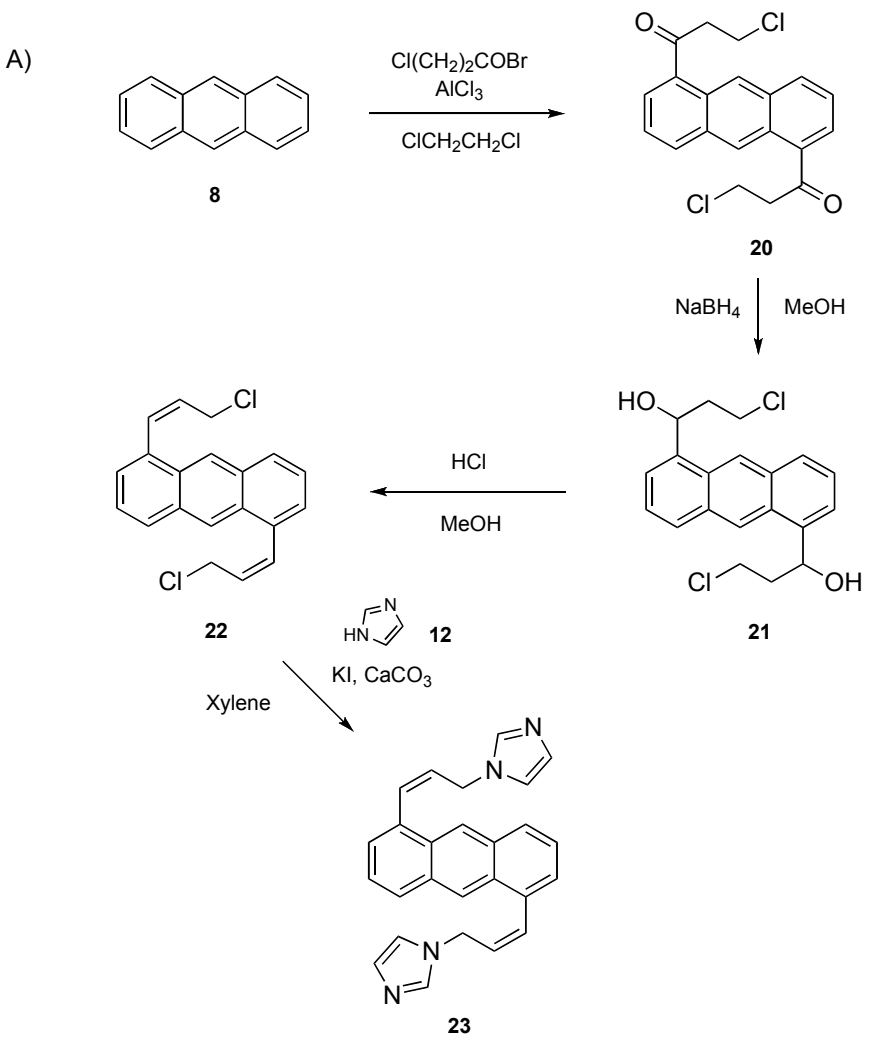
1H-NMR: δH (400 MHz, DMSO) 9.41 (2H, m, ArH), 9.32 (2H, ‘s’, ArH), 7.67 (2H, m, ArH), 7.22 (2H, ‘s’, ArH), 7.10 (2H, m, ArH), 6.97 (2H, m, ArH), 4.68 (4H, m, CH2), 3.88 (4H, m, CH2)

13C-NMR {1H}: δC (75 MHz, DMSO) 199.2, 151.5, 144.1, 136.7, 131.9, 131.3, 131.0, 126.5, 125.4, 121.5, 49.4, 39.5

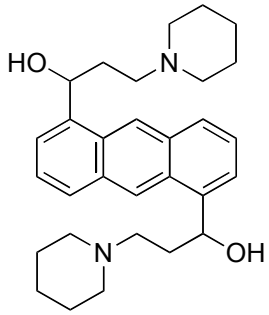
HRMS (ESI): exp. 425.1925 (M+1), calc. 425.1648



3.6 SCHEME II: PROCEDURES



1-[5-(1-Hydroxy-3-piperidinopropyl)-1-anthryl]-3-piperidino-1-propanol (76)



The keto- derivative **67** (11 mg, 0.024 mmol) was dissolved in 2 mL of methanol in a round-bottom flask and $NaBH_4$ was added to the solution (4 mg, 0.096). The reaction mixture was allowed to stir overnight at room temperature. The proceedings of the reaction were monitored through TLC (chloroform, methanol 6:1). After the end of the reaction, HCl 10% was added to acid pH. The reaction mixture was extracted with chloroform and the organic phase was washed with basic water. Evaporation of the organic phase gave the product as brown solid. Yield: 72%.

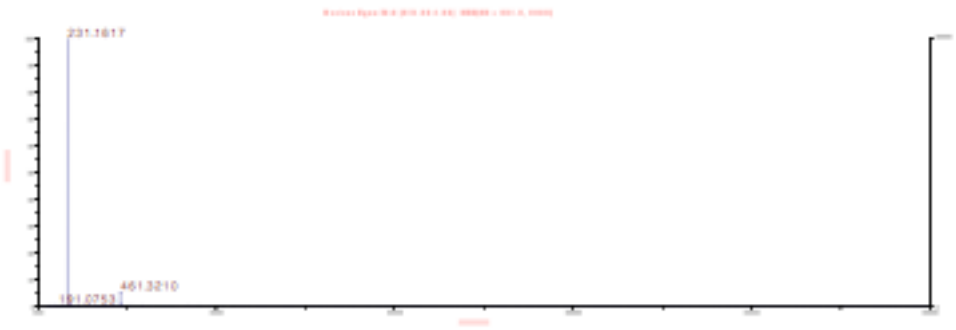
$C_{30}H_{40}N_2O_2$

MW 460.65

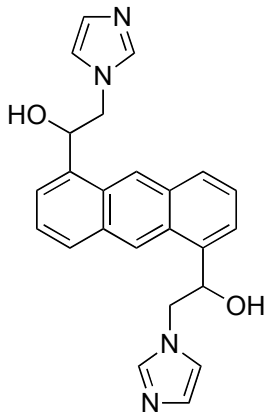
1H-NMR: δH (400 MHz, CDCl3) 8.57 (2H, d, J 5.4 Hz, ArH), 7.95 (2H, d, J 8.2 Hz, ArH), 7.77 (2H, d, J 6.5 Hz, ArH), 7.51 (2H, dd, J 8.4 Hz, J 7 Hz, ArH), 2.79 (4H, m, CH2), 2.67 (8H, m, CH2), 2.25 (4H, m, CH2), 2.08 (4H, m, CH2), 1.55 (8H, m, CH2)

13C-NMR {1H}: δC (75 MHz, CDCl3) 199.1, 136.9, 133.9, 132.3, 131.4, 126.9, 126.0, 122.1, 51.4, 50.0, 35.2, 26.3, 24.1

HRMS (ESI): exp. 231.1624 ((M+2)/2), calc. 231,1545; exp. 461.3203 (M+1), calc. 461,3090



1-{5-[1-Hydroxy-2-(1H-imidazol-1-yl)ethyl]-1-anthryl}-2-(1H-imidazol-1-yl)-1-ethanol (77)



The keto- derivative **74** (11 mg, 0.028 mmol) was dissolved in 2 mL of methanol in a round-bottom flask and $NaBH_4$ was added to the solution (4 mg, 0.096). The reaction mixture was allowed to stir overnight at room temperature. The proceedings of the reaction were monitored through TLC (chloroform, methanol 6:1). After the end of the reaction, HCl 10% was added to acid pH. The yellow solid obtained was collected by filtration, dissolved in chloroform and the solution was then washed with basic water (K_2CO_3). The organic phase was evaporated to give the final compound as a dark yellow solid. Yield: 63%.

$C_{24}H_{22}N_4O_2$

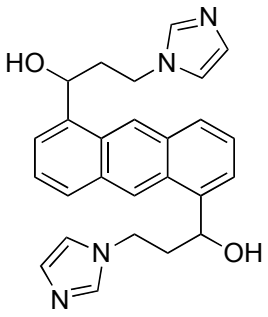
MW 398.46

1H-NMR: δH (400 MHz, DMSO) 8.39 (2H, m, ArH), 8.32 (2H, s, ArH), 7.82 (2H, m, ArH), 7.65 (2H, m, ArH), 7.52 (2H, m, ArH), 7.33 (2H, m, ArH), 7.23 (2H, m, ArH), 5.75 (2H, m, CHOH), 3.77 (4H, m, CH2)

13C-NMR {1H}: δC (75 MHz, DMSO) 140.9, 18.8, 136.5, 132.9, 130.7, 129.2, 126.4, 126.0, 121.5, 120.8, 66.3, 55.2

HRMS (ESI): exp. 200.1031 ((M+2)/2), calc. 200.0872; exp. 399.1956 (M+1), calc. 399.1743

1-{5-[1-Hydroxy-3-(1H-imidazol-1-yl)propyl]-1-anthryl}-3-(1H-imidazol-1-yl)-1-propanol (78)



The keto- derivative **19** (11 mg, 0.022 mmol) was dissolved in 2 mL of methanol in a round-bottom flask and $NaBH_4$ was added to the solution (3 mg, 0.089). The reaction mixture was allowed to stir overnight at room temperature. The proceedings of the reaction were monitored through TLC (chloroform, methanol 6:1). After the end of the reaction, HCl 10% was added to acid pH. The brown solid obtained was collected by centrifugation and then dissolved in water. The addition of NaOH 1M gave the product as a brown solid (free base). Yield: 43%.

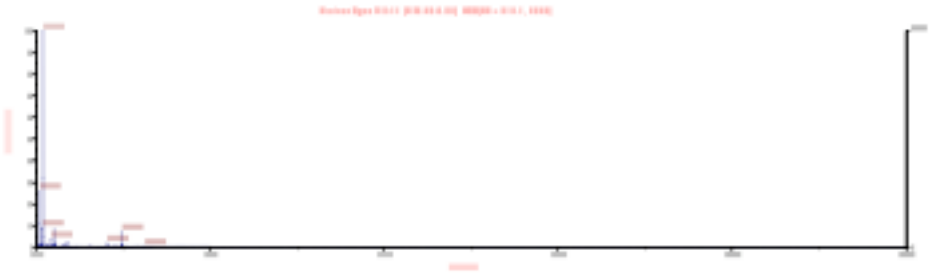
$C_{26}H_{26}N_4O_2$

MW 426.51

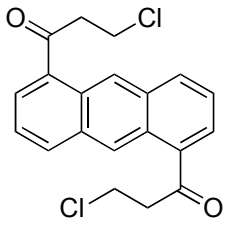
1H-NMR: δH (400 MHz, MeOD) 8.17 (2H, s, ArH), 7.86 (2H, m, ArH), 7.81 (2H, m, ArH), 7.71 (2H, d, J 7.0 Hz, ArH), 7.49 (2H, dd, J 7.0 Hz, 1.6 Hz, ArH), 7.30 (2H, m, ArH), 7.15 (2H, s, ArH), 5.5-5.4 (2H, m, CH), 4.5-4.4 (2H, m, CH-H), 4.3-4.2 (2H, m, CH-H) 2.5-2.4 (2H, m, CH-H), 2.3-2.2 (2H, m, CH-H)

13C-NMR {1H}: δC (100 MHz, MeOD) 138.92, 133.32, 129.47, 129.33, 128.92, 126.13, 123.30, 123.20, 121.15, 67.87, 45.21, 41.10

HRMS (ESI): exp. 2145.1125 ((M+2)/2) calc. 214.1101; exp. 427.2290 (M+1) calc. 427.2129



3-Chloro-1-[5-(3-chloropropionyl)-1-anthryl]-1-propanone (20)



A three-neck round-bottom flask was charged with $AlCl_3$ (23 g, 168 mmol) in 150 mL of dichloroethane. The stirred suspension was cooled in an ice bath and 3-Cl propionyl chloride (22 mL, 196 mmol) was added. Anthracene (10 g, 56 mmol) was added in portions avoiding the temperature to go over 0°C. After the addition of the reactants, the red mixture was allowed to stir for 4 hours at room temperature. The proceeding of the reaction was monitored through TLC (dichloromethane, ethyl acetate 98:2). The red solid collected from filtration was poured in a mixture of 37% HCl (40.0 mL) and ice (400 g) and let under vigorous stirring for additional 3 hours. A yellow solid was then collected by filtration, which was refluxed in 100 mL of acetic acid for 2 hours. The non dissolved solid was then dissolved in chloroform and the organic solution was washed with basic water (K_2CO_3). The organic phase was then evaporated to give the product. Yield: 35%.

$C_{20}H_{16}Cl_2O_2$

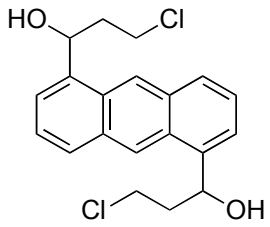
MW: 359.25

1H-NMR: δH (400 MHz, CDCl3) 9.56 (2H, s, ArH), 8.30 (2H, m, ArH), 8.09 (2H, dd, J 7.0 Hz, J 0.9 Hz, ArH), 7.57 (2H, dd, J 8.4 Hz, J 7.0 Hz, ArH), 4.65 (4H, t, J 6.2 Hz, CH2), 3.52 (4H, t, J 6.2 Hz, CH2)

13C-NMR {1H}: δC (75 MHz, CDCl3) 199.2, 136.5, 132.9, 132.2, 131.0, 126.7, 125.9, 121.2, 40.1, 38.7

HRMS (ESI): exp. 359.0611 (M+1), calc. 359.0535

3-Chloro-1-[5-(3-chloro-1-hydroxypropyl)-1-anthryl]-1-propanol (21)



The keto- derivative **20** (11 mg, 0.031 mmol) was dissolved in 2 mL of methanol in a round-bottom flask and $NaBH_4$ was added to the solution (5 mg, 0.122). The reaction mixture was allowed to stir overnight at room temperature. The proceedings of the reaction were monitored through TLC (chloroform, methanol 6:1). After the end of the reaction, HCl 10% was added to acid pH. The reaction mixture was extracted with chloroform and the organic phase was washed with basic water. Evaporation of the organic phase gave the product. Yield: 72%.

$C_{20}H_{20}Cl_2O_2$

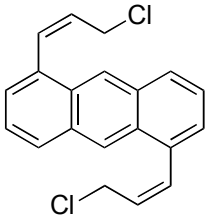
MW 363.28

1H-NMR: δH (400 MHz, MeOD) 8.81 (2H, s, ArH), 8.02 (2H, d, J 8.5 Hz, ArH), 7.69 (2H, d, J 7.0 Hz, ArH), 7.49 (2H, m, ArH), 5.81 (2H, dd, J 8.7 Hz J 3.3 Hz, CHOH), 3.95 (2H, m, HCH), 3.82 (2H, m, HCH), 2.27 (2H, m, HCH), 2.09 (2H, m, HCH)

13C-NMR {1H}: δC (75 MHz, MeOD) 140.7, 136.2, 123.9, 130.8, 126.4, 126.2, 121.5, 70.0, 41.7, 41.4

HRMS (ESI): exp. 363.0921 (M+1), calc. 363.0848

(Z)-1-{5-[(Z)-3-Chloro-1-propenyl]-1-anthryl}-3-chloro-1-propene (22)



A round-bottom flask was charged with the hydroxyl derivative **21** (130 mg, 0.358 mmol) in 50 mL of methanol. The mixture was allowed to stir at room temperature and 2 mL of HCl 37% were added. After the addition the reaction was heated at 60°C for 5 hours and then stirred at room temperature overnight. The solvent was then evaporated and the residue purified by flash chromatography (chloroform, methanol 95:5). Yield: 52%.

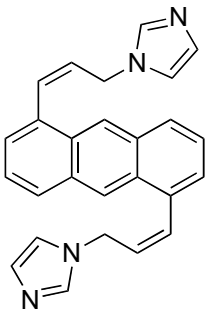
$C_{20}H_{16}Cl_2$

MW 327.25

1H-NMR: δH (400 MHz, CDCl3) 8.52 (2H, s, ArH), 7.91 (2H, d J 8.2 Hz, ArH), 7.61 (2H, dd J 8.5 Hz, J 8.2 Hz, ArH), 7.55 (2H, d J 8.5 Hz, ArH), 6.51 (2H, m J 9.1 Hz, C=CH), 6.34 (2 H, m J 9.1 Hz, C=CH), 3.92 (4H, "d", CH2)

13C-NMR {1H}: δC (100 MHz, CDCl3) 137.9, 133.4, 133.0, 131.5, 130.7, 127.9, 124.0, 123.6, 121.5, 45.2

(Z)-1-{5-[(Z)-3-(1H-Imidazol-1-yl)-1-propenyl]-1-anthryl}-3-(1H-imidazol-1-yl)-1-propene (23)



A round-bottom flask was charged with the anthracene-based starting material **22** (2 mg, 0.001 mmol) and 5 mL of xylene. $CaCO_3$ (3 mg, 0.003 mmol), KI (5 mg, 0.003 mmol) and imidazole (2 mg, 0.003 mmol) were then added. The reaction mixture was allowed to stir at reflux for 7 hours and the proceedings were monitored through TLC (chloroform, methanol 10:1). The obtained solution was extracted with chloroform and the organic phase was washed with aqueous NaOH. The organic phase was then evaporated to dryness to give the compound. Yield: 25%

$C_{26}H_{22}N_4$

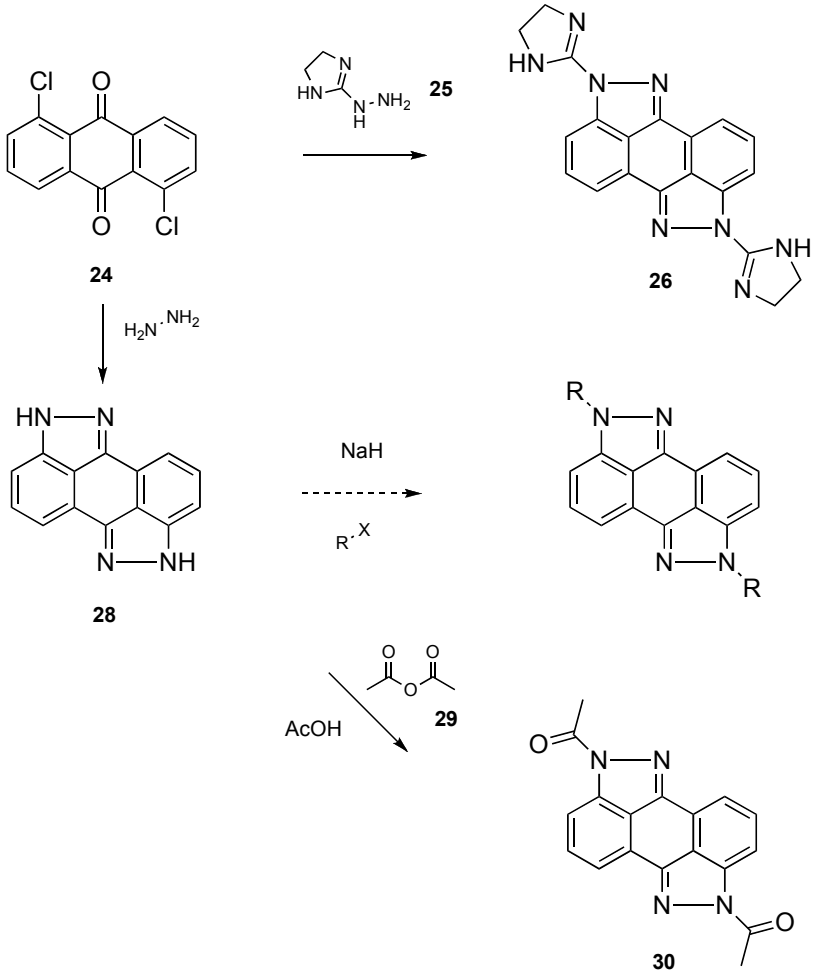
MW 390.48

1H-NMR: δH (400 MHz, MeOD) 8.76 (2H, s, ArH), 8.01 (2H, d J 8.6 Hz, ArH), 7.72 (2H, d J 7.2 Hz, NCH), 7.60 (2H, dd J 8.6 Hz, J 8.1 Hz, ArH), 7.50 (2H, d J 7.2 Hz, NCH), 7.43 (2H, d J 8.1 Hz, ArH), 7.01 (2H, s, NCH), 6.48 (2H, m J 9.1 Hz, C=CH), 6.15 (2 H, m J 9.1 Hz, C=CH), 3.43 (4H, "d", CH2)

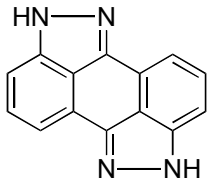
13C-NMR {1H}: δC (75 MHz, MeOD) 138.1, 136.5, 136.0, 132.9, 130.7, 130.5, 130.2, 127.4, 126.0, 121.5, 117.2, 112.0

HRMS (ESI): exp. 391.2064 (M+1), calc. 391.1852

3.7 SCHEME III: PROCEDURES



3.4.11.12-Tetrazapentacyclo[8.6.1.22,9.013,17]nonadeca-1(17),2,5,7,9(18),10,13,15-octaene (28)



A round-bottom flask was charged with 1,5-dichloro anthraquinone (**24**, 1 g, 3.583 mmol) and 25 mL of dimethylacetamide. After the addition of hydrazine (460 µL, 14.331 mmol) the solution was allowed to stir at reflux and monitored through TLC (dichloromethane, methanol 97.5:2.5). When the TLC showed no traces of starting material, the reaction mixture was poured into ice/water to give a dark solid. The purification of the compound was performed by flash chromatography (dichloromethane, methanol 97.5:2.5) followed by an additional preparative TLC (dichloromethane, methanol 95:5). Yield: 5%.

$C_{14}H_8N_4$

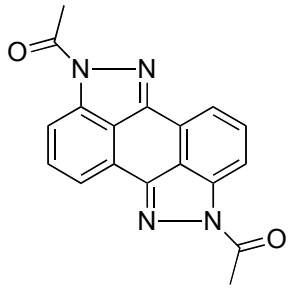
MW 232.24

1H-NMR: δH (400 MHz, Acetone) 7.60 (2H, dd, J 5.9 Hz J 1.7 Hz, ArH), 7.48 (2H, m, ArH)

13C-NMR {1H}: δC (100 MHz, Acetone) 143.8, 137.9, 130.6, 129.5, 127.2, 123.8, 113.9

HRMS (ESI): exp. 233.0887 (M+1), calc. 233.0700

1-{12-Acetyl-3.4.11.12-tetrazapentacyclo[8.6.1.22,5.013,17]nonadeca-1(17),2,5,7,9(18),10,13,15,18-nonaen-4-yl}-1-ethanone (30)



A round-bottom flask was charged with anthrapyrazole **28** (8 mg, 0.034 mmol) and 5 mL of acetic acid and acetic anhydride (10 µL, 0.103 mmol) was added to the solution. The mixture was allowed to stir at room temperature for 4 hours and then it was cooled in an ice bath for 2 additional hours. A yellow solid was then obtained and isolated by filtration. Yield: 46%.

$C_{18}H_{14}N_4O_2$

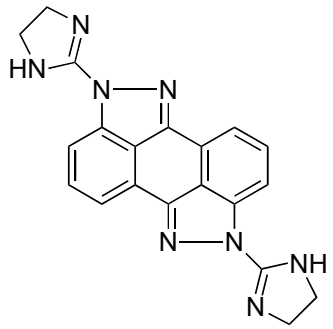
MW 318,33

1H-NMR: δH (400 MHz, CDCl3) 8.35 (2H, d, J 8.2 Hz, ArH), 8.02 (2H, d, J 7.4 Hz, ArH), 7.74 (2H, m, ArH), 2.90 (6H, s, CH3)

13C-NMR {1H}: δC (75 MHz, MeOD) 175.3, 153.4, 139.2, 130.7, 129.6, 127.2, 123.7, 113.8, 23.1

HRMS (ESI): exp. 316.2016 (M+1), calc. 316.1968

4,12-Bis(2-imidazolinyl)-3.4.11.12-tetrazapentacyclo[8.6.1.22,5.013,17]nonadeca-1(17),2,5,7,9(18),10,13,15,18-nonaene (26)



A round-bottom flask was charged with 1,5-dichloro anthraquinone (**24**, 100 mg, 0.365 mmol), 4,5-dihydro-1H-imidazol-2-yl hydrazine HBr (132 mg, 1.462 mmol) and K_2CO_3 (101 mg, 1.462 mmol) in 15 mL of pyridine. The mixture was refluxed for 10 h and then was stirred overnight at room temperature. The following day the solvent was removed under reduced pressure and the residue was purified by preparative TLC (n-hexan, ethyl acetate, triethylamine 49.5:49.5:1).

$C_{20}H_{16}N_8$

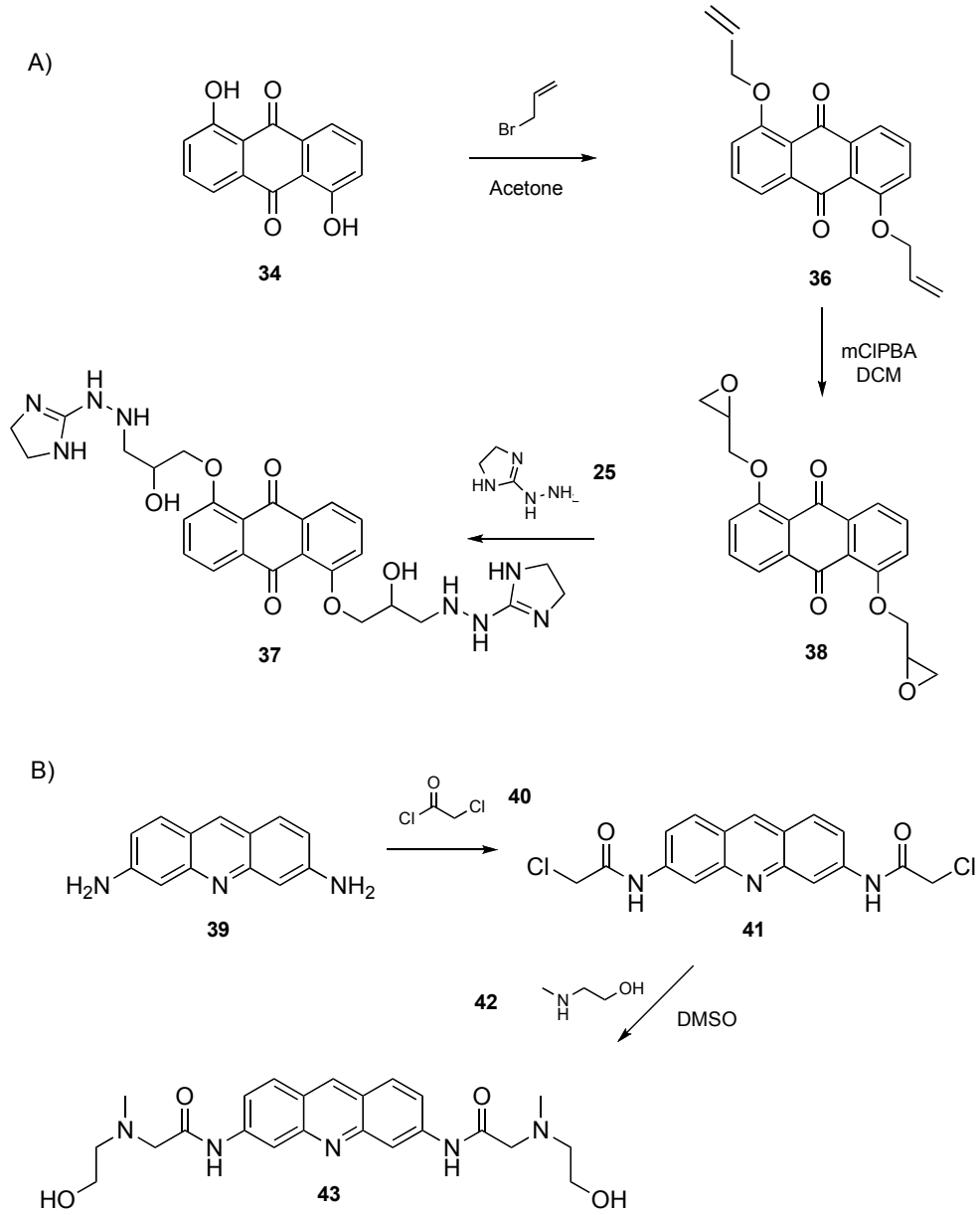
MW 368.39

1H-NMR: δH (400 MHz, DMSO) 8.05 (2H, m, ArH), 7.70 (4H, m, ArH), 3.59 (8H, s, 4x CH2)

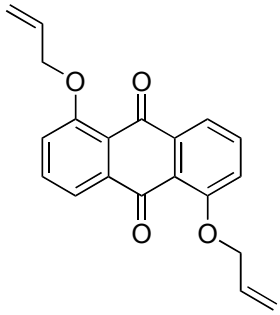
13C-NMR: δC (75 MHz, DMSO) 169.6, 160.6, 154.8, 135.0, 127.1, 120.0, 118.0, 111.5, 44.7

HRMS: exp.369.1400 (M+1), calc. 369.1498

3.8 SCHEME IV: PROCEDURES



1,5-Bis(allyloxy)-9,10-anthracenedione (36)



A round-bottom flask was charged with 1,5-bis-hydroxy anthraquinone (**34**, 500 mg, 2.08 mmol) and 50 mL of acetone. Subsequently, K_2CO_3 (1.725 g, 12.49 mmol) and KI (1.040 g, 6.24 mmol) were added to the solution. The suspension was allowed to stir at room temperature and the allylbromide (1.07 mL, 12.48 mmol) was added. The obtained reaction mixture was refluxed for 5 days and the proceedings of the reaction were monitored trough TLC (n-hexane, ethyl acetate 2:1). The reaction was then cooled to room temperature and the insoluble salts were filtered off. The liquid was washed with water and extracted with ethyl acetate. The evaporated organic phase gave the product as a solid. Yield: 46%.

$C_{20}H_{16}O_4$

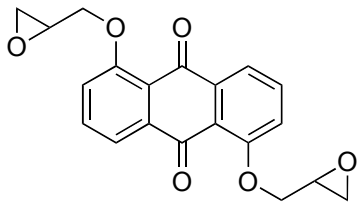
MW 320.33

1H-NMR: δH (400 MHz, DMSO) 7.94 (2H, dd J1 7,7 Hz J2 1,1 Hz, ArH), 7.68 (2H, “t”, J 8,1 Hz, ArH), 7.27 (2H, “d”, ArH), 6.16 (2H, m, =CH), 5.68 (2H, dd, J1 2,1 Hz J2 17,2 Hz, =CH2), 5.41 (2H, dd, J1 1,4Hz J2 10,4Hz, =CH2), 4.78 (4H, m, OCH2)

13C-NMR {1H}: δC (100 MHz, DMSO) 189.3, 158.9, 134.0, 132.9, 132.2, 119.8, 119.6, 118.0, 112.7, 69.2

HRMS (ESI): exp. 321,1123 (M+1) calc. 321,1082

1,5-Bis[(2-oxiranyl)methoxy]-9,10-anthracenedione (38)



A round-bottom flask was charged with **36** (300 mg, 0.850mmol) in 30 mL of dichloromethane. A solution of m-chloroperbenzoic acid (588 mg, 3.40 mmol) in dichloromethane was slowly added drop by drop to the reaction mixture. The reaction was monitored by TLC (n-hexane, ethyl acetate 2:1). After 4 hours another portion of m-chloroperbenzoic acid (588 mg, 3.40 mmol) was added and the mixture was allowed to stir at room temperature overnight. The reaction is quenched by the addition of a 20% aqueous solution of Na $NaHSO_3$ and then sodium carbonate. The organic phase, once evaporated, gave the solid product. Yield: 43%.

$C_{20}H_{16}O_6$

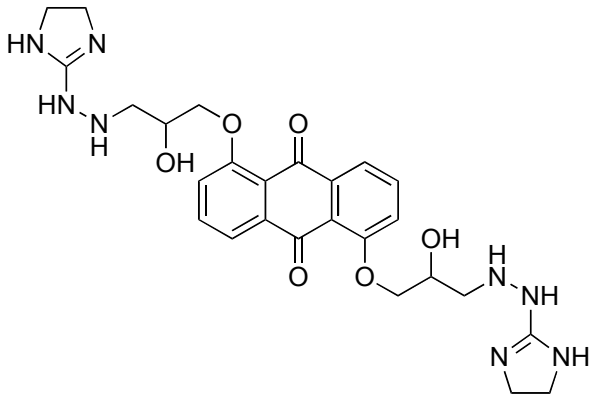
MW 352.3374

1H-NMR: δH (400 MHz, CDCl3) 7.95 (2H, dd, J1 7,9 Hz J2 1,1 Hz, ArH), 7.70 (2H, "t" J1 8,2 Hz, ArH), 7.31 (2H, dd, J1 8,4Hz J2 0,8Hz, ArH), 4.89 (2H, dd, J1 2,8 Hz J2 11,2 Hz, CH2O), 4,24 (2H, dd, J1 4,5 Hz J2 11,2 Hz, CH2), 3.50 (2H, m, CHO), 3.13 (2H, dd, J1 2,6 Hz J2 5,1 Hz, CH2O), 2,99 (2H, "t", CH2O)

13C-NMR: δC (75 MHz, CDCl3) 182.2, 158.1, 137.5, 121.7, 120.5, 118.9, 69.5, 53.4, 50.31

HRMS (ESI): exp. 353,1099 (M+1) calc. 353,0908

1,5-Bis{2-hydroxy-3-[2-(2-imidazolinyl)hydrazino]propoxy}-9,10-anthracenedione (37)



A round-bottom flask was charged with **38** (10 mg, 0.028 mmol) and 4,5-dihydro-1H-imidazol-2-yl hydrazine HBr (11 mg, 0.063 mmol) in 10 mL of ethanol. The mixture was refluxed overnight and then the solvent was evaporated. The residue was suspended in chloroform and the organic phase was washed with water, before being evaporated to give the final compound. Yield: 32%.

$C_{26}H_{32}N_8O_6$

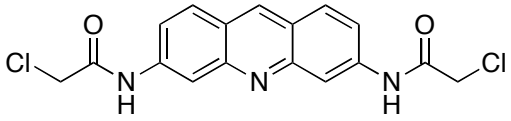
MW 552.58

1H-NMR: δH (400 MHz, DMSO) 7.95 (2H, m, ArH), 7.72 (4H, m, ArH), 7.55 (2H, m, ArH), 5.1-5.0 (6H, m, CH2, CH), 3.69 (8H, s, 4x CH2), 3.35 (4H, d, J 5.7 Hz, CH2)

13C-NMR: δC (75 MHz, DMSO) 189.5, 153.0, 146.6, 144.2, 141.4, 139.3, 131.3, 81.3, 66.8, 51.5, 19.2

HRMS (ESI): exp. 278.1847 ((M+2)/2), calc. 553.2445

2-Chloro-1-[6-(2-chloroacetylamino)-3-acridinylamino]-1-ethanone (41)



A round-bottom flask was charged with proflavine hydrochloride (**39**, 100 mg, 0.407 mmol) and the solid was suspended in chloroacetyl chloride (3 mL). The mixture was refluxed for 8 hours and then allowed to stand at room temperature overnight. The following day the mixture was refluxed for one additional hour and then, after being cooled to room temperature, the suspension was filtered. The brown solid collected is the desired product. Yield: 62%.

$C_{17}H_{13}Cl_2N_3O_2$

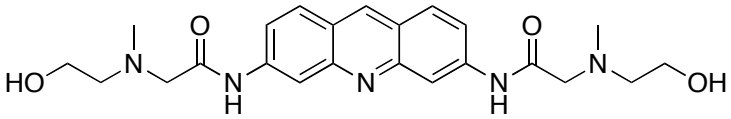
MW 362.21

1H-NMR: δH (400 MHz, DMSO) 8.51 (1H, s, ArH), 8.12 (2H, d J 8.4 Hz, ArH), 7.87 (2H, d J 8.1 Hz, ArH), 7.35 (2H, dd J 8.4 Hz J 8.1 Hz, ArH), 3.97 (4H, s, CH2)

13C-NMR {1H}: δC (100 MHz, DMSO) 164.9, 147.1, 140.2, 138.2, 135.9, 119.9, 117.3, 99.0, 43.2

HRMS: exp. 360.2017 (M-1), calc. 360.2021

2-[N-Methyl(2-hydroxyethyl)amino]-1-(6-{2-[N-methyl(2-hydroxyethyl)amino]acetylamino}-3-acridinylamino)-1-ethanone (43)



A reaction tube was charged with **41** (5 mg, 0.014 mmol), N-methyl aminoethanol (2.2 μL, 0.028 mmol) and triethylamine (6 μL, 0.042 mmol) in DMSO (140 μL). The reaction was incubated at 60°C for 6 h in a water bath. The proceeding of the reaction was checked by MS analysis until complete formation of the product.

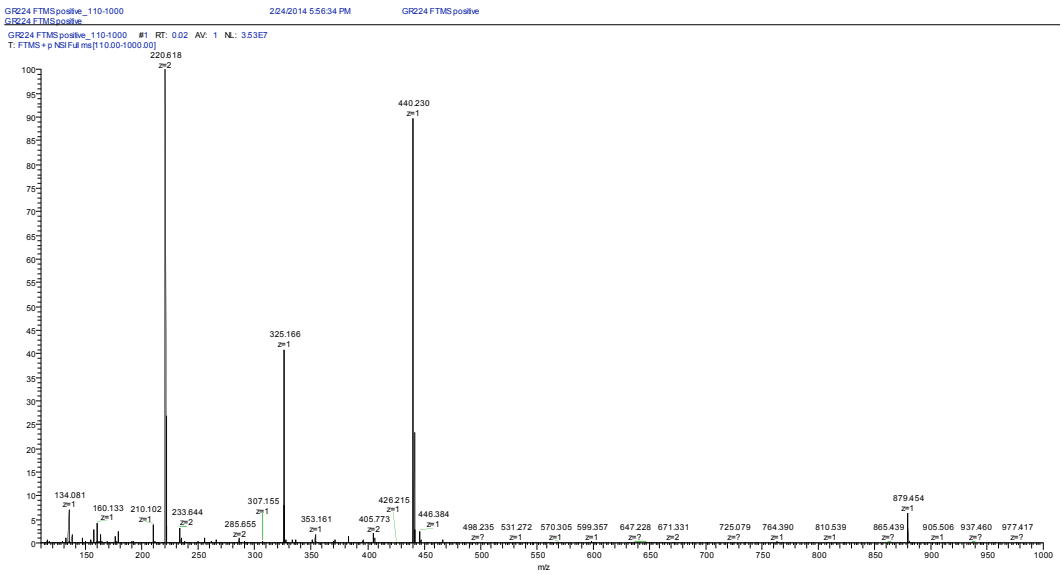
$C_{23}H_{29}N_5O_4$

MW 439.51

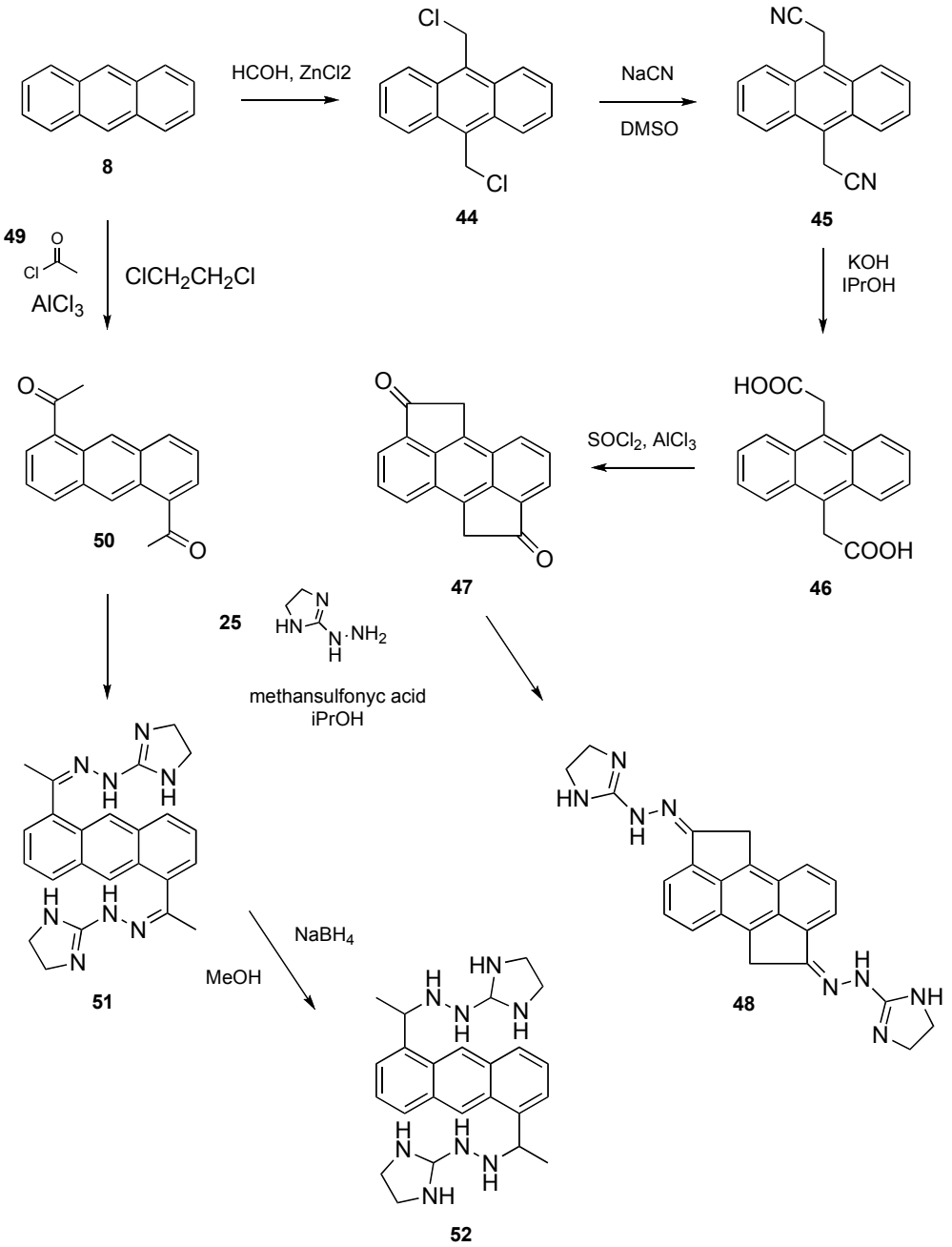
1H-NMR: δH (400 MHz, DMSO) 8.75 (1H, s, ArH), 8.01 (2H, d J 8.4 Hz, ArH), 7.92 (2H, d J 8.1 Hz, ArH), 7.32 (2H, dd J 8.4 Hz J 8.1 Hz, ArH), 3.58 (4H, s, CH2), 3.40 (4H, t, CH2), 2.82 (4H, t, CH2), 2.33 (6H, s, CH3)

13C-NMR {1H}: δC (100 MHz, DMSO) 163.1, 147.1, 140.2, 136.4, 135.9, 119.9, 117.9, 98.2, 59.3, 58.2, 49.8, 45.2

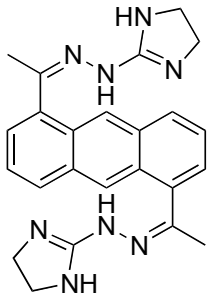
HRMS: exp. 220.6180 ((M+2)/2), calc. 220.6142; exp. 440.2271 (M+1), calc. 440.2298



3.9 SCHEME V: PROCEDURES



1-[2-(2-Imidazolinyl)hydrazono]-1-(5-{1-[2-(2-imidazolinyl)hydrazono]ethyl}-1-anthryl)ethane (51)



A round bottom flask was charged with **50** (100 mg, 0.381 mmol) and 4,5-dihydro-1H-imidazol-2-yl hydrazine HBr (205 mg, 1.145 mmol) in 19 mL of isopropanol with 1 mL of methansulfonic acid. The mixture was stirred at room temperature for twelve hours and then the yellow solid was filtered off. Flash chromatography was performed on this solid (isocratic, dichloromethane, methanol, triethylamine 70:29:1) gave light yellow crystals of the product. Yield: 25 %.

$C_{24}H_{26}N_8$

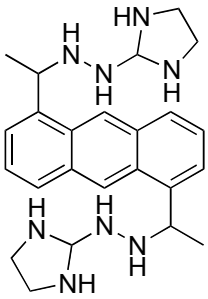
MW 426.52

1H-NMR: δH (400 MHz, DMSO) 8.76 (2H, s, ArH), 8.14 (2H, dd, J 5.6 Hz, ArH), 7.55 (2H, m, ArH), 3.56 (8H, s, CH2), 2.47 (6H, s, CH3)

13C-NMR {1H}: δC (100 MHz, DMSO): 153.4, 148.8, 147.9, 144.6, 144.1, 127.5, 126.6, 125.7, 113.3, 47.9, 42.3

HRMS (ESI): exp. 214.1260 ((M+2)/2), calc. 427.2280

1-[2-(2-Imidazolidinyl)hydrazino]-1-(5-{1-[2-(2-imidazolidinyl)hydrazino]ethyl}-1-anthryl)ethane (52)



Compound **51** (11 mg, 0.032 mmol) was dissolved in 2 mL of methanol in a round-bottom flask and $NaBH_4$ was added to the solution (4 mg, 0.096). The reaction mixture was allowed to stir overnight at room temperature. The proceedings of the reaction were monitored through TLC (chloroform, methanol 6:1). After the end of the reaction, HCl 10% was added to acid pH. The solvent was removed under reduced pressure and the residue was dissolved in dichloromethane and the organic solution was washed with basic water (K2CO3). The organic phase was evaporated to give the final compound as a dark yellow solid. Yield: 36%.

$C_{24}H_{34}N_8$

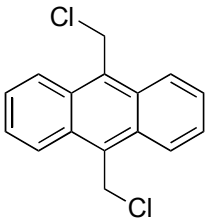
MW 434,58

1H-NMR: δH (400 MHz, DMSO) 8.82 (2H, s, ArH), 7.98 (2H, dd, ArH), 7.52 (2H, m, ArH), 3.53 (8H, s, CH2), 3.11 (4H, m, CH) 2.21 (6H, s, CH3)

13C-NMR {1H}: δC (100 MHz, DMSO): 136.4, 133.2, 130.7, 130.5, 126.4, 122.9, 121.4, 100.0, 50.2, 42.4, 20.8

HRMS (ESI): exp. 369.1099 (M+1) calc. 369.0908

9,10-Bis(chloromethyl)anthracene (44)



A round-bottom flask was charged with anthracene (1.78 g, 10 mmol), dry $ZnCl_2$ (1.64 g, 12 mmol), paraformaldehyde (1.50 g, 50 mmol) and dioxane (20 mL). HCl 37% was then slowly added to the mixture at room temperature. The reaction was allowed to stir at reflux for 3 h, and then allowed to stand for 16 h. The yellow solid was collected by filtration and washed with H_2O and dioxane to give the product, further purified by recrystallization from toluene. Yield: 36%

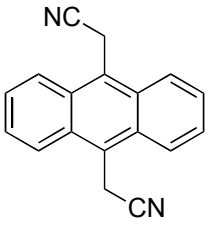
$C_{16}H_{12}Cl_2$

MW 275.17

1H-NMR: δH (300 MHz, DMSO) 8.51 (4H, q, J 3.4 Hz, ArH), 7.71 (4H, q, J 3.4 Hz, ArH), 5.87 (4H, s, CH2)

13C-NMR: δC (75 MHz, DMSO) 130.7, 129.2, 126.7, 124.7

[10-(Cyanomethyl)-9-anthryl]acetonitrile (45)



A round-bottom flask was charged with 9,10-bis(chloromethyl)anthracene (**44**, 1.8 g, 6.54 mmole) and sodium cyanide (1.6 g, 32.7 mmol) in DMSO and the mixture was heated to 50 °C for 3 h. The reaction was then cooled to room temperature and 200 mL of water were added. The resulting mixture was allowed to stir overnight at 4°C. The following day a solid precipitate was collected and washed with water to give the product as a yellow solid. Yield: 72%

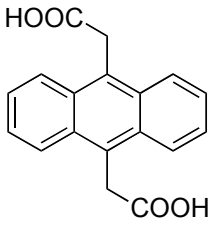
$C_{18}H_{12}N_2$

MW 256.30

1H-NMR: δH (300 MHz, DMSO) 8.52 (4H, q, J 3.3 Hz, ArH), 7.76 (4H, q, J 3.3 Hz, ArH), 5.01 (4H, s, CH2)

13C-NMR: δC (75 MHz, DMSO) 129.2, 126.9, 124.9, 124.6, 119.0, 15.9

[10-(Carboxymethyl)-9-anthryl]acetic acid (46)



A round-bottom flask was charged with **45** (1 g, 3.9 mmol) and KOH (2.2 g, 39 mmol) in butanol (30 mL). The mixture was refluxed for two days and, after cooling, most of the butanol was removed under reduced pressure. The residue was treated with HCl 37% and the precipitate was collected and washed with cold water. Yield: 32%.

$C_{18}H_{14}O_4$

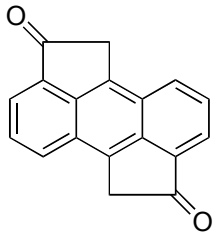
MW 294.30

1H-NMR: δH (300 MHz, DMSO) 12.5 (2H, s, COOH), 8.35 (4H, q, J 3.3 Hz, ArH), 7.59 (4H, q, J 3.3 Hz, ArH), 4.65 (4H, s, CH2)

13C-NMR: δC (75 MHz, DMSO) 172.7, 129.9, 127.6, 125.6, 125.3, 33.7

HRMS (ESI): exp. 292.0276 (M-1), calc. 292.0884

Pentacyclo[8.6.1.22,5.013,17]nonadeca-1(17),2(18),5,7,9,13,15-heptaene-4,12-dione (47)



A round-bottom flask was charged with **46** (0.88 g, 3 mmol) and $SOCl_2$ (10 mL). The mixture was refluxed for 4 h and then excess $SOCl_2$ was removed under reduced pressure. The crude acyl chloride was dissolved in 70 mL of dichloroethane and the solution was cooled in an ice bath. $AlCl_3$ (1.22 g, 9 mol) was added in three portions and the reaction was then stirred for 2 h at 0 °C and later refluxed for 20 min. The mixture was eventually cooled and poured into ice/water containing 100 mL of 4 N HC1. The solid obtained was collected to give the desired product. Yield: 60%.

$C_{18}H_{10}O_2$

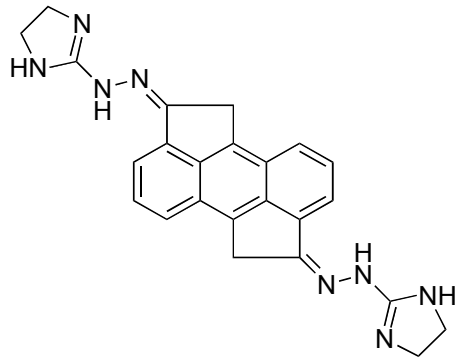
MW 258.27

1H-NMR: δH (400 MHz, DMSO) 8.40 (2H, d J 8.4, ArH), 8.03 (2H, d, J 6.7 Hz, ArH), 7.92 (2H, "t", J 6.8 Hz, ArH), 4.29 (2H, s, CH2)

13C-NMR {1H}: δC (100 MHz, DMSO): 196.8, 140.2, 130.8, 128.9, 126.4, 123.2, 121.0, 30.9

HRMS (ESI): calc. 258.0681 (M+1) , exp. 258.0679

4,12-Bis[2-(2-imidazolinyl)hydrazono]pentacyclo[8.6.1.22,5.013,17]nonadeca-1,5,7,9(18),10(17),13,15,18-octaene (48)



A round bottom flask was charged with **47** (40 mg, 0.152 mmol) and 4,5-dihydro-1H-imidazol-2-yl hydrazine HBr (82 mg, 0.458 mmol) in 9 mL of isopropanol with 1 mL of methansulfonic acid. The mixture was stirred at room temperature for twelve hours and then the yellow solid was filtered off. Flash chromatography was performed on this solid (isocratic, dichloromethane, methanol, triethylamine 70:29:1) gave the product as dark solid. Yield: 22 %.

$C_{24}H_{22}N_8$

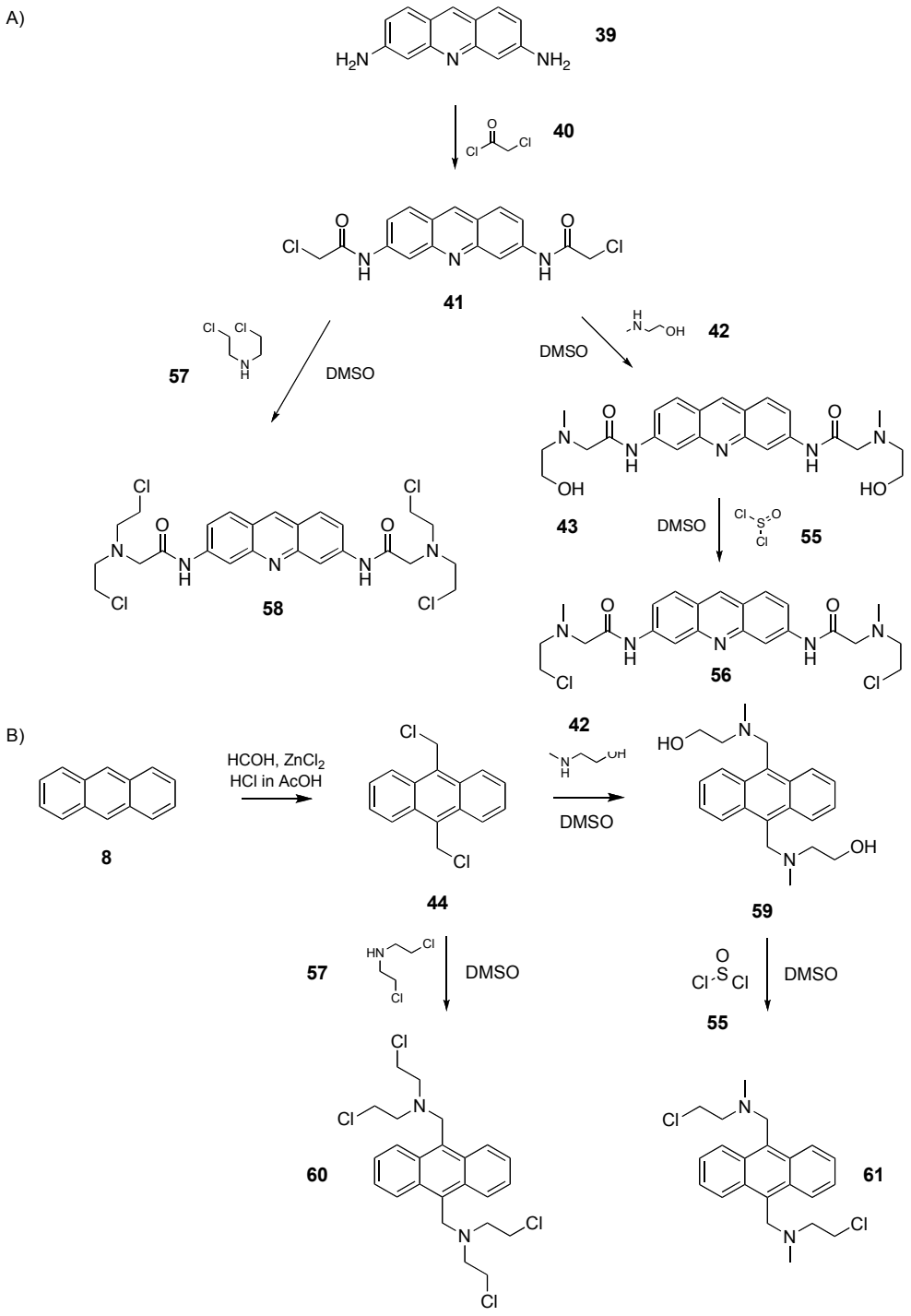
MW 422.49

1H-NMR: δH (400 MHz, DMSO) 7.22 (2H, d, J 5.1 Hz, ArH), 7.10 (2H, d, J 5.1 Hz, ArH), 6.96 (2H, m, ArH), 4.49 (4H, s, CH2), 3.83 (8H, s, CH2)

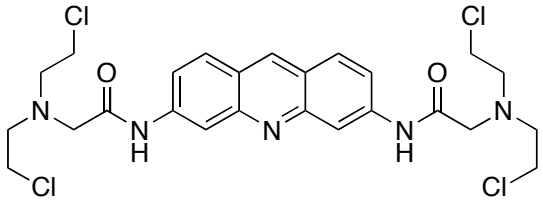
13C-NMR: δC (75 MHz, DMSO) 154.6, 150.4, 144.4, 137.6, 137.4, 126.7, 123.4, 111.1, 102.1, 68.1, 47.3

HRMS (ESI): exp. 212.1141 ((M+2)/2), calc. 423.1967

3.10 SCHEME VI: PROCEDURES



2-[Bis(2-chloroethyl)amino]-1-(6-{2-[bis(2-chloroethyl)amino]acetylamino}-3-acridinylamino)-1-ethanone (58)

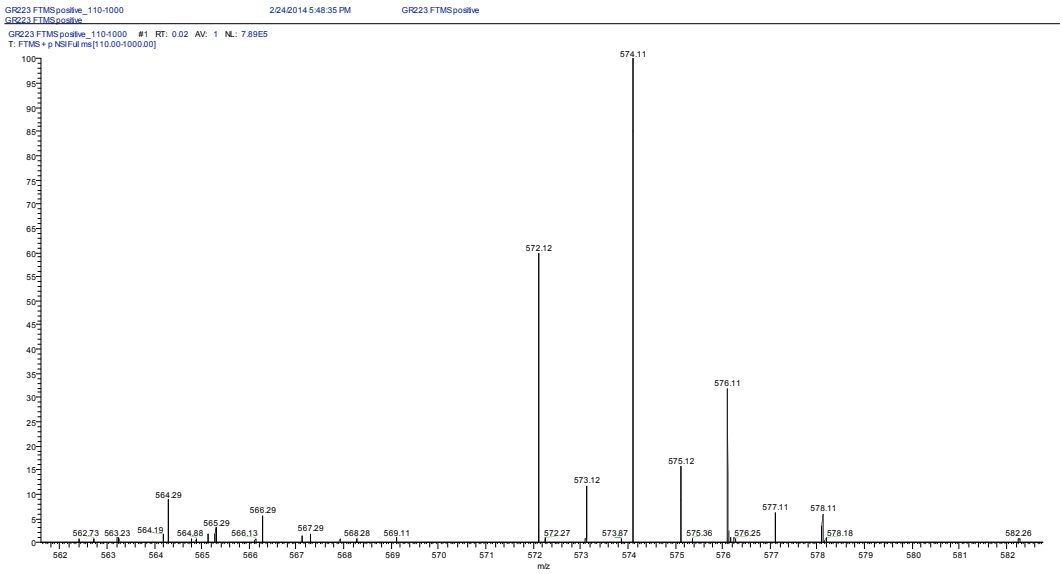


A reaction tube was charged with **41** (5 mg, 0.014 mmol), N,N-bis-chloroethylenamine hydrochloride (5 mg, 0.028 mmol) and triethylamine (6 μL, 0.042 mmol) in DMSO (140 μL). The reaction was incubated at 60°C for 6 h in a water bath. The proceeding of the reaction was checked by MS analysis until complete formation of the product.

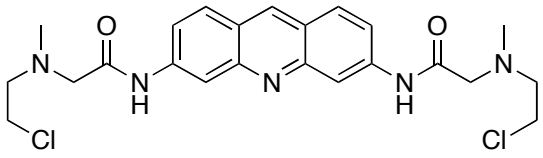
$C_{25}H_{29}Cl_2N_5O_2$

MW 573.34

HRMS: exp. 572.1168 (M-1), calc. 572.1175.



2-[N-Methyl(2-chloroethyl)amino]-1-(6-{2-[N-methyl(2-chloroethyl)amino]acetylamino}-3-acridinylamino)-1-ethanone (56)



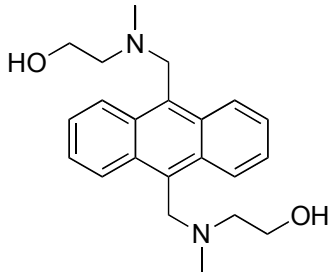
N,NI-2-(N-methylethanolamino)-acetyl-3,6-diamino acridine (**43**, 140 µL of a 100 mM DMSO solution, 0.028 mmol) was reacted with thionyl chloride (5 µL) in a reaction tube. The reaction mixture was incubated at 60°C for 6 h in a water bath. The proceeding of the reaction was checked by MS analysis until complete formation of the product.

$C_{23}H_{27}Cl_2N_5O_2$

MW 475.40

HRMS: exp. 474.1452 (M-1), calc. 474.1446.

2-{N-Methyl[(10-{[N-methyl(2-hydroxyethyl)amino]methyl}-9-anthryl)methyl]amino}ethanol (59)



Compound **44** (4 mg, 0.011 mmol), triethylamine (5 µL, 0.033 mmol), N-methylaminoethanol (2 µL 0.022 mmol) and DMSO (55 µL) were introduced in a reaction tube. The reaction mixture was incubated at 60°C for 6 h in a water bath. The proceeding of the reaction was checked by MS analysis until complete formation of the product.

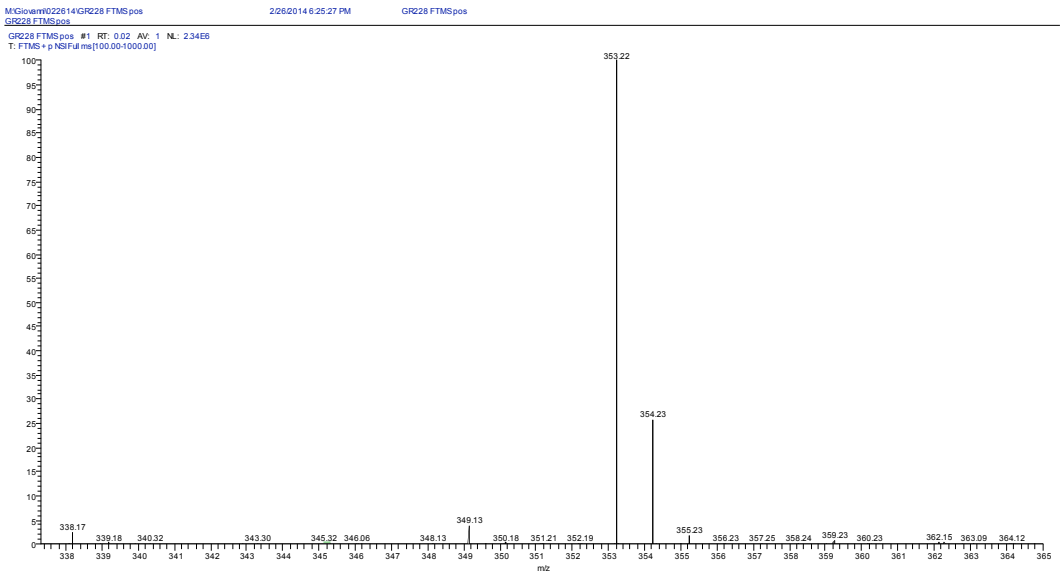
$C_{22}H_{28}N_2O_2$

MW 352.47

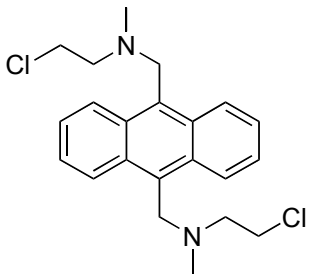
1H-NMR: δH (400 MHz, DMSO) 7.89 (2H, d, ArH), 7.60 (4H, m J 8.4 Hz, 4 x ArH), 3.92 (4H, s, CH2), 3.56 (4H, t, CH2), 2.87 (4H, t, CH2), 2.62 (6H, s, CH3)

13C-NMR {1H}: δC (100 MHz, DMSO): 133.2, 126.5, 125.9, 123.4, 59.8, 55.2, 39.7

HRMS: exp. 353.2175 (M+1), calc. 353.2168.



1-{N-Methyl[(10-{[N-methyl(2-chloroethyl)amino]methyl}-9-anthryl)methyl]amino}-2-chloroethane (61)

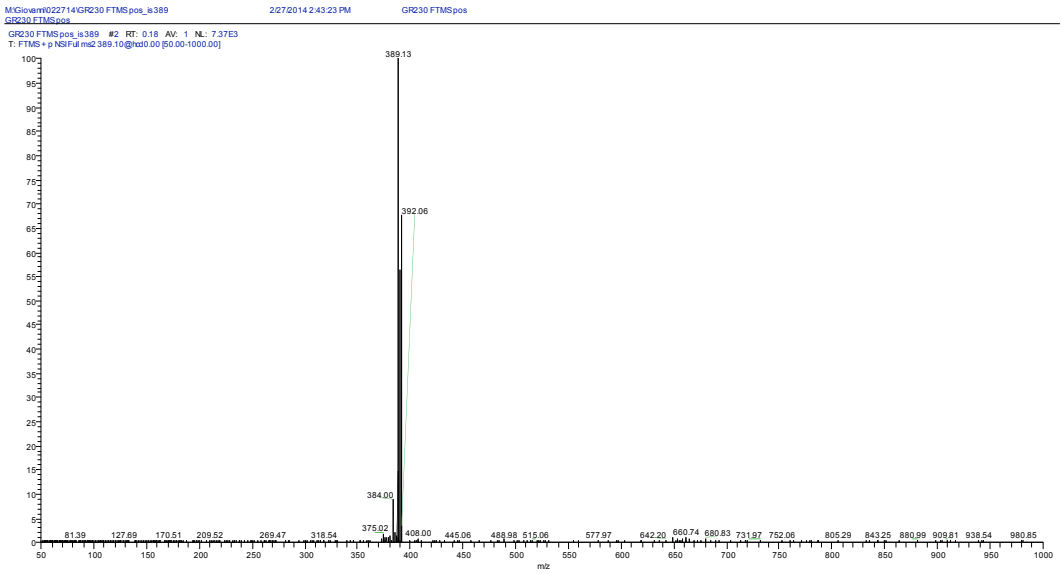


Compound **59** (55 µL of a 100 mM DMSO solution, 0.055 mmol) was reacted with thionyl chloride (5 µL) in a reaction tube. The reaction mixture was incubated at 60°C for 10 minutes and then allowed to stand at room temperature for 6 hours, followed by other 3 hours at 60°C in a water bath. The proceeding of the reaction was checked by MS analysis until complete formation of the product.

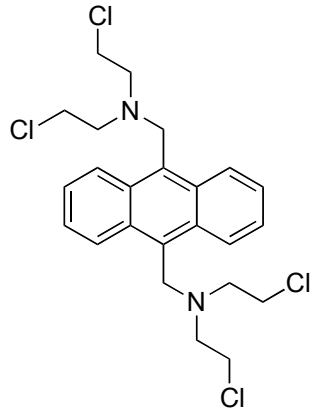
$C_{22}H_{26}Cl_2N_2$

MW 388.15

HRMS: exp. 389.1356 (M+1), calc. 389.1365



1-{[(10-{[Bis(2-chloroethyl)amino]methyl}-9-anthryl)methyl](2-chloroethyl)amino}-2-chloroethane (60)



A reaction tube was charged with 9,10-bis(chloromethyl)anthracene (3 mg, 0.011 mmol), N,N-bis-chloroethylenamine hydrochloride (4 mg, 0.022 mmol) and triethylamine (5 μL, 0.033 mmol) in DMSO (110 μL). The reaction was incubated at 60°C for 6 h in a water bath. The proceeding of the reaction was checked by MS analysis until complete formation of the product.

$C_{24}H_{28}Cl_4N_2$

MW 486.30

HRMS: exp. 484.0010 (M-1), calc. 483.0999.

3.10 BIOPHYSICAL EVALUATION

Fluorescence melting

The fluorescence melting assay evaluates the stability of a complex supermolecular structure according to the denaturation temperature of the structure itself. The efficacy in stabilization of the G-quadruplex by the compounds can be quantified from the variation of the melting temperature (Tm). Tm can be defined as the temperature showing 50 % of structured DNA and 50 % of denaturated DNA. At room temperature (structured DNA) 5' and 3' endings, respectively marked with the cromophore FAM (6-carboxy fluorescein) and the quencher Dabcyl, are close one to the other. Fluorescence observed in this condition is very low as long as Dabcyl has an absorbance maximum at the length of emission of FAM. On the other hand a destructured DNA shows a remarkable increase in fluorescence. Tm is measured at growing concentration of the compound and the ΔTm can be obtained. This experiment was performed in the lab of Prof. Sissi, University of Padova.

Other assay procedures

Fluorescence melting is adopted as a very preliminary screening procedure to identify potentially interesting stabilizers. Another technique is represented by Circular Dichroism (CD). According to the optical properties ("dichroism" literally means "double color"), this spectroscopy method reveals structural variations induced by the ligand. The TRAP assay, then, directly evaluates telomerase activity in the elongation of a DNA fragment. This sophisticated assay procedure are adopted in case of promising compounds individuated on the preliminary screening.

3.11 GEL ELECTROPHORESIS

Preparation of G-quadruplex stock solutions

Oligonucleotides were purchased from IDT, USA. The two sequences:

- GQm: 5' - AGG GTT AGG GTT AGG GTT AGG GT - 3'
- GQd: 5' - TAG GGT TAG GGT - 3'

were dissolved in water (200 μL) and allowed to stand in an ice bath for 2 hours. The obtained solutions were poured in two 3000 Da cut-off centrifuge tube and 4 mL of 150 mM ammonium acetate were added. The tubes were centrifuged for 1 hour at 4°C and the cleaning procedure was repeated another time with fresh ammonium acetate solution and for the third time with water. The final concentration of the cleaned solution obtained was measured with the UV-Vis spectrophotometer. The samples were then diluted to 100 μM with ammonium acetate and stored in a 4°C fridge for three weeks.

Preparation of buffers

10x TBE (Tris/Borate/EDTA) buffer was prepared dissolving the following species in 1 L of milliQ water:

- 108 g Tris (Tris hydroxymethyl aminoethane)
- 55 g boric acid
- 7.5 g EDTA (ethylendiamine tetraacetic acid) disodium salt.

TBE 1x gel running buffer was prepared by dilution of the concentrated buffer in 1:10 proportions. In some experiments also 60 mM ammonium acetate or 60 mM potassium chloride TBE 1x buffers were used and they were prepared from concentrated solutions

(600 mM) of respective salts and 10x TBE buffer by suitable dilution.

Preparation of native gels

10 mL of native gel for the electrophoresis experiments were freshly prepared before every experiment. The solution for a 20% polyacrylamide gel was obtained from these components:

- 5 mL AccuGel 19:1 40% w/v acrylamide/bis-acrylamide
- 1 mL TBE 10% buffer
- 1 mL 600 mM salt solution (ammoniun acetate or potassium chloride)
- 3 mL milliQ water
- 120 µL ammonium persulfate 10% solution
- 6 µL TEMED (tetramethyl ethylendiamine).

The solution for a 15% polyacrylamide gel was obtained from these components:

- 3.75 mL AccuGel 19:1 40% w/v acrylamide/bis-acrylamide
- 1 mL TBE 10% buffer
- 1 mL 600 mM salt solution (ammoniun acetate or potassium chloride)
- 4.25 mL milliQ water
- 120 µL ammonium persulfate 10% solution
- 6 µL TEMED (tetramethyl ethylendiamine).

The solutions were prepared in a 15 mL Falcon tube and rapidly poured into the Bio-Rad apparatus as long as the polymerization process starts to take place after the addition of APS and TEMED. 1 mm thick apparatus was used. The electrophoresis cell and the running buffer were conditioned before every usage and stored for one hour at the desired

temperature for the experiment before running the gels and a pre-run was carried out. The concentration of the DNA samples was checked every time just before the addition to the wells by UV measures. Native gels were run in 60 mM ammonium acetate or 60 mM potassium chloride TBE 1x buffers

Preparation of denaturing gels

10 mL of native gel for the electrophoresis experiments were freshly prepared before every experiment. The solution for a 20% polyacrylamide denaturing gel was obtained from these components:

- 8 mL UreaGel System Concentrated
- 1 mL UreaGel System Diluent
- 1 mL UreaGel System Buffer
- 120 μL ammonium persulfate 10% solution
- 6 μL TEMED (tetramethyl ethylendiamine).

The solutions were prepared in a 15 mL Falcon tube and rapidly poured into the Bio-Rad apparatus as long as the polymerization process starts to take place after the addition of APS and TEMED. 1 mm thick apparatus was used. The electrophoresis cell and the running buffer were conditioned before every usage and stored for one hour at the desired temperature for the experiment before running the gels and a pre-run was carried out. The concentration of the DNA samples was checked every time just before the addition to the wells by UV measures. Denaturing gels were run in TBE 1x buffer.

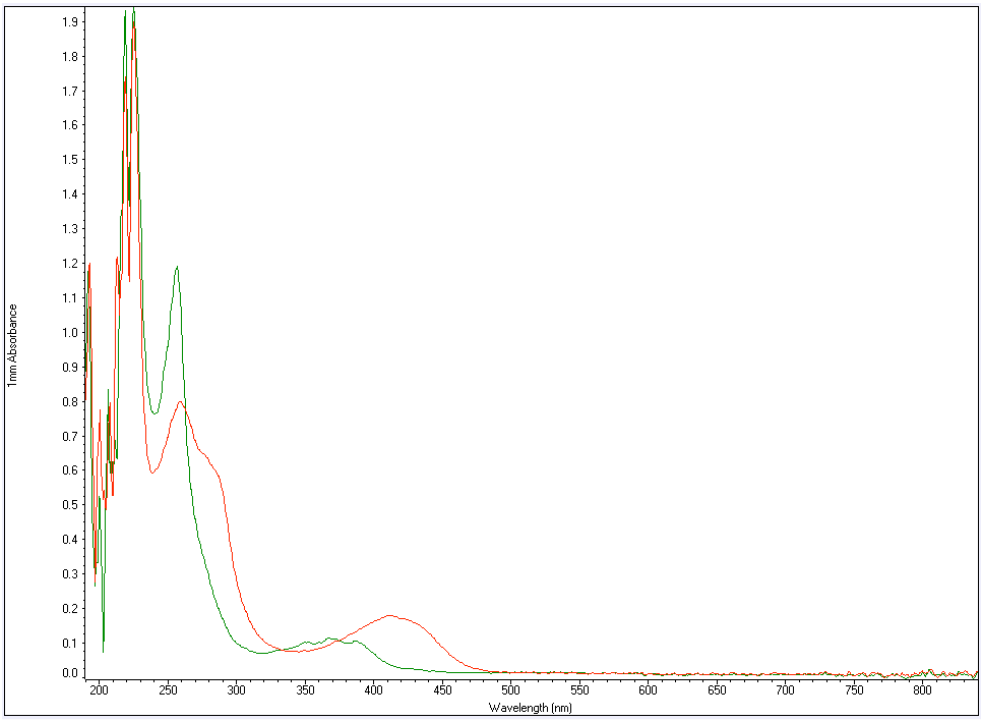
Staining and destaining

Gels were treated with a Sigma Stains All 1:4 solution for 45 minutes to achieve a complete

staining. Gels under treatment with the staining solution were covered and stored away from light. After recovering the staining solution the gels were washed with deionized water and allowed to stand in a little amount of deionized water under UV light.

3.12 UV-VIS SPECTROSCOPY

The concentration of DNA samples and of the stock solutions were analyzed in concentrations ranging from µM to mM. The DNA samples were quantified before every experiment. Stock solutions of synthesized compounds were prepared by dissolving in the calculated amount of solvent the weighted powder. The concentration of these samples was checked, when possible, by UV-vis analysis.



3.13 ELECTRONSPRAY IONIZATION MASS SPECTROMETRY

Preparation of samples

The preparation of G-quadruplex followed what already reported in the preparation of the samples for the gel electrophoresis experiments. Samples for the experiments were prepared just before the analysis from opportune dilutions of stock solutions. Solutions were sampled for the presence of G-quadruplexes after three weeks of storage at 4°C.

Tandem MS (MS-MS) experiments

Experiments were performed on a Thermo OrbiTrap LTQ Velos. Fragmentation experiments were performed using a 10 μM solution of the compound of interest.

ESI-MS of G-quadruplexes

Experiments were performed on a Thermo OrbiTrap LTQ Velos. Samples were prepared just before the analysis from opportune dilutions of stock solutions. The final concentration of the oligonucleotide was 3 μM in 150 mM ammonium acetate, with a ratio compound/oligo of 5:1 if a binding study was being performed.

Some of the G-quadruplex samples were also stored in a 150 mM solution of potassium acetate to promote the formation of a potassium including G-quadruplex. To achieve this, instead of following the desalting procedure described above, the oligonucleotide was treated as following: the oligonucleotide dissolved in water was stored in 150 mM potassium acetate and the slow annealing procedure was followed (3 min 100°C and then slow cooling to room temperature before putting the samples in a 4C fridge for a week

before the analysis). The samples were then desalted by washing with 150 mM triethylammonium acetate, following the same routine described for the samples above, using 3000 cut-off tubes. Even in this case, a diluition of the samples was performed before the MS analysis and the spraying buffer had a final concentration of 10 mM triethylammonium acetate.

3.14 ION MOBILITY MASS SPECTROMETRY

Preparation of samples

The preparation of G-quadruplex followed what already reported in the preparation of the samples for the gel electrophoresis experiments. Samples for the experiments were prepared just before the analysis from opportune dilutions of stock solutions.

Ion mobility experiment

Compounds were divided into groups of 3 or 4 molecules and then mixed together before the addition to the oligonucleotide solution in the described ratio. This experiment was designed in order to screen multiple compounds at the same time and evaluate a possible competition for the substrate, if any. The instrument collects both mass and mobility data at the same time and the spectra were acquired both as full scans and as isolation of single peaks. A peak fitting procedure was operated using a preformed algorithm to extrapolate mobility data in terms of collision cross sections for a single species. Experiments were run and recorded three times and an average value was calculated for every measure. Standard deviation was also calculated for every data set.

ACKNOWLEDGEMENTS

Un sincero ringraziamento al mio supervisore, Prof. Giuseppe Zagotto. Grazie a tutti i laureandi che ho avuto la fortuna di conoscere (Arianna, Francesca, Mariarita, Roberta, Jessica, Alberto M., Nicola, Marco, Chiara, Giulia, Gloria, Tamsin, Irina, Alina, Ana, Elisa, Lara, Raffaele, Laura, Alberto F.) ed ai dottorandi Riccardo, Valeria ed Elena.

Grazie al Prof. Daniele Fabris ed al Dott. Matteo Scalabrin per il sostegno durante il periodo di ricerca alla State University of New York.

Un grazie particolare alla mia famiglia.

I would like to sincerely thank my supervisor, Prof. Giuseppe Zagotto, the undergraduate students that I had the opportunity to encounter (Arianna, Francesca, Mariarita, Roberta, Jessica, Alberto M., Nicola, Marco, Chiara, Giulia, Gloria, Tamsin, Irina, Alina, Ana, Elisa, Lara, Raffaele, Laura, Alberto F.) and the PhD students Riccardo, Valeria and Elena.

I would also like to thank Prof. Daniele Fabris and Dr. Matteo Scalabrin for their support during the research activity carried out at the State University of New York.

I am especially and sincerely grateful to my family.

REFERENCES

1 Neidle S. Stephen Neidle on cancer therapy and G-quadruplex inhibitors. Interview by Joanna De Souza. Drug Discov Today. 2004 Sep 15;9(18):778-81.

2 Bryan TM, Baumann P. G-quadruplexes: from guanine gels to chemotherapeutics. Mol Biotechnol. 2011 Oct;49(2):198-208.

3 Sissi C, Gatto B, Palumbo M. The evolving world of protein-G-quadruplex recognition: a medicinal chemist's perspective. Biochimie. 2011 Aug;93(8):1219-30.

4 Folini M, Pivetta C, Zagotto G, De Marco C, Palumbo M, Zaffaroni N, Sissi C. Remarkable interference with telomeric function by a G-quadruplex selective bisantrene regioisomer. Biochem Pharmacol. 2010 Jun 15;79(12):1781-90.

5 Agrawal P, Hatzakis E, Guo K, Carver M, Yang D. Solution structure of the major G-quadruplex formed in the human VEGF promoter in K+: insights into loop interactions of the parallel G-quadruplexes. Nucleic Acids Res. 2013:41:10584-10592.

6 Huppert JL. Structure, location and interactions of G-quadruplexes. FEBS J. 2010 Sep;277(17):3452-8.

7 Rezler EM, Bears DJ, Hurley LH. DNA tetraplex-binding drugs. Annu. Rev. Pharmacol. 2003:43:359-379.

8 Cheung L, Schertzer M, Rose A, Landsdorp PM. Disruption of dog-1 in Caenorhabditis elegans triggers deletions upream in guanine-rich DNA. Nat. Genet. 2002:31:405-409.

9 Do NQ, Lim KW, Teo MH, Heddi B, Phan AT. Stacking of G-quadruplexes: NMR structure of a G-rich oligonucleotide with potential anti-HIV and anticancer activity. Nucleic Acids Res. 2011.

10 Este´JA, Cabrera C, Schols D, Cherepanov P, Gutierrez A, Witvrouw M, Pannecougue C, Debyser Z, Rando RF, Clotet B, Desmyter J, Clercg ED. Activity of different Bicyclamderivatives against Human Immunodefincency Virus on their interaction with the CXCR4 chemokine receptor. J. Mol. Pharmacol. 1998:53:340−345.

11 Wai LK. Telomeres, telomerase, and tumorigenesis-a review. MedGenMed. 2004 Jul

26;6(3):19.

12 Campisi J, d'Adda di Fagagna F. Cellular senescence: when bad things happen to good cells. Nat Rev Mol Cell Biol. 2007 Sep;8(9):729-40.

13 Pasqualino A, Panattoni GL. Anatomia umana. Citologia, istologia, embriologia. Utet Div. Scienze Mediche. 2009.

14 Watson JD et al. Molecular Biology of the Gene 7th Edition. Pearson. 2013.

15 Singleton MR, Scaife S, Wigley DB.Structural analysis of DNA replication fork reversal by RecG. Cell 2001:107:79

16 Beese LS, Derbyshire V, Steitz T. A Structure of DNA polymerase I Klenow fragment bound to duplex DNA.

17 Cong YS, Wright WE, Shay JW. Human telomerase and its regulation. Microbiol Mol Biol Rev. 2002 Sep;66(3):407-25.

18 Hayflick L, The limited in vitro lifetime of human diploid cell strains. Exp Cell Res. 1965 Mar;37:614-36.

19 Neidle S. Therapeutic application of quadruplex nucleic acids. Academic Pr. 2011

20 Skordalakes E. Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. Nat. Struct. Mol. Biol. 2010:17:513-518.

21 Collins K, Mitchell JR. Telomerase in the human organism. Oncogene. 2002 Jan 21;21(4):564-79.

22 Sekaran VG, Soares J, Jarstfer MB. Structures of telomerase subunits provide functional insights. Biochim Biophys Acta. 2010 May;1804(5):1190-201.

23 Kim NK, Zhang Q, Zhou J, Theimer CA, Peterson RD, Feigon J. Solution Structure and Dynamics of the Wild-type Pseudoknot of Human Telomerase RNA. J.Mol.Biol. 2008:384:1249-1261.

24 Zvereva MI, Shcherbakova DM, Dontsova OA. Telomerase: structure, functions, and activity regulation. Biochemistry (Mosc). 2010 Dec;75(13):1563-83.

25 Harkisheimer M, Mason M, Shuvaeva E, Skordalakes E. A Motif in the Vertebrate

Telomerase N-Terminal Linker of TERT Contributes to RNA Binding and Telomerase Activity and Processivity. Structure 2013:21:1870-1878.

26 Wyatt HD, West SC, Beattie TL. InTERTpreting telomerase structure and function. Nucleic Acids Res. 2010 Sep;38(17):5609-22.

27 Gillis AJ, Schuller AP, Skordalakes E. Structure of the Tribolium castaneum telomerase catalytic subunit TERT. Nature. 2008 Oct 2;455(7213):633-7.

28 Podlevsky JD, Chen JJ. It all comes together at the ends: telomerase structure, function, and biogenesis. Mutat Res. 2012 Feb 1;730(1-2):3-11.

29 Zvereva MI, Shcherbakova DM, Dontsova OA. Telomerase: structure, functions, and activity regulation. Biochemistry (Mosc). 2010 Dec;75(13):1563-83.

30 Steczkiewicz K, Zimmermann MT, Kurcinski M, Lewis BA, Dobbs D, Kloczkowski A, Jernigan RL, Kolinski A, Ginalski K. Human telomerase model shows the role of the TEN domain in advancing the double helix for the next polymerization step. Proc Natl Acad Sci U S A. 2011 Jun 7;108(23):9443-8.

31 Cifuentes-Rojas C, Shippen DE. Telomerase regulation. Mutat Res. 2012 Feb 1;730(1-2):20-7.

32 Mathad R, Hatzakis E, Yang D. c-MYC promoter G-quadruplex formed at the 5'-end of NHE III1 element: insights into biological relevance and parallel-stranded G-quadruplex stability. Nucleic Acids Res. 2011:39:9023-9033.

33 Mergny JL, Riou JF, Mailliet P, Teulade-Fichou MP, Gilson E. Natural and pharmacological regulation of telomerase. Nucleic Acids Res. 2002 Feb 15;30(4):839-65.

34 Günes C, Lichtsteiner S, Vasserot AP, Englert C. Expression of the hTERT gene is regulated at the level of transcriptional initiation and repressed by Mad1. Cancer Res. 2000 Apr 15;60(8):2116-21.

35 O'Sullivan RJ, Karlseder J. Telomeres: protecting chromosomes against genome instability. Nat Rev Mol Cell Biol. 2010 Mar;11(3):171-81.

36 Matulić M, Sopta M, Rubelj I. Telomere dynamics: the means to an end. Cell Prolif. 2007 Aug;40(4):462-74.

37 Rao T, Armstrong GS, Wuttke DS. Structure of Est3 reveals a bimodal surface with differential roles in telomere replication. Proc Natl Acad Sci USA. 2014:111:214-8

38 Verdun RE, Karlseder J. Replication and protection of telomeres. Nature. 2007 Jun 21;447(7147):924-31.

39 Denchi EL. Give me a break: how telomeres suppress the DNA damage response. DNA Repair (Amst). 2009 Sep 2;8(9):1118-26.

40 Smogorzewska A, de Lange T. Regulation of telomerase by telomeric proteins. Annu Rev Biochem. 2004;73:177-208.

41 Luu KN, Phan AT, Kuryavyi V, Lacroix L, Patel DJ. Structure of the human telomere in K+ solution: an intramolecular (3 + 1) G-quadruplex scaffold. J Am Chem Soc. 2006 Aug 2;128(30):9963-70.

42 Campbell NH, Smith DL, Reszka AP, Neidle S, O'Hagan D. Fluorine in medicinal chemistry: β-fluorination of peripheral pyrrolidines attached to acridine ligands affects their interactions with G-quadruplex DNA. Org Biomol Chem. 2011 Mar 7;9(5):1328-31.

43 Micco M, Collie GW, Dale AG, Ohnmacht SA, Pazitna I, Gunaratnam M, Reszka AP, Neidle S. Structure-based design and evaluation of naphthalene diimide G-quadruplex ligands as telomere targeting agents in pancreatic cancer cells. J Med Chem. 2013 Apr 11;56(7):2959-74.

44 Burge S, Parkinson GN, Hazel P, Todd AK, Neidle S. Quadruplex DNA: sequence, topology and structure. Nucleic Acids Res. 2006;34(19):5402-15.

45 Huppert JL. Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. Chem Soc Rev. 2008 Jul;37(7):1375-84.

46 Zhang N, Gorin A, Majumdar A, Kettani A, Chernichenko N, Skripkin E, Patel DJ. V-shaped scaffold: a new architectural motif identified in an A x (G x G x G x G) pentad-

containing dimeric DNA quadruplex involving stacked G(anti) x G(anti) x G(anti) x G(syn) tetrads. J Mol Biol. 2001 Aug 31;311(5):1063-79.

47 Haider S, Parkinson GN, Neidle S. Crystal structure of the potassium form of an Oxytricha nova G-quadruplex. J Mol Biol. 2002 Jul 5;320(2):189-200.

48 Zhou W, Brand NJ, Ying L. G-quadruplexes-novel mediators of gene function. J Cardiovasc Transl Res. 2011 Jun;4(3):256-70.

49 Dai J, Chen D, Carver M, Yang D. NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. Nucleic Acids Res. 2006:34:5133-5144.

50 Patel DJ, Phan AT, Kuryavyi V. Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. Nucleic Acids Res. 2007;35(22):7429-55.

51 Kettani A, Bouaziz S, Gorin A, Zhao H, Jones RA, Patel DJ. Solution structure of a Na cation stabilized DNA quadruplex containing G.G.G.G and G.C.G.C tetrads formed by G-G-G-C repeats observed in adeno-associated viral DNA. J Mol Biol. 1998 Sep 25;282(3):619-36.

52 Amodeo P, Castiglione Morelli MA, Ostuni A, Cristinziano P, Bavoso A. Structural Features of the C-Terminal Zinc Finger Domain of the HIV-2 Nc Protein (Residues 23-49). To be published, PDB ID: 2IWJ

53 Rajendran A, Endo M, Hidaka K, Thao Tran PL, Mergny JL, Gorelick RJ, Sugiyama H. HIV－1 Nucleocapsid Proteins as Molecular Chaperones for Tetramolecular Antiparallel G－Quadruplex Formation. J Am Chem Soc. 2013 135:18575-85

54 Read M, Harrison RJ, Romagnoli B et al . Structure-based design of selective and potent G-quadruplex-mediated telomerase inhibitors. Proc Natl Acad Sci USA 2001; 98: 4844-9

55 Perrone R, Butovskaya E, Daelemans D, Palu` G, Pannecouque C and Richter SN. Anti-HIV-1 activity of the G-quadruplex ligand BRACO-19. J Antimicrob Chemother.

2014 doi:10.1093/jac/dku280

56 Martidanata H, Phan AT. Structure of Propeller-Type Parallel-Stranded RNA G-Quadruplexes, Formed by Human Telomeric RNA Sequences in K+ Solution. J Am Chem Soc. 2009, 131:(7):2570-78

57 Xu Y, Kaminaga K, Komiyama M. G-quadruplex formation by human telomeric repeats-containing RNA in Na+ solution. J Am Chem Soc. 2008 Aug 20;130(33):11179-84.

58 Biffi G, Tannahill D, McCafferty J, Balasubramanian S. Quantitative visualization of DNA G-quadruplex structures in human cells. Nat Chem. 2013 Mar;5(3):182-6.

59 Bazzicalupi C, Ferraroni M, Bilia AR, Scheggi F, Gratteri P. The crystalstructure of human telomeric DNA complexed with berberine: an interesting case of stacked ligand to G-tetrad ratio higher than 1:1. Nucleic Acids Res. 2013 Jan 7;41(1):632-8.

60 Zimmermann S, Martens UM. Telomeres and telomerase as targets for cancer therapy. Cell Mol Life Sci. 2007 Apr;64(7-8):906-21.

61 Ou TM, Lu YJ, Tan JH, Huang ZS, Wong KY, Gu LQ. G-quadruplexes: targets in anticancer drug design. ChemMedChem. 2008 May;3(5):690-713.

62 Di Leva,FS, Zizza P, Cingolani C, D'Angelo C, Pagano B, Amato J, Salvati E, Sissi C, Pinato O, Marinelli L, Cavalli A, Cosconati S, Novellino E, Randazzo A, Biroccio A. Exploring the Chemical Space of G−Quadruplex Binders: Discovery of a Novel Chemotype Targeting the Human Telomeric Sequence. J. Med. Chem. 2013:56:9646−54

63 Perry PJ, Read MA, Davies RT, Gowan SM, Reszka AP, Wood AA, Kelland LR, Neidle S. 2,7-Disubstituted amidofluorenone derivatives as inhibitors of human telomerase. J Med Chem. 1999 Jul 15;42(14):2679-84.

64 Parkinson GN, Ghosh R, Neidle S. Structural basis for binding of porphyrin to human telomeres. Biochemistry. 2007 Mar 6;46(9):2390-7.

65 Collie GW, Promontorio R, Hampel SM, Micco M, Neidle S, Parkinson GN. Structural

basis for telomeric G-quadruplex targeting by naphthalene diimide ligands. J Am Chem Soc. 2012 Feb 8;134(5):2723-31.

66 Sissi C, Lucatello L, Paul Krapcho A, Maloney DJ, Boxer MB, Camarasa MV, Pezzoni G, Menta E, Palumbo M. Tri-, tetra- and heptacyclic perylene analogues as new potential antineoplastic agents based on DNA telomerase inhibition. Bioorg Med Chem. 2007 Jan 1;15(1):555-62.

67 Hurley LH, Wheelhouse RT, Sun D, Kerwin SM, Salazar M, Fedoroff OY, Han FX, Han H, Izbicka E, Von Hoff DD. G-quadruplexes as targets for drug design. Pharmacol Ther. 2000 Mar;85(3):141-58.

68 Chung WJ, Heddi B, Tera M, Iida K, Nagasawa K, Phan AT. Solution structure of an intramolecular (3 + 1) human telomeric G-quadruplex bound to a telomestatine derivative. J Am Chem Soc. 2013 Sep 11;135(36):13495-501

69 Alcaro S, Musetti C, Distinto S, Casatti M, Zagotto G, Artese A, Parrotta L, Moraca F, `Costa G, Ortuso F, Maccioni E, Sissi C. Identification and Characterization of New DNA G−Quadruplex Binders Selected by a Combination of Ligand and Structure-Based Virtual Screening Approaches. J. Med. Chem. 2013:56:843−55.

70 Fox MA, Whitesell JK. Organic Chemistry - Third Edition. 2004, Jones and Barlett.

71 Adams C, Earle MJ, Roberts G, Seddon KR. Friedel-Craft reactions in room temperature ionic liquids. Chemical communications. 1998.

72 Clayden J, Greeves N. Organic Chemistry. 2000. Oxford University Press.

73 Smith MB, March J. March's Advanced Organic Chemistry: reactions, mechanisms and structures - 6th Edition. 2007. Wiley Interscience.

74 US2010/0062460 A1

75 G. Capranico, F. Guano, S. Moro, G. Zagotto, C. Sissi, B. Gatto, F. Zunino, E. Menta and M. Palumbo, J. Biol. Chem., 1998, 273, 12732

76 Neuhaus D, Williamson MP. The Nuclear Overhauser Effect in Structural and Conformational Analysis. Wiley-VCH, second edition, 2000.

77 D. Ryu, E. Park, D.S. Kim, S. Yan, J.Y. Lee, B.Y. Chang and K.H. Ahn, , J. Am. Chem. Soc., 2008, 130, 2394

78] A.R. Mohebbi, C. Munoz and F. Wudl, Org. Lett., 2011, 13, 2560

79 Tan W, Zhou J and Yuan G. Electrospray ionization mass spectrometry probing of binding affinity of berbamine, a flexible cyclic alkaloid from traditional Chinese medicine, with G-quadruplex DNA. Rapid Commun. Mass Spectrom. 2014, 28, 143–7

80 Collie GW, Promontorio R, Hampel SM, Micco M, Neidle S, Parkinson GN. Structural basis for telomeric G-quadruplex targeting by naphthalene diimide ligands. J.Am.Chem.Soc. 2012, 134, 2723-2731

81 Collie GW, Promontorio R, Hampel SM, Micco M, Neidle S, Parkinson GN. Structural basis for telomeric G-quadruplex targeting by naphthalene diimide ligands. J.Am.Chem.Soc. 2012, 134, 2723-2731

82 Di Antonio M, McLuckie KIE, Balasubramanian S. Reprogramming the Mechanism of Action of Chlorambucil by Coupling to a G−Quadruplex Ligand. J. Am. Chem. Soc. 2014:136:5860-63

83 Yin LN, Liu QX, Wu XM, Cheng FG, Guo JH. 9,10-Bis(dichloroethylaminomethyl)anthracene. Acta Cryst. 2006:62:2510-11.

84 Wang GZ, Zhuang YY, Zhou CH. N,N-[(2,3,5,6-Tetramethyl-p-phenylene)dimethylene]bis[2-chloro-N-(2-chloroethyl)ethanamine]. Acta Cryst. 2009:65:1850

85 Evans, Mendez, Turner, Keating, Grimes, Melchoir, Szalai. End-stacking of copper cationic porphyrins on parallel-stranded guanine quadruplexes. J Biol Inorg Chem 12:2007:1235-49.

86 Rosu F, Gabelica V, Houssier C and De Pauw E. Determination of affinity, stoichiometry and sequence selectivity of minor groove binder complexes with double-stranded oligodeoxynucleotides by electrospray ionization mass spectrometry. Nucleic Acid Research 2002, 30, 16

87 D.E. Clemmer and M. F. Jarrold, J. Mass Spectrom., 1997, 32, 577

88 E. W. McDaniel, W. S. Barnes and D. W. Martin, Rev. Sci. Instrum., 1962, 33, 2.

89 Lopez A, Tarrago T, Vilaseca M, Giralt E. Applications and future of ion mobility mass spectrometry in structural biology. New J. Chem. 2013:37:1283-89

90 Ferreira, Marchland, Gabelica. Mass spectrometry and ion mobility spectrometry of G-quadruplexes. A study of solvent effects on dimer formation and structural transitions in the telomeric DNA sequence d(TAGGGTTAGGGT). Methods 57, 2012, 56-63

91 Monchaud, Allain, Bertrand, Smargiasso, Rosu, Gabelica, De Cian, Mergny, Teulade-Fichou. Ligands playing musical chairs with G-quadruplex DNA: A rapid and simple displacement assay for identifying selective G-quadruplex binders. Biochimie 90, 2008, 1207-1223

92 Doria Nadai, Folini, Scalabrin, Germani, Mella, Palumbo, Zaffaroni, Fabris, Freccero, Richter. Targeting loop adenines in G-quadruplex by a selective oxirane. Chem. Eur. 19, 2013, 78 – 81.

93 Rosu, De Paw, Gabelica. Electrospray mass spectrometry to study drug-nucleic acids interactions. Biochimie 90, 2008, 1074-1087

94 http://www.rcsb.org

95 http://www.scilligence.com/web/jsdraw.aspx

96 http://avogadro.cc

97 http://www.pymol.org

98 http://www.cgl.ucsf.edu/chimera/

99 web.chemdoodle.com

100 http://autodock.scripps.edu/resources/references