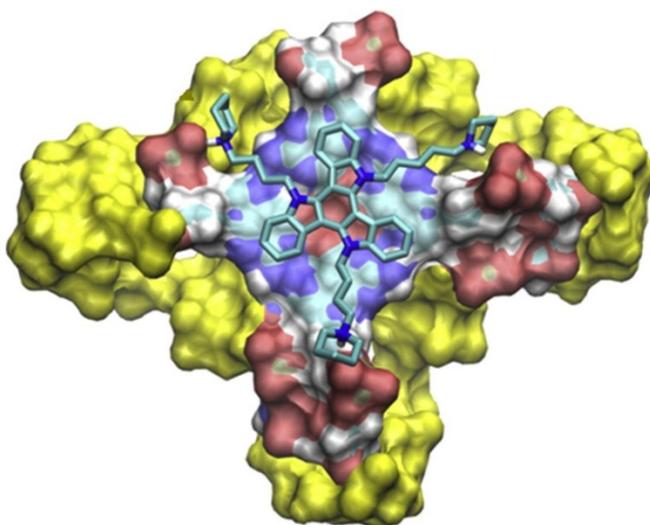


---

# New Anti-Telomerase, Anti-Cancer Drugs: A physico-chemical approach

---

Iolanda Fotticchia



Dottorato di Ricerca in Scienze Chimiche  
XXIV Ciclo  
Università degli Studi di Napoli "Federico II"



2011

UNIVERSITÀ DEGLI STUDI DI NAPOLI

“FEDERICO II”

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI



TESI DI DOTTORATO IN  
SCIENZE CHIMICHE  
XXIV CICLO

New Anti-Telomerase, Anti-Cancer Drugs:  
A physico-chemical approach

Dottoranda

Iolanda Fotticchia

TUTORE

Prof.ssa Concetta Giancola

RELATORE

Prof.ssa Daniela Montesarchio

CO-TUTORE

Dr. Luigi Petraccone

COORDINATORE DEL XXIV CICLO

Prof. Lucio Previtiera

# Contents

<b>ABSTRACT</b>	<b>4</b>
<b>INTRODUCTION</b>	<b>7</b>
<b>CHAPTER 1 G-QUADRUPLEX STRUCTURES AS ANTI-TELOMERASE TARGETS</b>	<b>10</b>
1.1.    TELOMERE AND TELOMERASE	10
1.2.    GENERAL STRUCTURAL FEATURES OF G-QUADRUPLEX	14
1.3.    HUMAN TELOMERIC G-QUADRUPLEX	17
1.4.    G-QUADRUPLEX LIGANDS	22
1.5.    THESIS PURPOSE	23
<b>CHAPTER 2 METHODS</b>	<b>25</b>
2.1.    ISOTHERMAL TITRATION CALORIMETRY	25
2.2.    CIRCULAR DICHROISM	30
2.3.    STEADY-STATE FLUORESCENCE	36
2.4.    DOCKING AND MOLECULAR DYNAMICS	38
<b>CHAPTER 3 TMPYP4 AND HUMAN TELOMERIC G-QUADRUPLEXES</b>	<b>44</b>
3.1.    INTERACTION BETWEEN TMPYP4 AND FOUR-REPEATS HUMAN TELOMERIC SEQUENCES	44
3.1.1. <i>Results</i>	47
3.1.2. <i>Discussion and Conclusions</i>	57
3.2.    INTERACTION BETWEEN TMPYP4 AND EIGHT-REPEATS HUMAN TELOMERIC SEQUENCE	62
3.2.1. <i>Results</i>	63
3.2.2. <i>Discussion and Conclusions</i>	71
<b>CHAPTER 4 A HYDROPHILIC THREE SIDE-CHAINED TRIAZATRUXENE AS A NEW G-QUADRUPLEX LIGAND</b>	<b>74</b>
4.1.    INTERACTION BETWEEN TRIAZATRUXENE DERIVATIVE (AZATRUX) AND HUMAN TELOMERIC G-QUADRUPLEX UNDER MOLECULAR CROWDING CONDITIONS.	74
4.1.1. <i>Results</i>	75
4.1.2. <i>Discussion and Conclusions</i>	89
<b>CHAPTER 5 NOVEL CLASSES OF POTENTIAL ANTI-CANCER AGENTS BINDING TO THE GROOVES OF G QUADRUPLEX STRUCTURES</b>	<b>93</b>
5.1.    SELECTIVE BINDING OF DISTAMYCIN A DERIVATIVE TO G-QUADRUPLEX STRUCTURE (TGGGGT) <sub>4</sub>	93
5.1.1. <i>Results</i>	95
5.1.2. <i>Discussion and Conclusions</i>	98
5.2.    VIRTUAL SCREENING-DERIVED COMPOUNDS AND THE G-QUADRUPLEX STRUCTURE (TGGGGT) <sub>4</sub> : AN ITC STUDY	100
5.2.1. <i>Results</i>	104
5.2.2. <i>Discussion and Conclusions</i>	107
<b>CONCLUDING REMARKS</b>	<b>109</b>

## Abstract

Human telomeric DNA terminates with a 3' single-stranded overhang containing tandem repeats of the sequence TTAGGG. These G-rich overhangs are prone to fold back to form G-quadruplex structures stabilized by consecutive G-tetrads each containing four guanines involved in Hoogsteen hydrogen bonds. G-quadruplex structures can inhibit the activity of telomerase, the enzyme that adds telomeric repeats to the ends of chromosomes and maintains the proliferation of cancer cells. Inhibition of telomerase can stop tumour growth and thus small molecules capable to interfere with telomere maintenance inducing the formation of quadruplex structures, are considered to be potential anti-cancer agents.

In this perspective my thesis project concerns the characterization of the human telomeric DNA and aims to the study of new anti-telomerase agents that specifically bind telomeric DNA stabilizing the G-quadruplex structures.

The first part of the thesis project concerns the study of  $\pi$ - $\pi$  stacking ligands. The aim was to solve the nature of the binding mode and stoichiometry of the cationic porphyrin TMPyP4 to G-quadruplex structures. The cationic porphyrin TMPyP4 has been extensively studied as a quadruplex-binding ligand since it induces telomerase inhibition. A systematic study was performed, in dilute solution and under molecular crowding conditions, of the binding reactions between TMPyP4 and four different human telomeric G-quadruplexes structures. The results clearly indicate that all G-quadruplexes are able to bind up to four TMPyP4 molecules. CD studies show that interaction with TMPyP4 promotes the conversion of the hybrid structures to an antiparallel conformation in dilute solution, while under molecular crowding conditions the interaction does not promote any conformational change. ITC reveals in both cases that the binding process comprises two sequential events, a first one in which one molecule of TMPyP4 interacts with the quadruplex structures and second one in which three other molecules bind to the structures. The selectivity of TMPyP4 for the quadruplex relative to duplex DNA was also investigated under molecular crowding conditions showing that TMPyP4 has enhanced selectivity for quadruplex DNA compared to the duplex structure. After then, taking into account that the 3'-terminal single-stranded human telomeric DNA is actually 100-200 bases long and can form higher-order quadruplex structures, the binding of TMPyP4 with a longer telomeric sequence able to form two consecutive quadruplex units (multimer), was explored. The

number of binding sites per quadruplex unit available in the multimeric structure is smaller than the one expected on the basis of the results obtained from individual quadruplex binding studies, suggesting that the quadruplex units along a multimer structure do not behave as completely independent. The presence of adjacent quadruplexes results in a diverse binding ability not predictable from single quadruplex binding studies. The existence of quadruplex-quadruplex interfaces in the full length telomeric overhang may provide an advantageous factor in drug design to enhance both affinity and selectivity for DNA telomeric quadruplexes.

Another possible quadruplex drug studied in this thesis was the three side-chained triazatruxene derivative, termed azatrux. The binding of azatrux to the human telomeric G-quadruplex, was explored in presence of 40 % PEG 200 to simulate the crowding conditions existing inside the cell. The binding of azatrux to the tetramolecular parallel [d(TGGGGT)]<sub>4</sub> quadruplex and to another biologically relevant G-quadruplex (oncogene promoter c-Kit87up) and to duplex DNA in the presence and absence of crowding conditions was characterized. The data indicate that azatrux binds in an end-stacking mode to the parallel G-quadruplex scaffold and highlights the key structural elements involved in the binding, showing a good selectivity for the human telomeric G-quadruplex over the other investigated DNA structures.

The last part of the Ph.D. study was focused on the characterisation of groove binder ligands. The binding of a dicationic derivative of distamycin A (compound **1**) with the [d(TGGGGT)]<sub>4</sub> quadruplex were explored by isothermal titration calorimetry (ITC), and compared to the binding behaviour of netropsin to the same target. The experiments show that netropsin and compound **1** are able to bind to [d(TGGGGT)]<sub>4</sub> with good affinity and comparable thermodynamic profiles. In both cases the interactions are entropically driven processes with a small favourable enthalpic contribution. Interestingly, the structural modifications of compound **1** decrease the affinity of the ligand toward the duplex, enhancing the selectivity. Furthermore the affinity of some new ligands with the grooves of DNA-quadruplex [d(TGGGGT)]<sub>4</sub> were studied. These ligands were derivatives of a molecule obtained from a previous virtual screening study. The interaction has been evaluated, analysing the displacement of these ligands from the grooves of DNA-quadruplex by distamycin A, the best groove binder identified until now, by means of isothermal titration calorimetry (ITC). The competition ITC experiments performed, represent a possible strategy to evaluate the efficiency of a groove binder compared to distamycin A. The derivatives have been tested over double-stranded DNA to

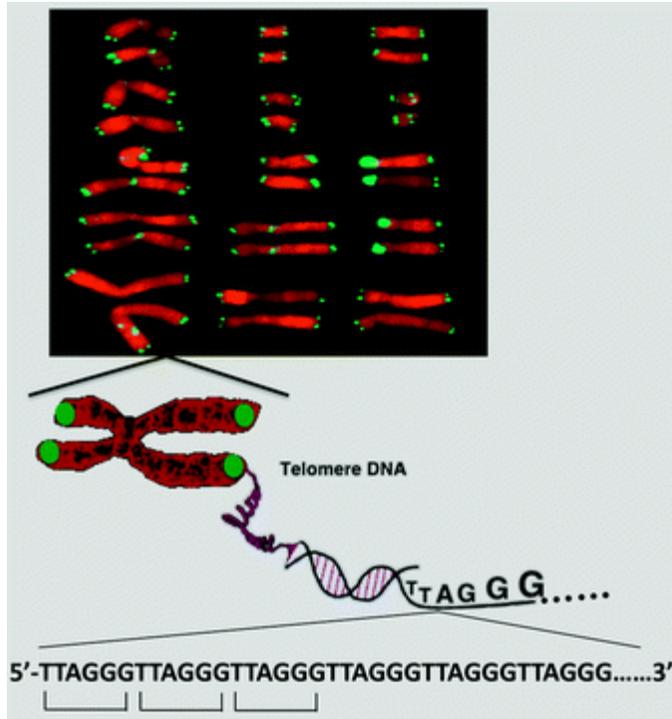
demonstrate whether there is an enhanced selectivity for quadruplex DNA compared to the duplex structure.

## INTRODUCTION

It was in the 1930 that telomeres (from the Greek telos = end and meros = part) were first recognized as essential structures at the ends of the chromosomes and were shown to be important for chromosomal stability. However, it was only in 1978 that the first telomeric sequence was identified in the Tetrahymena, a single cell organism that at a certain stage of development has many identical minichromosomes with twice as many telomeres [1]. Today we know that telomeres, the ends of the chromosomes, are essential for chromosomal stability and integrity. Telomeres exist as DNA protein complexes in which telomeric DNA is packaged by a series of unique telomere proteins.

Telomeres consist of many thousand repeats of the hexanucleotide TTAGGG giving rise to 4–15 kb of repetitive non-coding DNA in humans. Several kilobases of that sequence are paired with a complementary strand to form duplex DNA, but approximately 200 nucleotides of the sequence remain unpaired and forms a single-stranded overhang (Figure 1). The single-stranded overhang of telomeres progressively shortens during the replication process because DNA polymerase cannot fully copy the ends of telomeric DNA at each replicative cycle. Human cells are estimated to have the potential to undergo on average 60–70 divisions, and at this point the cells growth arrest and enter senescence [2]. So the telomere can be considered as a biological clock, and shortens with age.

Unfortunately integrity of telomeres in most cancer cells is maintained by the action of the telomerase enzyme complex, which catalyzes the synthesis of telomeric DNA repeats in order to replace those lost during replication. Telomerase is an enzyme overexpressed in ~80–85% of cancer cells and primary tumours [3,4], it maintains telomere length homeostasis acting as a tumour promoter. The catalytic subunit of telomerase (called hTERT in humans) has reverse transcriptase enzymatic activity and synthesizes TTAGGG repeats on the end of the 3' single-stranded overhang.



**Figure 1** *Fluorescent hybridization in situ showing human telomeres. The green spots mark the telomeres (adapted from ref. 5).*

The G-rich telomeric overhang is able to fold in unusual DNA structures, called G-quadruplex and stabilized by consecutive G-tetrads each containing four guanines involved in Hoogsteen hydrogen bonds [6]. It has been shown that the folding of the single-stranded telomeric DNA into a four-stranded G-quadruplex structure inhibits the enzyme's catalytic activity [7]. Inhibition of telomerase can stop tumor growth, and thus small molecules that induce the formation of quadruplex structures by binding to them are considered to be potential anticancer agents. In order to optimize the design of new anticancer agents, in my Ph.D thesis, the structural and energetic aspects of the binding of several small molecules to G-quadruplex structures formed by DNA telomeric sequences have been explored. In the study, all the known human telomeric quadruplex topologies were employed. Further, taking into account that the 3'-terminal single-stranded human telomeric DNA is 100-200 bases long, a longer telomeric sequences able to fold in more quadruplex units was also employed. Lastly, in order to mimic more closely the physiological conditions, the effect of crowding on the human

telomeric G-quadruplex structure and binding properties was also explored.

The obtained results provides new insights to understand the binding properties of the human telomeric quadruplex structures and could be useful for future optimization of new anti- telomerase, anti-cancer drugs.

# CHAPTER 1 G-quadruplex structures as anti-telomerase targets

## 1.1. Telomere and Telomerase

Telomeres exist as DNA-protein complexes in which telomeric DNA is packaged by a series of unique telomere proteins. Human telomeric DNA contains tandem repeats of the sequence TTAGGG and in all eukaryotic cells the telomeric region is double-stranded for most of its length, but the extreme 3'-end consists of single-stranded G-rich DNA overhangs approximately 200 nt in length.

The TTAGGG repeats of human telomeres are associated with the six-protein complex shelterin (Figure 1.1). Three shelterin proteins, TRF1, TRF2, and POT1, directly recognize telomere DNA sequences, and are interconnected by three additional shelterin proteins, TIN2, TPP1, and Rap1.

**Telomeric repeat-binding factors 1 and 2 (TRF1 and TRF2)** directly bind the TTAGGG sequences in double-stranded telomeric DNA by a C-terminal DNA binding domain [8,9].

**Protection of telomeres 1 (POT1)** directly binds the single stranded overhang of the chromosome end in its N-terminus [10]. The single-stranded DNA-binding domain in the N terminus allows the proteins to bind to arrays of the sequence TAGGGTTAG with great sequence specificity. It has been suggested that POT1 is able to disrupt the highly folded intramolecular G-quadruplex structures [11].

**TRF2- and TRF1-interacting nuclear protein 2 (TIN2)** bridges TRF1 and TRF2 to the TPP1–POT1 heterodimer linking the duplex component of the telomeres to the single-stranded overhang [12,13].

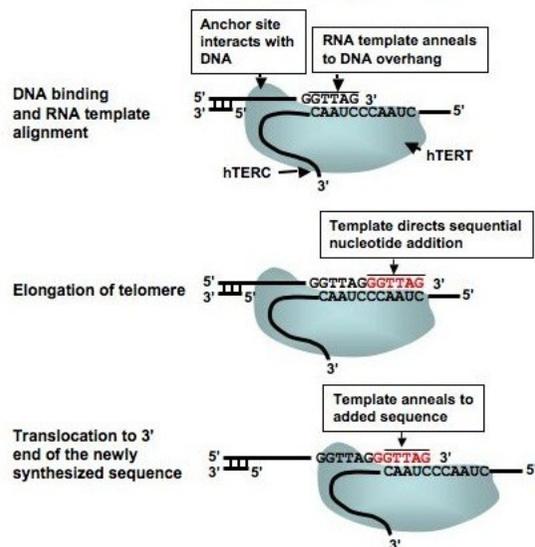
**Repressor/Activator protein 1 (Rap1)** lacks DNA binding activity in human and is dependent on its interaction with TRF2 for telomere binding [14,15].

**TIN2 and POT1 interacting protein 1 (TPP1)** interacts with POT1 and TIN2 by the POT1-binding domain and C-terminal of TIN2 [16]. The connection with TPP1 is critical for the association of POT1 with telomeres, providing a main way by which POT1 is recruited to telomeres [17,18]. Several telomeric DNA-binding proteins have been found to associate with telomere G-quadruplex. POT1 is known to



repeats directly onto the ends of chromosomes. The enzyme adds the first six bases, then translocates the template RNA for the synthesis of the next six bases (Figure 1.2) [32]. This extension at the 3'-end of the DNA template compensates for telomere shortening that results from nuclease action and incomplete terminal DNA replication. A multi component "telomere homeostasis" system promotes this telomeric extension when a telomere becomes shortened, thus making telomeric repeats accessible to the telomerase [29].

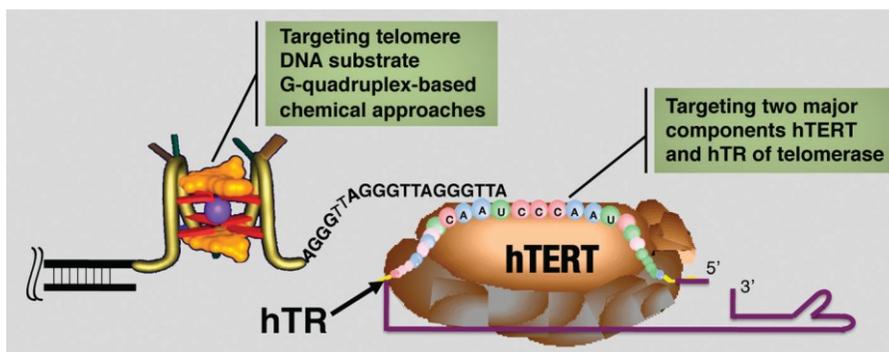
Hence, cancer cells maintain telomere length because telomerase assures the addition of telomeric repeats to chromosomal ends. These cells have high telomerase activity whereas somatic cells do not [33,34]. These findings are the base for the evaluation of telomerase inhibitors as potential anticancer agents [35,36,37]. The aim is to preserve from toxic side effects normal cells, which do not express the telomerase [38]. Different telomerase inhibitor strategies exist, based on targeting the different telomerase components or inhibiting telomerase activity.



**Figure 1.2** *Telomerase reaction cycle. In every reverse transcription cycle, the enzyme adds six bases, and then translocates the template RNA for the synthesis of the next six bases.*

Numerous approaches have been developed, including the targeting of the two major components hTERT and hTR of telomerase, and the

telomeres themselves (Figure 1.3). A possible approach is to design and synthesize drugs able to bind and stabilize the folding of the telomeric DNA G-rich overhangs into G-quadruplex structures (Figure 1.3) [7].

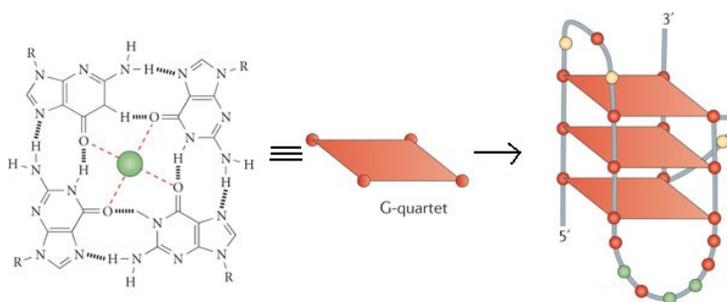


**Figure 1.3** *Inhibition mechanisms of telomerase (adapted from ref.5).*

G-quadruplex structures were detected *in vivo* at the ends of human chromosomes through the use of a titrated ligand that binds specifically to quadruplex DNA [39] and through detection using a fluorescent carbazole derivative that binds selectively to quadruplex structures [40]. Additionally in 1991, Zahler and coworkers found that folding of a telomeric G-rich single strand in a DNA quadruplex structure inhibits *in vitro* telomerase activity [7]. Hence, G-quadruplex structures within the telomere could effectively inhibit telomerase by blocking its binding to its substrate DNA and preventing elongation during replication. The development of new G-quadruplex ligands, as possible anticancer drug, has been the main focus of this Ph.D. thesis. A detailed structural knowledge of the G-quadruplex structures (the target structures) is critical for structure-based drug design and thus has been the subject of intense investigation.

## 1.2. General structural features of G-quadruplex

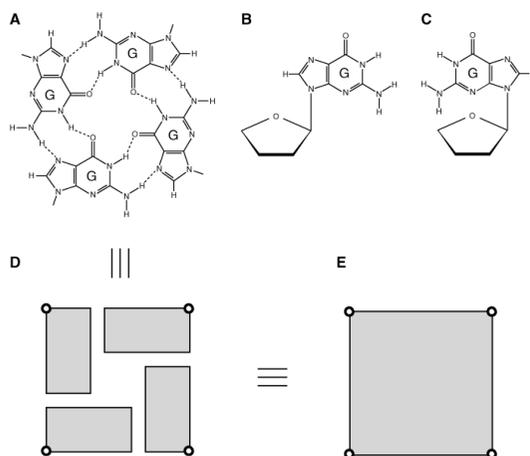
G-quadruplexes are higher-order DNA and RNA structures formed from G-rich sequences that are stabilized by tetrads of hydrogen-bonded guanine bases (G-quartets) [41]. Stacking of G-quartets and coordination of metal cations, preferentially  $\text{Na}^+$  and  $\text{K}^+$ , between two adjacent G-quartet planes yield a stable G-quadruplex structure (Figure 1.4) [42].



**Figure 1.4** Schematic representation of the G-quartet and G quadruplex.

G-quadruplexes can be formed from one, two or four separate G-rich strands of DNA (or RNA) and can display a wide variety of topologies, which are in part a consequence of various possible combinations of strand direction, as well as variations in loop size and sequence. They can be defined in general terms as structures formed by a core of at least two stacked G-quartets (or G-tetrads), which are held together by loops that are not involved in the quartets themselves. The G-quartet consists of a planar arrangement of four guanine bases associated through a cyclic array of Hoogsteen-like hydrogen bonds in which each guanine base both accepts and donates two hydrogen bonds (Figure 1.4). The G-tetrads are not stacked linearly, but adopt a right-handed helix. The combination of the number of stacked G-quartets, the polarity of the strands and the location and length of the loops lead to a plurality of G-quadruplex structures. G-quadruplexes can be classified on the basis of: 1) the number of strands (one, two or four); 2) the pattern of strand orientation (parallel, antiparallel); 3) the conformation of guanine glycosidic torsion angles (*anti* or *syn*); 4) the orientation of the loops (lateral, diagonal or propeller) [43].

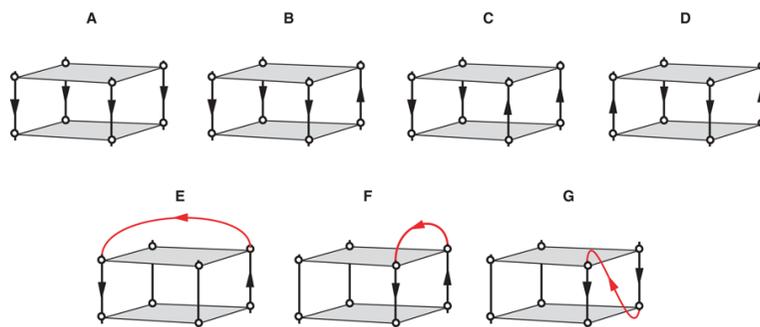
The G-tetrad core can be classified with regard to two mutually related factors, the relative orientations of the strands and the glycosidic conformations [anti (Figure 1.5B) or syn (Figure 1.5C)] of guanines, which in turn define specific patterns of groove dimensions.



**Figure 1.5** (A) G-tetrad alignment. (B,C) Guanine in (B) anti and (C) syn glycosidic conformations. (D, E) Schematic presentation of a G-tetrad (D) each guanine shown as a rectangular and (E) the G-tetrad shown as a square (adapted from ref. 44).

There are four different possibilities for the relative strand orientations in the G-tetrad core (Figure 1.6 A–D): (a) four strands are oriented in the same direction (designated a parallel-stranded core) (Figure 1.6 A); (b) three strands are oriented in one direction and the fourth in the opposite direction [designated a (3 + 1) core, also called a hybrid core in the literature] (Figure 1.6 B); (c) two neighbouring strands are oriented in one direction and the two remaining strands in the opposite direction (designated an up–up–down–down core, also called an antiparallel-stranded core in the literature) (Figure 1.6 C); and (d) two strands across one diagonal are oriented in one direction and the two remaining strands across the other diagonal in the opposite direction (designated an up–down–up–down core, also called an antiparallel-stranded core in the literature) (Figure 1.6 D). The glycosidic conformations of guanines within a G-tetrad are geometrically associated with the relative strand orientations. The hydrogen-bond directionality of a G-tetrad in the core can be clockwise or anticlockwise, and this is directly related to the

glycosidic conformations of guanines for each type of strand orientations. The stacking patterns between adjacent G-tetrads of the same hydrogen-bond directionality differ from those between adjacent G-tetrads of opposite hydrogen-bond directionalities [44].



**Figure 1.6** (A–D) Four types of G-tetrad cores: (A) parallel G-tetrad core, (B) (3 + 1) G-tetrad core, (C) antiparallel G-tetrad core (up–up–down–down) and (D) antiparallel G-tetrad core (up–down–up–down). (E–G) Three types of loops (colored red): (E) diagonal loop, (F) edgewise loop and (G) double-chain-reversal loop. Arrows indicate the strand orientations, from 5' to 3' direction (adapted from ref.44).

When one or more G-tetrads are stacked, a cylindrical central cavity is produced [45]. This cavity, lined with the guanine O6 carbonyl oxygens, forms a specific binding site for metal ions [46]. The coordination of potassium [47], sodium [48], and strontium [49] all provide both thermodynamic and kinetic stability to the G-quadruplex structure. Experiments demonstrate that the G-quadruplex is stabilized by the alkali series in the following order:  $K^+ \gg Na^+ > Rb^+ > NH_4^+ > Cs^+ \gg Li^+$ , and for the alkaline earth series in the order:  $Sr^{2+} \gg Ba^{2+} > Ca^{2+} > Mg^{2+}$  indicating that the atomic radii of 1.3 Å of potassium and strontium fit best in the coordination site between adjacent G-tetrads [50]. The precise location of the cations between the tetrads is dependent on the nature of the ion.  $Na^+$  ions within the channel have been observed in a range of geometries; in some structures, a  $Na^+$  ion is in plane with a G-tetrad whereas in others it is between two successive G-tetrads.  $K^+$  ions are always equidistant between each G-tetrad plane, interacting equally with each of eight carbonyl oxygens in a bipyramidal antiprismatic coordination geometry [51].

All quadruplex structures have four grooves, defined as the cavities bounded by the phosphodiester backbones. A consequence of variations in the glycosidic torsion angles is the altering of the spacing between the DNA strands of the G-quadruplex, resulting in varying groove-widths along the side of the quadruplex [52]. In cases where a guanosine in the *syn* conformation donates hydrogen bonds to a neighboring guanosine in the *anti* conformation, the groove formed between the two is extremely narrow, with a phosphate to phosphate distance as small as 7-9 Å [53]. In contrast, when the hydrogen bonding polarity between adjacent *syn*- and *anti*-guanosines is reversed, a very wide groove is formed [53]. An intermediate width groove results when adjacent guanosines adopt the same glycosidic conformation [53]. Indeed, in a parallel four-stranded quadruplex, where all the guanine bases are in the *anti* conformation, the four grooves are approximately the same medium width. For example, in a symmetric bimolecular structure having lateral loops, the G-tetrads are formed from guanosines of alternating *anti-syn-anti-syn* conformation, with each *syn*-guanosine donating hydrogen bonds to an adjacent *anti*-guanosine, and accepting hydrogen bonds from the other adjacent guanosine [54]. This results in a rectangular G-tetrad core with grooves of alternating wide-narrow-wide-narrow widths. In contrast, a bimolecular quadruplex with diagonal loops, has guanosines that adopt the *syn-syn-anti-anti* conformations, which results in G-tetrads that adopt a parallelogram arrangement and the formation of alternative wide, medium, narrow, medium width grooves between strands. Groove dimensions are then variable, and depend on overall topology and the nature of the loops. Grooves in quadruplexes with only lateral or diagonal loops are structurally simple, but, in contrast, grooves that incorporate propeller loops have more complex structural features that reflect the insertion of the variable-sequence loops into the grooves.

### 1.3. Human telomeric G-Quadruplex

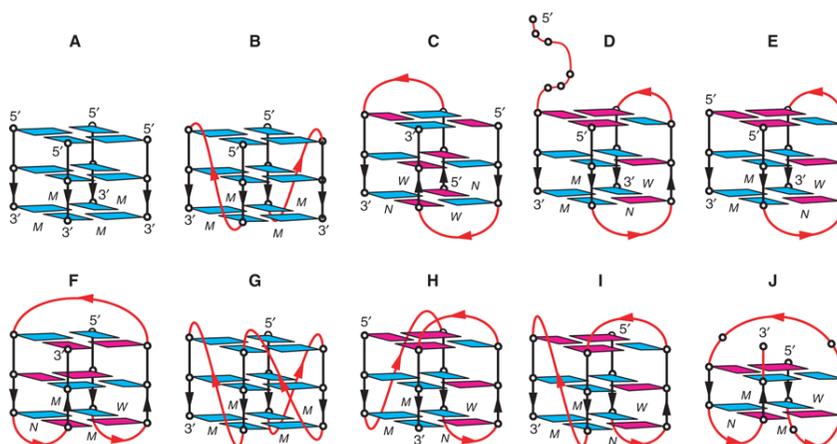
The rational design of G-quadruplex-binding drugs requires knowledge of the detailed structures of the intramolecular human telomeric G-quadruplexes to be successful in a potential anticancer therapy. In this perspective extensive researches have been dedicated to G-quadruplex structures formed by human telomeric DNA. Several native structures, formed by human telomeric G-quadruplex sequences, have been

determined by both NMR and X-ray techniques [55,56,57]. Interestingly, it has been shown that the conformation of the intramolecular human telomeric G-quadruplexes is dependent upon the type of counter ions and the presence of flanking bases. An antiparallel basket conformation was found for human telomeric d[AGGG(TTAGGG)<sub>3</sub>] sequence in Na<sup>+</sup> containing solution (Figure 1.7F) [55]. A hybrid-type conformation, the so-called (3+1) mixed conformation, appears to be a major form for various human telomeric oligonucleotides in solution in the presence of K<sup>+</sup> [58,59]. Depending on the flanking segments, those telomeric sequences can form two distinct but related hybrid-type structures, known as hybrid-1 or Form 1 (Figure 1.7H) and hybrid-2 or Form 2 (Figure 1.7I) [60,61]. It has been shown that some human telomeric sequences are in equilibrium between hybrid-1 and hybrid-2 structures in K<sup>+</sup> solution, which is largely determined by the 3'-flanking sequence. The hybrid-2 type structure appears to be a plausible conformation for the extended human telomeric sequences [61]. The structure of a new intramolecular G-quadruplex conformation formed by the human telomeric d[(GGGTTA)<sub>3</sub>GGGT] sequence in K<sup>+</sup> solution has recently been reported (Figure 1.7J). This is an antiparallel basket-type G-quadruplex involving only two layers of G-tetrads (Form 3) and has been reported to be of greater stability than the two hybrid forms [57]. The parallel form has been observed in the crystal structure of the four-repeat sequence d[AGGG(TTAGGG)<sub>3</sub>] in K<sup>+</sup> (Figure 1.7G) [62], and may be present in more crowded solution conditions, such as in a cell nucleus [63].

G-quadruplex structures determined for human telomeric DNA sequences show different configurations for TTA loops in various contexts involving three G-tetrads, as well as GTTA and GTTAG loops in Form 3 involving two G-tetrads. Base pairing and stacking are generally observed in these loops [64,65]. It has been suggested that these loops are dynamic and may be good targets for specific ligand recognitions [66,67].

In summary, the crucial information provided by all the structural studies is that the particular DNA sequences and the solution conditions employed are relevant in determining the structure observed. Hence to determine the most physiological relevant human telomeric DNA structures, studies must be performed with DNA in experimental conditions mimicking more closely the living cellular environment. On this regard, two points should be considered: 1) that the full length telomeric DNA is much longer than the short DNA sequences commonly employed in the biophysical studies and can, in principle, form more than one consecutive quadruplex unit [68,69,70], 2) the cellular environment

is “crowded” with other biomolecules and the behavior of the telomeric DNA in this “crowding conditions” could be significantly different from the one in the buffers, commonly employed in laboratory experiments (also called “dilute conditions”).



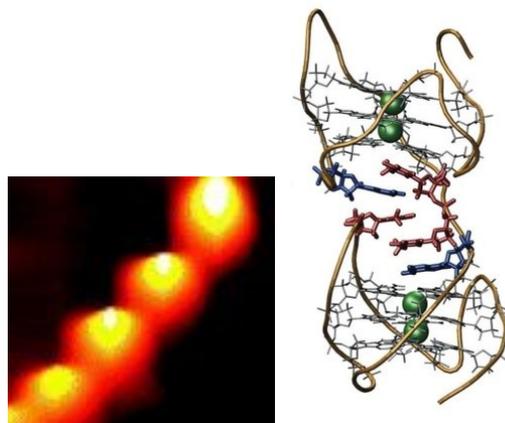
**Figure 1.7** Schematic structures of G-quadruplexes. (A) Tetrameric parallel-stranded G-quadruplex [71]. (B) Dimeric parallel-stranded G-quadruplex [72]. (C) Dimeric antiparallel-stranded G-quadruplex [72]. (D) (E) Asymmetric dimeric (3+1) G-quadruplex [73]. (F) Basket-type form [55]. (G) Propeller-type form [62], (H) (3+1), Form 1 observed for  $d[TA(GGGTTA)_3GGG]$  in  $K^+$  solution. (I) (3+1) Form 2 observed for  $d[TA(GGGTTA)_3GGGT]$  in  $K^+$  solution [59]. (J) Basket-type form observed for  $d[(GGGTTA)_3GGGT]$  in  $K^+$  solution [57]. anti-guanines are colored cyan; syn-guanines are colored magenta; loops are colored red. M, N and W represent medium, narrow and wide grooves, respectively (Adapted from ref.44).

### 1.3.1. Structures of longer DNA telomeric sequences

Structural studies on human telomeric DNA has been mainly performed employing four-repeats telomeric sequences, little is known about the structure of longer DNA telomeric sequences. However, it might be more biologically relevant to take into account the real length of the 3'-terminal single-stranded overhang of human telomeric DNA in vivo (100–200 nt). The telomeric single-stranded overhang is actually constituted by tens of TTAGGG repeats and, a growing evidence exists that it forms several consecutive quadruplex structures (multimers) in the overhang region [67,69,74,75,76,65]. Recently, G-quadruplexes were

visualized at the end of telomeres by atomic force microscopy (AFM) [76]. The AFM image shows blob-shaped protrusions arranged end-to-end, suggesting the presence of adjacent G-quadruplex units (Figure 1.8) [58]. The formation of two contiguous quadruplex in the GGG(TTAGGG)<sub>8</sub> DNA telomeric sequence and its RNA analogous has been recently demonstrated on the basis of ESI-MS experiments [77]. Several models have been proposed regarding the arrangements of these G-quadruplex blocks [69]. In one model, G-quadruplex blocks are arranged like ‘beads-on-a-string’ [69], they can move relatively independently of each other and are constrained only by the connecting linkers. Alternatively, G-quadruplex blocks can stack on each other to form a higher order structure. Different multimeric structures have been proposed for the longer DNA telomeric sequences depending on the type of quadruplexes involved in the multimers and on the kind of interactions between the quadruplex units. Neidle and coworkers proposed, on the basis of a computational study, a multimeric model formed by consecutive quadruplex units in all-parallel-stranded topology from the observed crystal structure of the four-repeat sequence d[AGGG(TTAGGG)<sub>3</sub>] in K<sup>+</sup> [67]. Later Petraccone *et al.* have suggested that multimeric structures could be easily formed by quadruplex units with the mixed-type G-quadruplex topology (called “Hybrid-1” and “Hybrid-2” or “Form-1” and “Form-2”) on the basis of NMR structures observed for d[TA(GGGTTA)<sub>3</sub>GGG] [62]. (I) (3+1) and d[TA(GGGTTA)<sub>3</sub>GGGTT] in K<sup>+</sup> solution [59] (right side Figure 1.8) [78,65]. In further support of this hypothesis, it has been proposed a dimeric model for the (TTAGGG)<sub>8</sub>TT telomeric sequence consisting of two consecutive mixed-type quadruplexes, one having the Hybrid-1 conformation and the other having the Hybrid-2 conformation, this model has been validated experimentally [69,79].

Despite the increasing number of structural information on longer DNA telomeric sequences, very few data are available on the binding properties of these sequences compared with the shorter DNA telomeric sequences. However, the existence of quadruplex-quadruplex interfaces in the longer telomeric sequences could be a predominant factor in determining the recognition properties of the telomeric DNA and the results of binding studies performed on single quadruplex may not be necessarily useful to predict the binding properties of multimeric structures. To address this question, in this Ph.D. thesis, were compared the binding properties of multi-quadruplex structures with the one of the single quadruplex structures (discussed in chapter 3).



**Figure 1.8** *AFM image of higher-order telomeric DNA structures (left side), human telomeric DNA, Hybrid-12 model [64] (right side).*

### 1.3.2. *G-quadruplex under molecular crowding conditions*

Biomolecules have evolved to function within living cells, which contain a variety of macromolecules including nucleic acids, proteins, polysaccharides, and metabolites. These molecules make the intracellular environment extremely crowded, so becomes of biological interest to study biochemical reactions in systems that simulate the living environment that contain substantially a total concentration of ~400 mg/ml of macromolecules in solution [69,80,81]

It has been reported that molecular crowding is a critical factor determining the structure, stability and function of nucleic acids [63]. Recent studies have raised significant controversy regarding the exact structure of the quadruplex formed by human telomeric DNA in a physiological relevant environment. Particularly, it has been reported that different G-quadruplex conformations are converted to a propeller-type parallel-stranded G-quadruplex in  $K^+$  containing crowded solution due to water depletion [82]. Human telomeric G-quadruplexes have been studied in different molecular agents such as putrescine, cadaverine, spermine, glycerol, and polyethylene glycol [83,84]. The most common crowding agent is polyethylene glycol (PEG 200). PEG is a neutral polymer and does not interact with G-quadruplex because the interaction is thermodynamically unfavorable. The volume occupied by PEG grows

large as it is added increasingly, therefore, excluded volume could be the major factor inducing quadruplex structure [85].

Molecular crowding plays an important role in the regulation of human telomeric G-quadruplexes conformation, for this reason becomes very important to perform binding studies in solution conditions that more closely mimic the cell environment.

#### 1.4. G-quadruplex ligands

G-quadruplex stabilizing compounds have recently received increased interest due to their potential application as anticancer therapeutics, substantial efforts have been made to design synthetic compounds that could lock or induce telomeric DNA in a G-quadruplex conformation and prevent telomere elongation. Recently a large number of small molecules (G-quadruplex ligands or G4-ligands), have been designed [86,87,88].

A desired ligand would recognize a G-quadruplex structure formed by human telomeric sequences with high affinity and specificity. Different G-quadruplex recognition modes are possible: (a) stacking on the ends of the G-tetrad core, (b) groove binding, (c) taking place of one or more strands in the core, (d) interacting with the backbone (core and loops), and (e) interacting with the loop bases. Many of the reported G-quadruplex ligands contain planar aromatic rings [89,90], which can interact with human telomeric G-quadruplex by stacking on the terminal G-tetrads. In addition to the end-stacking binding mode of the aromatic rings, some ligands also contain other moieties that can recognize loops by stacking with loop bases or forming intermolecular hydrogen bonds [89,90,91] or recognize the backbone with electrostatic interactions [92,93].

**( $\pi$ )-stacking ligands:** For efficient stacking, these ligands should possess a large aromatic surface, much larger than duplex binders, for efficient aromatic–aromatic overlap and selectivity. In addition, the ligands must exhibit reasonable water solubility, usually achieved by incorporating side chains around the aromatic core, which can be protonated. The side chains also contribute to the stability of quadruplexes by establishing electrostatic interactions with atoms in grooves and central channel [94].

**Groove/loop binding ligands:** There are significant chemical and conformational differences between quadruplex and duplex grooves and

therefore recognition of these grooves by ligands is expected to provide higher quadruplex *vs.* duplex selectivity. Distamycin A is the only compound for which groove binding has been proved until today. Martino *et al.* investigated the interaction of distamycin A with d[TGGGGT]<sub>4</sub> and found that the ligand binds with a 4/1 ratio to the quadruplex as a dimer in an anti-parallel fashion [95]. Using the above principles, should be relatively easy to design compounds that bind G-quadruplexes, although not necessarily with high affinity and more importantly selectivity (ability to interact only with quadruplex DNA and not duplex DNA).

## 1.5. Thesis purpose

Despite a wealth of current knowledge about human telomeric G-quadruplexes, there remain many challenges associated with the structure and molecular recognition in the human telomeres.

These include: (a) the structure and dynamics of all possible DNA G-quadruplexes formed by short and long human telomeric sequences; (b) the structural basis for molecular recognition of human telomeric G-quadruplexes by different small molecules and (c) the detection of G-quadruplex structures and conformational transitions in the human telomeres in the physiological environment of living systems.

The general purpose of this thesis was to study the binding properties between different topologies of DNA G-quadruplex and selected ligands in different buffer conditions.

The energetic of the interaction with different sequences of human telomeric DNA with cation porphyrin TMPyP4 and three side-chained triazatruxene derivative, was explored.

Additionally the interaction between the truncated sequence of *Tetrahymena* telomeric DNA, d(TGGGGT), which forms a tetramolecular quadruplex in solution, was explored with new groove binder ligands like distamycin A derivatives and new virtual screening derived compounds.

Lastly the binding properties of the selected ligands have been studied over other biologically relevant DNA structures, in order to verify, whether there is, an enhanced selectivity for human telomeric quadruplexes over the others [96,97,98,99].

The binding stoichiometry and the thermodynamic parameters related to the association processes, have been studied by means of isothermal titration calorimetry (ITC) and other spectroscopic techniques such as circular dichroism, fluorescence spectroscopy and UV-vis spectroscopy, in order to understand the overall nature of the forces that drove the binding reactions and to get more insight into the structural aspects of binding.

## CHAPTER 2 Methods

A wide variety of techniques are employed to study interactions of ligands with G-quadruplex DNA. These methods vary from simple methods, used for evaluating properties like ligand affinity, to more sophisticated methods, used for obtaining kinetic, thermodynamic, stoichiometric and conformational data for structure–activity relationship studies. The key challenge in G4-ligand design is to make the ligand selective for quadruplex structures in comparison to duplex DNA and the biophysical and biochemical method being used to investigate these interactions should be able to detect and measure the ligand selectivity for quadruplexes over duplexes or other secondary structures. Because of high degree of structural polymorphism in quadruplexes, it is important to determine binding sites of the ligand and its mode of interaction. Each technique has its own advantages and disadvantages and usually more than one method is needed to obtain complete information about quadruplex DNA/ligand interactions. The methodology and experimental procedures applied in this thesis are, essentially, belonging to two different families

- Spectroscopic techniques
- Calorimetric techniques

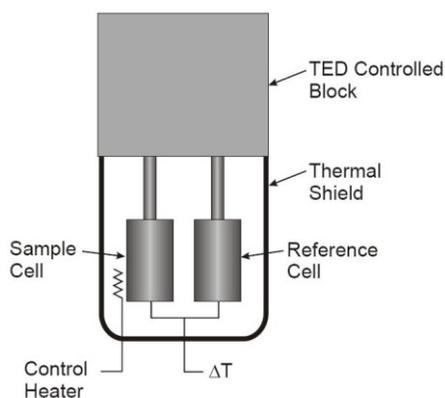
The techniques belonging to the first family, like UV-visible (UV-Vis), fluorescence and circular dichroism (CD) spectroscopy are rapid methods and require small amounts of material. These techniques allow to investigate on the conformational proprieties of the studied macromolecules. On the other hand the techniques belonging to the second family allow the study of the energetics aspects of the studied interactions.

### 2.1. Isothermal Titration calorimetry

Isothermal titration calorimetry (ITC) is a valuable tool for characterizing interactions of G-quadruplexes with other biomolecules, including small ligands, thanks to its general applicability and precision [100,101]. ITC is a high-accuracy method for measuring binding affinities and

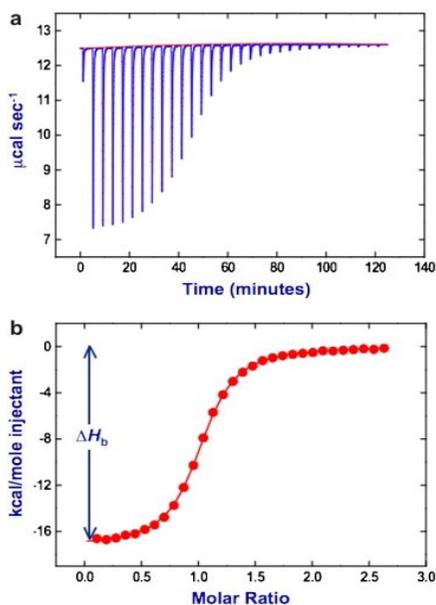
stoichiometry; moreover, it is the only technique that directly measures the binding enthalpy. ITC also allows dissecting the free energy of binding into enthalpic and entropic components to reveal the overall nature of the forces that drive the binding reaction [102,103]. In ITC, one component of a complex (such as DNA or a protein) is present in the calorimeter's sample cell, and the second component (for example, a drug) is slowly added in an incremental, stepwise fashion. Since all binding events are accompanied by the evolution or absorption of heat (a change in enthalpy), the analysis of these extremely small thermal effects arising from the binding allows a full thermodynamic characterization of the reaction and provides fundamental information about the molecular interactions driving the process. The calorimeter used in the experiments is the CSC 5300 nano ITC III TA Instruments.

The calorimeter holds two vessels (reference and sample cells), with a nominal volume of 1.0 ml. The reference vessel contains water or buffer solution, the sample vessel holds a solution containing one of the reactants. The two vessels are constantly kept in thermal equilibrium with the bath during the experiment. The heat flow between the reaction vessel and the isothermal block is precisely measured by thermoelectric device sensors that surround the vessel and is monitored as a function of time. The nano ITC uses a differential power compensation design for maximum sensitivity and responsiveness. Semiconducting thermoelectric devices (TED) are used for temperature control as well as to detect temperature differences between the sample and reference cell. A computer controlled PID (a proportional–integral–derivative controller) loop uses a heater on the sample cell to maintain a zero temperature difference between the sample and the reference cell. The ITC measures the heat generated/absorbed when two solutions of reactants are mixed.



**Figure 2.1** Schematic representation of a CSC Model 5300 nano ITC III.

The power required to maintain the zero difference between the cells represents the calorimeter signal and is monitored as function of time. If a reaction occurs, the sample cell produces heat, the heat required to maintain the zero difference decreases by the amount of heat supplied by the reaction, resulting in a peak in the thermogram. A typical experiment consists of several injections spaced minutes apart, resulting in several peaks in the thermogram. The reactant is contained in a precision syringe (25-250  $\mu\text{l}$ ) connected to the sample cell and is added in computer-controlled injections (1-20  $\mu\text{l}$ ). The titrant syringe needle also functions as the stirrer and extends down into the reaction vessel from the top when the burette is mounted. The nano ITC can detect heat effects as small as 0.4 milli Joule, allowing titrations to be done with nanomoles of biopolymer, ITC may be run at any temperature between 0 and 110°C. In a typical ligand-macromolecule titration, the chemical reaction generated by each injection either releases or absorbs a certain amount of heat ( $q_i$ ) proportional to the amount of ligand that binds to the macromolecule in a particular injection ( $V \times \Delta L_i$ ) and to the characteristic binding enthalpy ( $\Delta H$ ) for the reaction:  $q_i = V \times \Delta H \times \Delta L_i$  where  $V$  is the volume of the reaction cell and  $\Delta L_i$  is the increase in the concentration of bound ligand after the  $i^{\text{th}}$  injection [104]. The heat after each injection is therefore obtained by calculating the area under each peak. Because the amount of uncomplexed macromolecule available progressively decreases after each successive injection, the magnitude of the peaks becomes progressively smaller until complete saturation is achieved. Once saturation is reached, subsequent injections produce similar peaks corresponding to dilution or constant mechanical effects that need to be subtracted from all the injection peaks before analysis. The top panel of Figure 2.2 depicts the signal produced by the sequence of injections, and the bottom panel shows the data after integration of each injection peak. The sigmoidal shape of the bottom panel (Figure 2.2), with numerous data points throughout the curved rise portion of the plot, facilitates estimation of the midpoint of the transition, and thus the stoichiometry of the binding reaction (in this case, 1:1).  $K_b$  and  $\Delta H$  are calculated by iterative approximation.



**Figure 2.2** (a) Signal produced by the sequence of injections for a typical ITC experiment. (b) Data after integration of each injection peak.

A value for  $K_b$  is initially estimated, and then the concentration of bound complex is calculated for each injection. In combination with the measured heat, these values are used to determine the average of  $\Delta H$ . The  $\Delta H$  and the calculated concentration are then used to determine an expected heat per injection, and the error square sum between the measured and expected heat for each peak is calculated. The value of  $K_b$  is then adjusted and the process repeated until a minimum error square sum is obtained.

The values of  $K_b$ ,  $\Delta H$  and  $n$  (stoichiometry) can often be calculated from a single experiment as long as the concentration of both macromolecule and ligand are accurately known [105] and chosen so that:

$$10 < K_b [M]_T < 1000$$

where  $[M]_T$  is the total concentration of macromolecule in the sample cell titrated by ligand [106]. If concentrations are not within this range, the curvature of the titration plot can be so low as to be almost linear, or so high as to produce a step-like profile. In these cases  $K_b$  may not be estimated accurately. Typically, macromolecule concentrations in the order of 10-100  $\mu\text{M}$  are used, permitting  $K_b$  values in the range  $10^2$ – $10^9$   $\text{M}^{-1}$  to be accurately estimated.

Since temperature (T) is held constant throughout the entire experiment, the free energy ( $\Delta G^\circ$ ) of the binding reaction can be determined by:

$$\Delta G^\circ = -RT \ln K_b$$

where R is the gas constant.

ITC directly measures  $\Delta H^\circ$ , so the change in entropy ( $\Delta S^\circ$ ) can be determined by:

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

Quantification of these thermodynamic parameters reveals the physical processes involved in the binding reaction. A spontaneous binding process must have a negative  $\Delta G^\circ$ , and  $\Delta G^\circ$  will become increasingly negative as binding becomes tighter. As seen above, free energy changes have both an enthalpic and entropic component. The enthalpic contribution to binding is primarily due to an increased number of hydrogen bonds at the ligand-target interface, and to more favorable van der Waals interactions between the two interacting molecules; the hydrophilicity of the system will determine how important electrostatic, polar and dipolar interactions will be in driving the reaction. The entropic contribution has two primary components: conformational changes, such as folding or unfolding of the macromolecules, and the release of bound solvent as hydrophobic groups interact. The large number of ordered water molecules released into the bulk solvent when the hydrophobic surfaces of the ligand and target interact provides the main driving force for hydrophobic interactions. This driving force is sufficient to compensate for the unfavorable conformational entropy of the macromolecule and ligand caused by decreased conformational and rotational freedom following binding. In addition to the entropic effect, burial of surface area also affects the heat capacity of the sample, since water molecules ordered at hydrophobic surfaces have a different heat capacity from that of water that has been released into the bulk solvent following binding.

Of paramount importance when conducting ITC experiments with nucleic acids is sample preparation. It is now routine to obtain large amounts of synthetic DNA from commercial suppliers. However it is very important to make sure that the sequence has been subject to purification by HPLC or capillary electrophoresis to ensure that the sample is better than 99% pure. The presence of contaminating sequences from the synthesis will yield erroneous data in the ITC

experiment, especially in terms of stoichiometry if there is competitive binding to the non-standard sequence. The binding curve might also appear to be biphasic or suggestive of some complex binding mechanism (as opposed to being a simple sigmoid). With a relatively impure sample it is difficult to establish whether the origin of the complexity is indeed genuine or whether it is due to contaminating sequences present in the solution mixture. With regard to solution conditions and nucleic acids it is worth emphasizing that ITC is rarely if ever used as an isolated technique, when studying DNA. Generally ITC is used in conjunction with a series of experiments that typically involve circular dichroism (CD), UV/vis and fluorescence spectrophotometry.

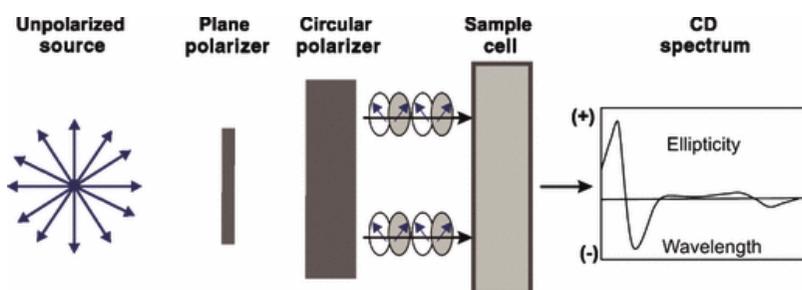
For meaningful experimental design and accurate data analysis it is clearly necessary to know the concentrations of nucleic acid (and ligand solutions) used in experiments. In order to obtain meaningful data it is necessary to have a clear idea about the following parameters: (1) DNA concentration in the calorimeter cell, (2) ligand concentration in the ITC syringe, (3) total volume of the ligand solution to be injected, (4) volume of each serial injection, (5) time between injections, and (6) experimental temperature. To select values for each of these parameters it is useful to know something about the system under study, especially the approximate value for the binding constant and the stoichiometry. If, however, nothing is known a good starting point for small molecule–nucleic acid interactions is to assume a binding constant of  $10^5 \text{ M}^{-1}$  and a stoichiometry of 3:1 per ligand. The aim is to produce a sigmoidal binding curve, which can be fit with a model. Once the DNA concentration is established, the ligand concentration can be chosen by remembering that it is necessary to ensure that the total ligand concentration at the end of the titration should be twice that of the initial concentration of binding sites in the cell [107].

Due to the commercial availability of calorimeters sensitive enough to measure low energy interactions involved in a small amount of biological materials, ITC is being increasingly used to investigate the recognition of large biomolecules, like nucleic acids [107,95,108].

## **2.2.Circular Dichroism**

The optical activity of small molecules arises from their lack of symmetry, particularly, from the presence of asymmetric carbon atoms

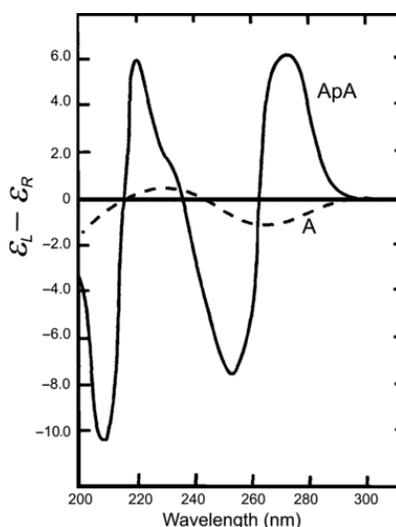
and from the effect these atoms have on any nearby chromophores. On the other hand, the optical activity of the biological samples derives, essentially, from the conformation of the macromolecules in solution. There are at least four ways that an optically active sample can alter the properties of transmitted light: optical rotation, ellipticity, circular dichroism and circular birefringence. Circular Dichroism (CD) is the phenomenon that is recorded when an optical active sample interacts with the plane-polarised light, particularly, when the left circularly polarised light is absorbed in a different amount with respect to the right circularly polarised light. After passing through the sample each component is still circularly polarised, but the radii of the circle traced out by the electric vector of each are now different. When these two opposite circularly polarised light waves are combined together the result will be elliptically polarised light because the two components have different amplitude. Recording the CD at different wavelength allows obtain the CD spectrum that is strongly suggestive, for the biological samples, to the secondary structure content of the macromolecule (Figure 2.3).



**Figure 2.3** Scheme of CD in which a difference in absorption is measured (adapted from ref. 109).

So, any conformational changes in the structure of macromolecules could be assessed using circular dichroism. As a consequent, unfolding of biomolecules (such as proteins, nucleic acids, glycosides, etc.) is measured as a change in circular dichroism (CD) spectra, and serves to give a measure of the relative quantities of changes have been made in the components. According to the literature, for reporting the CD of a sample, mean residue ellipticity ( $\text{degree cm}^2 \text{dmol}^{-1}$ ) and molar circular dichroism or delta epsilon ( $\text{L mol}^{-1} \text{cm}^{-1}$ ) are common. Nucleotides are

the main building blocks, which have been used in asymmetric structure of nucleic acids (RNA/DNA). The chiral sugars of nucleosides have an intrinsic asymmetry and the interaction of the strong  $\pi \rightarrow \pi^*$  transitions of the chromophoric bases with the higher energy in the sugars yield a circular dichroism of low intensity. In fact, circular dichroism of nucleic acids is mainly dependent on the stacking geometry of the bases. The difference in CD between a nucleoside, adenosine (A), and a dinucleoside phosphate, adenylyl-3'-5'-adenosine (ApA), is illustrated in Figure 2.4 [110].



**Figure 2.4** CD spectra of adenylyl-3'-5'-adenosine (ApA) compared with adenosine (A). For both molecules, the spectra are given per mole of nucleoside. The spectra are for aqueous solutions at pH 7 and room temperature (adapted from ref.109).

CD of the dinucleoside phosphate per adenosine is about a factor of 10 larger than the CD of adenosine. In adenosine, the CD depends on the interaction of adenine with its ribose and phosphate groups; whereas in the dinucleoside phosphate, the CD is mainly originated from the chiral adenine–adenine interaction. The combination of positive and negative extrema on either side of 260 nm is called an exciton band; there is another exciton band at 215 nm. The positive signs (long wavelength component positive, short wavelength component negative) of these two

bands indicate that the two adenines are forming a right-handed stack in ApA [110].

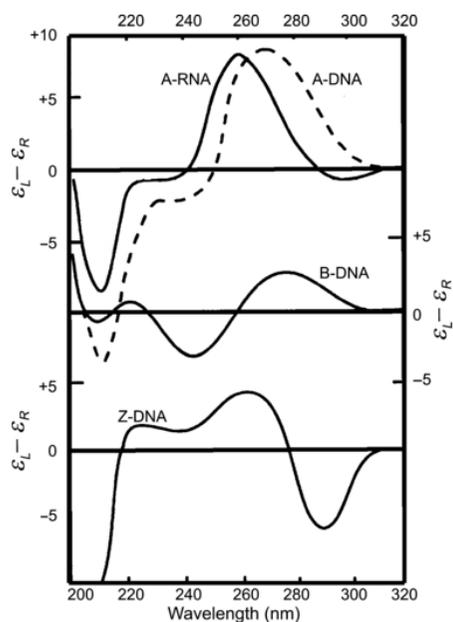
Today measurements of circular dichroism are widely used to quantify the association reactions between DNA and ligands. The association between the quadruplex structures and ligands can be studied by means of CD titrations. In particular, the saturation fraction ( $\alpha$ ) of the each quadruplex sample is measured, following the changes in molar ellipticity at a characteristic wavelength upon ligand addition, using the following relationship:

$$\alpha = \frac{CD - CD^{free}}{CD^{bound} - CD^{free}} \quad (\text{eq. 1})$$

where CD represents the CD signal at the wavelength chosen at different molar ratios (ligand/quadruplex sample) investigated,  $CD^{free}$  represents the CD signal of the quadruplex sample alone and  $CD^{bound}$  is the CD signal of the saturated quadruplex samples. The resulting binding curves, obtained plotting the saturation fraction versus the molar ratio between the ligand and the quadruplex in solution, can be interpolated with an equation describing an equilibrium binding model. The interpolation analysis generally provides values for the stoichiometry (n) and binding constant ( $K_b$ ). Two different models can be used in the interpolating procedure, the independent and equivalent-sites model and a cooperative-sites model based on the Hill approach [111].

### ***2.2.1. Determination of nucleic acid conformations***

The circular dichroism spectra of A-RNA, A-DNA, B-DNA and Z-DNA in the 200–320 nm range are shown in Figure 2.5. Both A-RNA and A-DNA have spectra that are similar in shape. The A-RNA has a maximum near 260 nm, a minimum near 210 nm, and a small negative CD between 290 and 300 nm. The A-DNA has a maximum at 270 nm, a minimum near 210 nm and zero CD at 300 nm and beyond. B-DNA has a conservative CD spectrum above 220 nm with approximately equal positive (275 nm) and negative (245 nm) components centered around 260 nm [112].

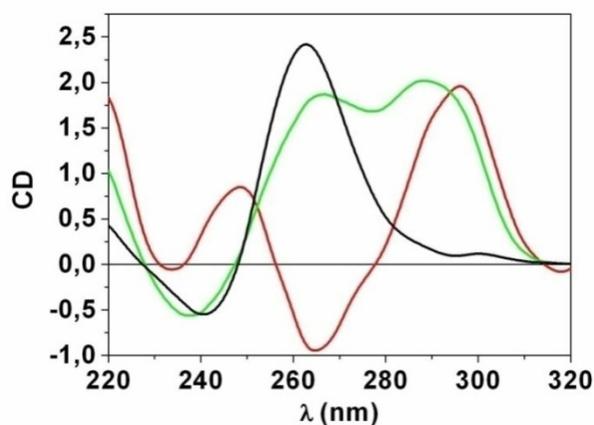


**Figure 2.5** Circular dichroism spectra above 200 nm for right-handed A-RNA and A-DNA, right-handed B-DNA, and left-handed Z-DNA (units are per M/cm/mol of nucleotide) (adapted from ref.113).

The B-DNA maximum has less than one-half the magnitude of the A-DNA maximum. Of course, the exact shapes and magnitudes of the CD spectra will depend on the base sequences, but the overall patterns will remain constant. The Z-DNA has a conservative spectrum above 240 nm with approximately equal negative (290 nm) and positive (260 nm) components centered around 280 nm. In summary, CD is most useful technique to compare DNA or RNA conformations and to detect changes when the solvent or temperature is changed [113], for this reason is extremely useful in conformational studies of G-quadruplex conformation in which the CD features are mainly influenced by stacking interactions between adjacent G-quartets [114]. The titration of quadruplexes with G4-ligands could induce change in the CD spectrum of quadruplexes and this change could be correlated to properties of the ligand under investigation. Although the topology of the folding of G4-DNA strands is very complex and many types of quadruplexes have been reported, there are two basic types of CD spectra, which have been associated to a single specific difference in the features of the strand folding, the relative orientation of the strands, parallel (all strands have

the same 5' to 3' orientation) or antiparallel. The spectra of “parallel” quadruplexes (a typical example is represented by the tetramolecular quadruplex formed by  $[d(TGGGGT)]_4$ , in which four strands with all glycosyl bonds in anti conformation run parallel to each other) have a dominant positive band at 264 nm, and a negative peak at 240 nm (Figure 2.6) [115,116]. On the contrary, the spectra of “antiparallel” quadruplexes (a typical example is represented by two folded-back strands of  $d[AG_3(T_2AG_3)_3]$  quadruplex in  $Na^+$  solution, where guanines alternate syn and anti glycosyl conformation along each strand) have a negative band at 264 nm and positive band at 290 nm (Figure 2.6) [117].

Interestingly, the hybrid-type (3+1) human telomeric G-quadruplexes contain both the alternating and non-alternating glycosidic conformations for adjacent guanines on a G-strand (a typical example is represented by  $d[AG_3(T_2AG_3)_3]$  quadruplex in  $K^+$  solution), and thus exhibit a hybrid CD profile, with a positive peak around 290 nm and a smaller peak around 264 nm, and a smaller negative peak at 240 nm. The positive peak around 290 nm is likely due to the alternating guanine glycosidic conformations along G-strands between the top and middle G-tetrads in the hybrid-type telomeric G-quadruplexes, while the positive peak around 264 nm and the negative peak around 240 nm are likely due to the non-alternating guanine glycosidic conformations between the middle and bottom G-tetrads (Figure 2.6) [118].



**Figure 2.6** CD spectra for the  $d[AG_3(T_2AG_3)_3]$  quadruplex in  $Na^+$  solution (red line) and the  $K^+$  solution (green line); CD spectrum of the  $[d(TGGGGT)]_4$  quadruplex in the  $K^+$  solution (black line).

### 2.3. Steady-State Fluorescence

Fluorescence spectroscopy has long been one of the most useful biophysical techniques available to scientists studying the structure and the function of biological macromolecules. Fluorescence is a spectroscopic method of analysis where the molecules are excited by irradiation at a certain wavelength and emit radiation of a different wavelength. The emission spectrum provides information for both qualitative and quantitative analysis. As shown in Figure 2.7, when light of an appropriate wavelength is absorbed by a molecule (excitation), the electronic state of the molecule changes from the ground state to one of many vibrational levels in one of the excited electronic states. The excited electronic state is usually the first excited singlet state,  $S_1$  (Figure 2.7). Once the molecule is in this excited state, relaxation can occur via several processes. Fluorescence is one of these processes and results in the emission of light.

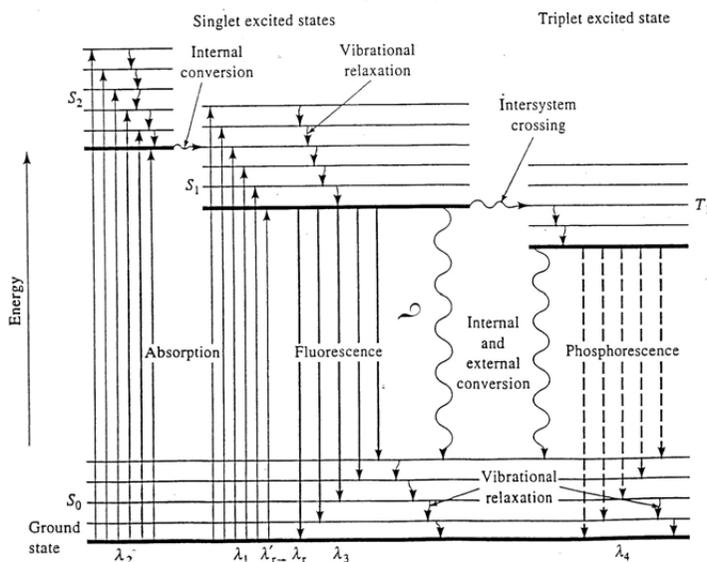


Figure 2.7 Electronic transition energy level diagram.

In order to study a molecule by fluorescence spectroscopy it needs to contain a fluorophore. Fluorophores are divided into two classes: intrinsic and extrinsic. Intrinsic fluorophores are those that occur

naturally. Extrinsic are those added to a sample. In proteins the dominant fluorophore is the indole group of tryptophan. The emission spectra of proteins are sensitive environment. Tryptophan residues are highly sensitive to the polarity of environment, which affects the energy levels of the first excited state with the result that the emission maximum can range from 330 nm in hydrophobic environment to 355 in water. DNA typically doesn't display intrinsic fluorescence, DNA contains nitrogenous bases that look like fluorophores but is weakly or not fluorescent. However a wide variety of dyes spontaneously bind to DNA such as acridines. For this reason staining of cells with dyes that bind to DNA is widely used to visualize and identify chromosomes. Measurement of fluorescence has been widely used to quantify the association reactions between DNA and fluorescent ligands. The observed fluorescence intensity is considered as the sum of the weighted contributions from free ligand and ligand bound to DNA:

$$F = (1 - \alpha_b)F_0 + \alpha_b F_b \quad ; \quad \alpha = \frac{I_\lambda - I_\lambda^{free}}{I_\lambda^{bound} - I_\lambda^{free}} \quad (\text{eq.2})$$

where  $F$  is the observed fluorescence intensity at each titrant concentration,  $F_0$  and  $F_b$  are the respective fluorescence intensities of the initial and final states of titration, and  $\alpha_b$  is fraction of bound ligand at each point of the titration and can be calculated, following the changes of fluorescence intensity at the wavelength of the maximum of emission.  $I_\lambda$  represents fluorescence intensity at the wavelength of the maximum of emission at the different DNA concentrations investigated,  $I_\lambda^{free}$  represents the fluorescence intensity of the drug alone and  $I_\lambda^{bound}$  is the fluorescence intensity signal of the saturated sample.

Assuming 1:1 stoichiometry for the interaction, it can be shown that:

$$[L]_0 \alpha_b^2 - ([L]_0 + [Q] + 1/K_b) \alpha_b + [Q] = 0 \quad (\text{eq.3})$$

where  $K_b$  is the binding constant,  $[L]_0$  is the total ligand concentration, and  $[Q]$  is the added DNA concentration. From combination of previous equations, it can be shown that:

$$\Delta F = \left( \frac{\Delta F_{max}}{2[L]_0} \right) \{ ([L]_0 + [Q] + 1/K_b) + \sqrt{([L]_0 + [Q] + 1/K_b)^2 - 4[L]_0[Q]} \} \quad (\text{eq.4})$$

where  $\Delta F = F - F_0$  and  $\Delta F_{max} = F_{max} - F_0$  and  $F$  and  $F_0$  are the initial and subsequent fluorescence intensities of the ligand at wavelength of maximum emission, upon DNA addition [119].

Finally as I already mentioned, the intrinsic fluorescence of most nucleic acids is too low for structural studies. Therefore, in recent studies a suitable fluorescent reporter group is incorporated to impart fluorescence to the polynucleotide structure. Highly fluorescent adenine analogue, 2-aminopurine (2-AP), has been widely used as a probe for loop structures in quadruplex DNA. Its quantum yield is highly sensitive to environmental conditions and it can be used to probe the interaction of quadruplexes with small organic molecules. This purine can be incorporated in any quadruplex nucleic acid without changing the quadruplex conformation and only slightly affecting its stability. The 2-AP fluorescence intensity tends to increase with solvent exposure and decrease with base stacking. The overall fluorescence of a quadruplex changes if a small molecule interacts in the proximity of 2-AP. The 2-AP modified human telomeric quadruplexes have been used to monitor specific complex formation with porphyrin TMPyP4 [120]. 2-AP-modified quadruplexes provided information about the stoichiometry and the binding mode of ligands.

## 2.4. Docking and Molecular Dynamics

Molecular modelling methods are nowadays routinely used to investigate structure, dynamics, surface properties and thermodynamics of biomolecular complexes. Docking and molecular dynamics methods are used to predict binding modes and affinities of G-quadruplexes and small organic molecules.

*Docking* is the process by which two molecules fit together in 3D space. This methodology is often used to predict protein-protein, protein-DNA (or RNA), and DNA (or RNA)-small molecule interactions. Predictive docking procedures start from the individually determined (unbound) structures of two molecules and they aim to predict the complex structure. All docking programs contain a scoring function to discriminate between near-native docked orientations and incorrect orientations, and a search algorithm that is designed to sample rapidly all possible docking orientations. Even though structural flexibility is mostly restricted to surface side chains [121], the innate flexibility renders

docking extremely difficult. The algorithm by Abagyan and colleagues [122,123] allows for surface side chain flexibility, however, such algorithms typically require hundreds of hours for each complex [122]. The alternative is to adopt the *rigid-body* approach, which samples only the six rotational/translational degrees of freedom, but uses target functions that are tolerant to some overlap of the two molecules being matched. Affinity docking module in INSIGHT II package ([www.accelrys.com](http://www.accelrys.com)) was used. In docking, the interaction energy is computed by summing the energy contributions between all atoms of the two molecules. The contribution between atoms interacting with other atoms in the same molecule is ignored. The objective of a docking type calculation is to evaluate the interaction energies of many orientations of one molecule relative to the other, while searching for the orientations that result in low interaction energies.

As proposed by Langridge [124] an energy grid approximating the larger of the two molecules can be precomputed. Since the interaction energy can now be approximated by calculating the energy between the atom of the moving molecule and the appropriate grid points, the docking can be done in real time.

The construction of energy grids depends on the forcefield in use.

The CFF (Cornell Force Field) forcefields use the nonbonded interaction potential

$$E_{ij} = \sum_{i < j} \frac{A_{ij}}{R_{ij}^9} \frac{B_{ij}}{R_{ij}^6} \quad (\text{eq.5})$$

where  $A_{ij}$  and  $B_{ij}$  are parameters with units of kcal mol<sup>-1</sup> angstrom<sup>9</sup> and kcal mol<sup>-1</sup> angstrom<sup>6</sup> respectively and  $R_{ij}$  is the distance between atoms  $i$  and  $j$  in  $\text{\AA}$ . A completely equivalent representation is:

$$E_{ij} = \varepsilon_{ij}^* \left[ 2 \left( \frac{R_{ij}^*}{R_{ij}} \right)^9 - 3 \left( \frac{R_{ij}^*}{R_{ij}} \right)^6 \right] \quad (\text{eq.6})$$

where  $\varepsilon_{ij}^*$  is the potential well depth in kcal mol<sup>-1</sup> and  $R_{ij}^*$  is the interatomic distance in angstroms at which the minimum occurs. The conversion between the two representations is straightforward:

$$\varepsilon_{ij}^* = \frac{4B_{ij}^3}{27A_{ij}^2} \quad ; \quad A_{ij} = 2\varepsilon_{ij}^* R_{ij}^9$$

$$R_{ij}^* = \left[ \frac{3A_{ij}}{2B_{ij}} \right]^{\frac{1}{3}}; \quad B_{ij} = 3\varepsilon_{ij}^* R_{ij}^6$$

The docking module of INSIGHT II provides facilities for calculating the nonbond energy between two molecules using the methods described before. This calculation is performed using the Affinity program module in INSIGHT II. A key step before starting Affinity is to define a binding site subset. The binding site is defined to include important residues suggested by other studies. Several approaches for non bonds are used: grid based approach, hydrogen bonding contacts, tethering atoms, and confining ligands. For the hydrogen bonding contacts feature, the most relevant subsets of atoms that are potential hydrogen bond partners are defined manually by picking them with the mouse.

Then the general strategy involves a two-step process. First, use the Monte Carlo minimization approach to obtain reasonably placed ligands. The second phase of the docking involves a more refined approach. First, import the initial structures generated using Monte Carlo minimization. Then, use a more realistic nonbond like the Cell\_Multipole method. The starting structures have reasonable nonbond contacts, after a brief minimization (100 steps), a Simulated Annealing is perform over a reasonable number (50) of stages. These processes are repeated for all the relevant subsets involved in the hydrogen bonding contacts. The last step is to perform molecular dynamics.

*Molecular Dynamics* (MD) is a computer simulation technique in which the time evolution of a set of interacting atoms is computed by integrating their equations of motion. Therefore, MD allows us to view how a molecular system evolves through time and to derive average properties of the system, given a simulation of sufficient length.

The equations of motion can only be solved numerically for a multi-body problem. To calculate the dynamics of the system (the position of each atom as a function of time), Newton's classical equations of motion are solved for each atom given an empirical force field:

$$F_i = m_i a_i \tag{eq.7}$$

where  $F_i$  is the force exerted on particle  $i$ ,  $m_i$  is the mass of particle  $i$  and  $a_i$  is the acceleration of particle  $i$ . The force on each atom is the negative of the derivative of the potential energy (V) with respect to the position of the atom ( $r_1, r_2, \dots, r_N$ ):

$$F_i = -\frac{\partial V}{\partial r_i} \quad (\text{eq.8})$$

Once the coordinates of the atoms of a starting structure and their velocities are defined, the force acting on each atom can be calculated for each point in time  $t + dt$  and a new set of coordinates can be generated. The repetition of this procedure generates a molecular trajectory corresponding to the time-dependent fluctuations of the atomic positions. The accuracy of the simulations is directly related to the potential energy function that is used to describe the interactions between particles. In MD a classical potential energy function is used that is defined as a function of the coordinates of each of the atoms. The potential energy function is separated into terms representing covalent interactions and non covalent interactions. The covalent interactions may be described by the following terms:

$$V_{bond} = \sum_{i=1}^{N_b} \frac{1}{2} k_i^b (r_i - r_{0,i})^2 \quad (\text{eq.9})$$

$$V_{angle} = \sum_{i=1}^{N_\theta} \frac{1}{2} k_i^\theta (\theta_i - \theta_{0,i})^2 \quad (\text{eq.10})$$

$$V_{dihedral} = \sum_{i=1}^{N_\Phi} \frac{1}{2} k_i^\Phi \cos(n_i(\Phi_i - \Phi_{0,i})) \quad (\text{eq.10})$$

$$V_{bond} = \sum_{i=1}^{N_\xi} \frac{1}{2} k_i^\xi (\xi_i - \xi_{0,i})^2 \quad (\text{eq.11})$$

the equations correspond to two, three, four and four body interactions, respectively. These interactions are represented by harmonic potentials for the bond lengths  $r_i$ , for the bond angles  $\theta_i$ , for the improper dihedral (out of the plane) angle  $\xi_i$  and by a more complex potential for the dihedral angles  $\Phi_i$ . The non-covalent (non-bonded) interactions, which correspond to interactions between particles separated by more than three covalent bonds, are usually described by Coulomb's law:

$$V_{Coulomb} = \sum_{i < j} \frac{1}{4\pi\epsilon_0\epsilon_r} \frac{q_i q_j}{r_{ij}} \quad (\text{eq.12})$$

for the electrostatic interactions (for a pair of atoms carrying the partial charges  $q_i$  and  $q_j$ ), and by a Lennard-Jones potential:

$$V_{LJ} = \sum_{i < j} \frac{A_{ij}}{r_{ij}^{12}} \frac{B_{ij}}{r_{ij}^6} \quad (\text{eq.13})$$

for the van der Waals interactions, where  $r_{ij}$  is the atomic distance between particles  $i$  and  $j$ .

The force field parameters describe the strength of the interactions. For bonded interactions parameters are defined for bond stretching, bond bending and torsional rotation. Another set of parameters determines the strength of non-bonded electrostatic and van der Waals interactions. Electrostatic interactions are generally represented by point charges located at the center of the atom. It should be kept in mind, however, that MD is affected by several limitations. Firstly, MD is computationally very demanding and the computational load scales with the square of the system size. Simulation times are currently limited to hundreds of nanoseconds or a few microseconds at most. The phenomena that can be explored must occur with sufficient statistical significance within time scales that are encompassed by the computation.

How the system evolves through time is specified by the force field and by an integration time step that determines where the atoms will be positioned at time  $t + dt$ . MD requires the use of a very small time-step (1-2 fs) to achieve accurate results, because small time-steps limit the approximations that are introduced by the numerical integrator. This limits the overall scope of the simulated time and the computable properties.

According to statistical thermodynamics, physical quantities are represented by averages over configurations belonging to a certain statistical ensemble. A trajectory obtained by molecular dynamics provides such an ensemble. Therefore, a measurement of a physical quantity by simulation is simply obtained as an arithmetic average of the various instantaneous values adopted by that quantity during the MD run. Statistical thermodynamics is the link between the microscopic ensembles and the macroscopic properties. In the limit of an exact force field and very long simulation times, one could expect the phase space to be fully sampled and in that limit the averaging process would yield exact thermodynamic properties. In practice, MD runs are always of finite length and one should exert caution when judging the sampling quality. An important constraint in deriving average properties is to extract configurations only from an ensemble at thermal equilibrium. Therefore, MD simulations start generally with an equilibration phase. Once equilibrium is reached, the simulation enters the production phase. The production run should be long enough to sample the property of interest with sufficient statistical significance.

Molecular modeling and dynamics studies are very important for the recognition of a quadruplex motif, formed by human telomeric sequence. These studies enabled the improvement of a ligands, which possess a

moderate affinity for quadruplexes and no selectivity over double-stranded DNA, by developing a derivative, which could represent a potent quadruplex binder. The first step of most computational studies involves manual docking of the ligand under investigation onto a quadruplex and it is imperative that one must carefully consider whether the chosen binding site is in agreement with known data about this ligand.

## CHAPTER 3 TMPyP4 and human telomeric G-quadruplexes

### 3.1. Interaction between TMPyP4 and four-repeats human telomeric sequences

Recently cationic porphyrins have emerged as a promising class of quadruplex-binding molecules with potential applications as anti-tumor therapeutic agents [125,126]. In particular, the cationic *meso*-tetrakis-(*N*-methyl-4-pyridyl)-porphyrin (TMPyP4) has been the subject of extensive investigations since it can induce telomerase inhibition upon binding to telomeric DNA quadruplexes [127]. Interestingly, it was found that TMPyP4 can induce the formation of a DNA quadruplex from a single-stranded oligonucleotide [128,129,130] The geometric features of the TMPyP4 ring (Figure 3.1) have been assumed to be appropriate to interact with the G-tetrads of quadruplex structures through  $\pi$ - $\pi$  stacking interactions, however, the structure of TMPyP4-quadruplex complexes, as well as the binding stoichiometry, continue to be controversial.

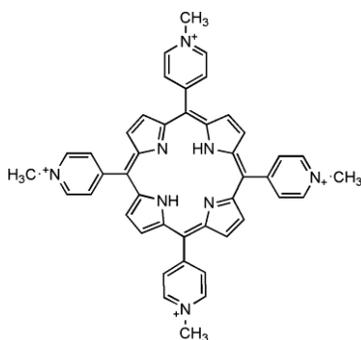


Figure 3.1 Chemical Structure of TMPyP4.

Several different models of TMPyP4 binding to quadruplexes have been proposed on the basis of structural, spectroscopic, calorimetric or computational studies. Features of the various models include

intercalative binding of the drug between two adjacent G-tetrads and/or stacking with the external G-tetrads or with the loops of the quadruplexes [131,132,133,134,135,136]. This is further complicated by the intrinsic structural polymorphism of human telomeric quadruplex sequences.

The binding stoichiometry of TMPyP4 to quadruplex structures is itself not fully established, especially for the quadruplexes from human telomeric sequences. For example, Hurley and co-workers determined by UV absorption spectroscopy a stoichiometric ratio of 2:1 for the binding of TMPyP4 to the 22-mer d[AGGG(TTAGGG)<sub>3</sub>] quadruplex in 100 mM Na<sup>+</sup> buffer [127], while, more recently, Wei *et al.* using the same technique determined a 4:1 ratio for the same interaction in identical solution conditions [134]. Haq *et al.* have investigated by ITC the stoichiometry and thermodynamics for binding of TMPyP4 to the d[AGGG(TTAGGG)<sub>3</sub>] quadruplex structure in two different buffered solutions containing either K<sup>+</sup> or Na<sup>+</sup> at a fixed ionic strength (200 mM) [131]. The binding stoichiometry was the same for the two different solutions, in agreement with the 2:1 stoichiometry value previously determined for the human sequence by Hurley and co-workers [127]. However, the binding behavior in the Na<sup>+</sup> buffer was found to be considerably different compared to the results obtained in K<sup>+</sup> solution. Indeed, the interaction with d[AGGG(TTAGGG)<sub>3</sub>] in K<sup>+</sup> was characterized by a single binding event, whereas the binding in Na<sup>+</sup> solution was characterized by an initial binding event requiring only low [TMPyP4]/[DNA] ratios to reach saturation, followed by a secondary process requiring a much higher ligand concentration for saturation. This data has shown that the interactions can be affected by the ionic environment used in the experiment, which may be one of the key factors determining the binding mode of TMPyP4 and the stoichiometry of the complexes, as also observed for other ligands [28]. Zhang *et al.* have recently characterized the binding of TMPyP4 to the human telomeric G-quadruplex formed by d[TAGGG(TTAGGG)<sub>3</sub>T] for a range of K<sup>+</sup> concentrations [128]. They found that, in the presence of 0, 10 and 100 mM K<sup>+</sup>, TMPyP4 forms complexes with a drug-quadruplex molar ratio of 5, 5 and 3, respectively. They also suggested that an increase in K<sup>+</sup> concentration would reduce the binding affinity of TMPyP4 and that it can transform the K<sup>+</sup> induced hybrid-type quadruplex structure into an anti-parallel one. The latter finding was also suggested by Gaynutdinov *et al.* using <sup>125</sup>I-radioprobng for a similar DNA sequence labeled with <sup>125</sup>I [137]. Indeed, they showed that d[GGG(TTAGGG)<sub>3</sub> <sup>125</sup>I-CT] is a mixture of two or more conformations in K<sup>+</sup> solution, but adding TMPyP4 stabilizes a basket conformation. They also inferred that the same sequence adopts a basket conformation in the presence of Na<sup>+</sup> and

the drug addition does not change the conformation of the quadruplex. This disagrees with a recent study by Gray *et al.* which shows that TMPyP4 addition promotes the conversion of d[AGGG(TTAGGG)<sub>3</sub>] from the basket to the hybrid form [78]. All the previous studies do not take into account binding of TMPyP4 to the parallel structure, possibly because it may only be present in dilute solution as a minor component. However, it has been shown that molecular crowding induces a structural transition of human telomeric G-quadruplexes to a parallel-stranded structure [63]. Moreover, in a very recent paper, Wei *et al.* studied the interaction of the d[AGGG(TTAGGG)<sub>3</sub>] sequence with TMPyP4, in both dilute solution and crowding condition [138]. They found that the crowding agent (PEG 200) induces the formation of a parallel-stranded quadruplex structure and two independent binding sites for the TMPyP4 were found. It is therefore apparent that the binding mode of TMPyP4 to quadruplex molecules depends on the chemical features of the porphyrin, on the structure and composition of the host DNA, and on the environment used in the experiment. In this thesis is reported the use of Isothermal Titration Calorimetry (ITC) and Circular Dichroism (CD) in a systematic study of the interaction between TMPyP4 and G-quadruplex molecules formed by four different sequences with 5'- or 3'-flanking bases (Table 3.1), in both dilute solution and under molecular crowding conditions, to shed light on the recognition processes involving the human telomeric G-quadruplex structures and TMPyP4, and to evaluate the effect of the flanking bases addition. In order to investigate the selectivity of TMPyP4 for the quadruplex molecules relative to duplex, a study on the interaction between the drug and a duplex-forming oligonucleotide (Table 3.1), was also performed.

All the oligonucleotide sequences used for this study were purchased from the Primm Company (Milan, Italy). Quadruplex and duplex samples have been prepared by dissolving the lyophilized compound in two different buffers, the first one containing 20 mM phosphate buffer with 70 mM KCl, 0.1 mM EDTA at pH 7.0, and the second one, used to reproduce the molecular crowding conditions, consisting of 20 mM phosphate buffer with 150 mM KCl, 0.1 mM EDTA, 40% (w/v) of PEG 200 at pH 7.0. The concentration of the dissolved oligonucleotides was evaluated by UV measurement at 95 °C, using as molar extinction coefficient the value calculated at 260 nm, by the nearest-neighbor model (Table 3.1) [139]. The solutions of TMPyP4 (Sigma-Aldrich, USA) have been prepared in the same buffers used for the oligonucleotides, and the concentration has been estimated by UV spectroscopy using the reported extinction coefficient at 424 nm (226000 M<sup>-1</sup>cm<sup>-1</sup>).

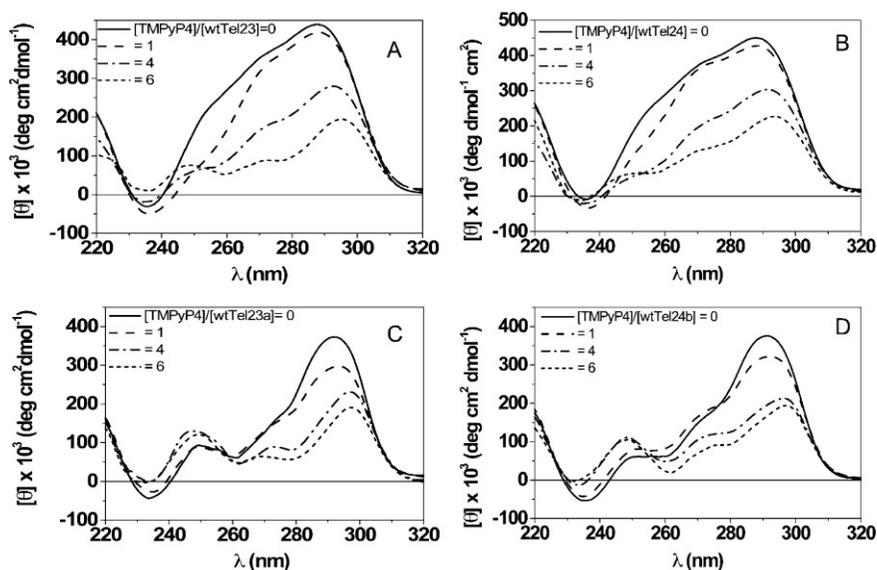
Sequences	$\epsilon_{260 \text{ nm}}^a$ $\text{M}^{-1}\text{cm}^{-1}$	Name <sup>b</sup>	Major conform- ation in dilute solution	Major conformation <sup>c</sup> under molecular crowding condition
TTAGGGTTAGGGTTAGGGTTAGGG	244600	wtTel24	hybrid-1	Parallel
TAGGGTTAGGGTTAGGGTTAGGG	236500	wtTel23	hybrid-1	Parallel
AGGGTTAGGGTTAGGGTTAGGGT	237000	wtTel23a	hybrid-2	Parallel
AGGGTTAGGGTTAGGGTTAGGGTT	245100	wtTel24b	hybrid-2	Parallel
CGCGAATTCGCG	110700	duplex	duplex	duplex

**Table 3.1** Thermodynamics of the TMPyP4 interaction (<sup>a</sup>The error on the extinction coefficients is less than 4%, <sup>b</sup>Names reflect the convention used by Dai J. et al.[118], <sup>c</sup>Evaluated from the CD spectrum).

### 3.1.1. Results

#### 3.1.1.a. CD experiments

CD experiments were carried out to analyze the structural changes of the quadruplex samples upon TMPyP4 interaction. In the absence of 40% PEG 200, the sequences investigated show the typical CD spectrum of the hybrid conformation (Figure 3.2), with a maximum around 290 nm, a shoulder centered around 270 nm and a weak minimum around 240 nm [58]. The sequences wtTel23 and wtTel24 produce a stronger shoulder in the CD spectrum at 270 nm compared to the sequences wtTel23a and wtTel24b. The first two sequences have been shown to form in solution the hybrid-1 structure as major conformation, while the last two sequences form the hybrid-2 structure as major conformation [118].

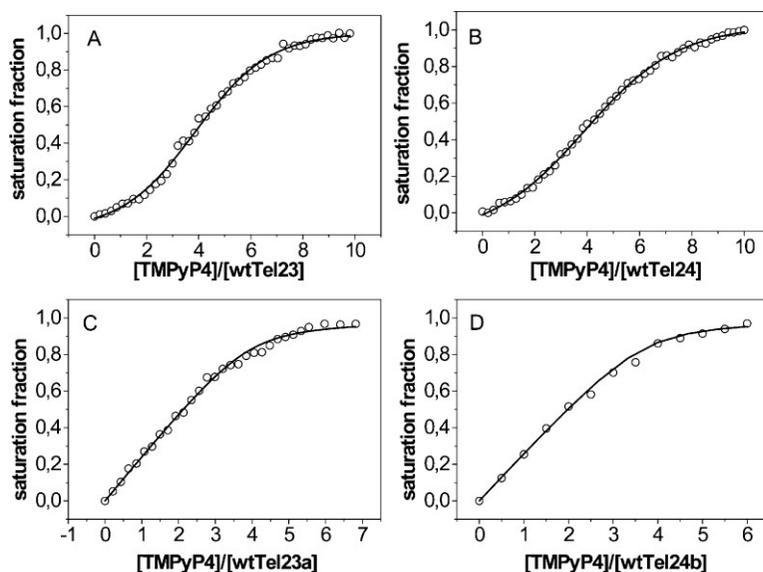


**Figure 3.2** CD spectra of the quadruplexes in  $K^+$  solution. wtTel23 (A), wtTel24 (B), wtTel23a (C) and wtTel24b (D). Each panel shows spectra at increasing molar ratios of TMPyP4. The spectra, recorded at the higher TMPyP4 concentration, are representative of the final complexes deriving from the interaction between the quadruplex samples and the porphyrin molecules.

The more intense signal at 270 nm in the spectra of the sequences wtTel23 and wtTel24 could be due to the presence of a significant amount of a parallel form in solution. Upon TMPyP4 addition, dramatic changes in the CD spectra are observed, suggesting that the TMPyP4 molecule causes a conformational change in the quadruplex structures. In particular, the CD spectrum of each sample evolves toward the characteristic CD spectrum of an antiparallel quadruplex structure, with a maximum centered at 290 nm [118].

In order to obtain the stoichiometry of the complexes and to get a deeper understanding of the TMPyP4-quadruplex interaction, CD titration experiments were performed. CD spectra were collected at different TMPyP4/quadruplex molar ratios to calculate the saturation fractions (see above) of the quadruplex at different concentrations of TMPyP4. Binding curves for all the investigated systems were obtained by plotting the saturation fractions as a function of the molar ratio (Figure 3.3). Intriguingly the sequences wtTel23 and wtTel24, forming the hybrid-1 structure as major conformation in solution [118], show sigmoid-shaped binding curve centered on a stoichiometry of four (Figure 3.3, panels A

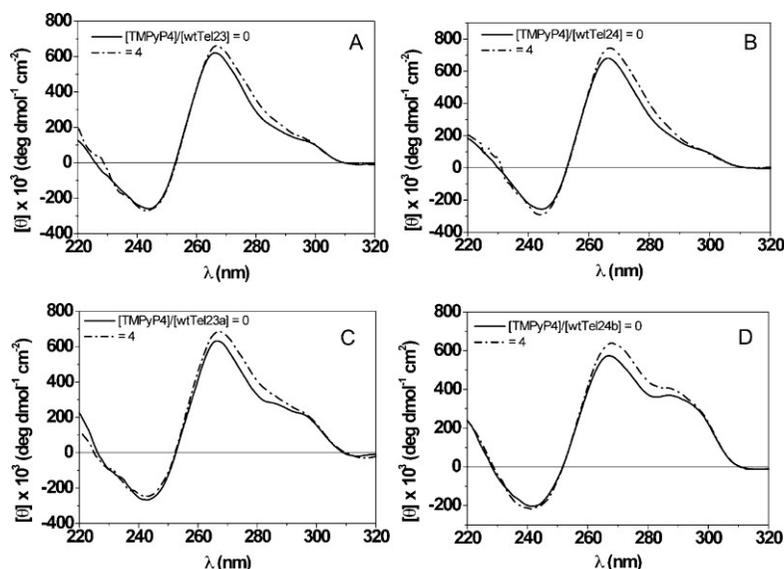
and B). On the other hand the sequences wtTel23a and wtTel24b, forming the hybrid-2 structure as major conformation, show a hyperbolic binding curve, suggesting a stoichiometry of four bound TMPyP4 molecules (Figure 3.3, panels C and D). The different shape of the binding curves suggests a different mechanism of interaction of TMPyP4 with the quadruplexes obtained with the sequences wtTel23 and wtTel24 compared to the ones formed with the sequences wtTel23a and wtTel24b. Particularly, a sigmoid-shaped rather than a hyperbolic binding curve indicates a cooperative mechanism.



**Figure 3.3** CD titration experiments on the interaction between TMPyP4 and the quadruplex structures formed by the sequences wtTel23 (A), wtTel24 (B), wtTel23a (C) and wtTel24b (D). The open circles represent the experimental data obtained following the change of the CD signal at 290 nm, upon TMPyP4 addition. The graph line represents the theoretical curve obtained on the basis of a cooperative-sites model (A, B) or equivalent-sites model (C, D).

On this basis an interpolation analysis has been carried out using two different models for the two classes of binding curves. The experimental binding curves of the sequences wtTel23 and wtTel24 are well modelled to an equation describing cooperative-sites, while the binding curves obtained with the sequences wtTel23a and wtTel24b have been modelled to an equation assuming equivalent-sites. The results of the interpolation

analysis clearly indicate that all of these quadruplexes are able to bind up to four TMPyP4 molecules; intriguingly, the sequences wtTel23a and wtTel24b possess a ten-fold higher affinity ( $K_b = 1.4 \times 10^6 \text{ M}^{-1}$  for both systems) for TMPyP4 compared to the sequences wtTel23 and wtTel24 ( $K_b = 1.0 \times 10^5 \text{ M}^{-1}$  and  $K_b = 1.5 \times 10^5 \text{ M}^{-1}$ , respectively). It is worth noting that the  $K_b$  values obtained are indicative of the global events and they take into account both binding and conformational change. In the presence of 40% PEG 200, the CD spectra of the quadruplex-forming oligonucleotides show dramatic differences compared to the ones previously observed (Figure 3.4).

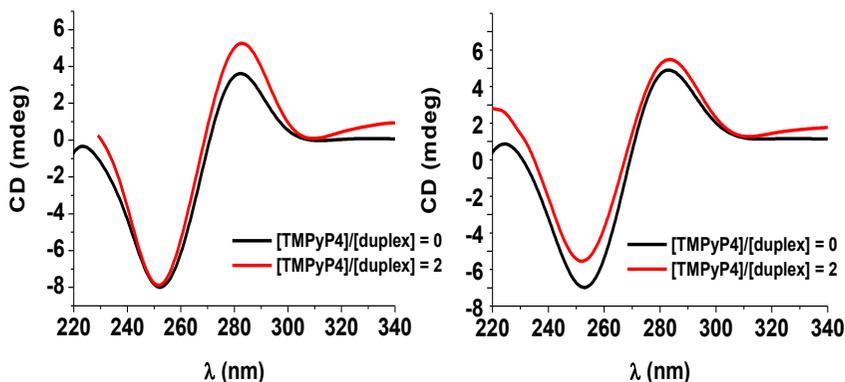


**Figure 3.4** CD spectra of the quadruplexes under molecular crowding condition. wtTel23 (A), wtTel24(B), wtTel23a (C) and wtTel24b (D). Each panel shows the spectrum of the quadruplex structure in absence of TMPyP4 and in presence of a high molar ratio of TMPyP4.

Interestingly, all the spectra show a well-defined maximum around 268 nm and a minimum at 241 nm, suggesting that a higher fraction of the oligonucleotides fold in a parallel quadruplex structure. A shoulder, centered on 290 nm, is also present and its intensity becomes higher as follows: wtTel23  $\approx$  wtTel24 < wtTel23a < wtTel24b, those results are consistent with the possible presence of additional minor conformations. Interestingly, under molecular crowding condition, the addition of

TMPyP4 does not cause dramatic changes in the CD spectra. The global shape of the spectra is conserved and the molar ellipticity becomes slightly higher, since, probably, the TMPyP4 interaction does not cause a conformational change of the quadruplex structure under these conditions. Since the changes in the CD signals, upon TMPyP4 addition, are small they did not allow to obtain a binding curve as was possible in the absence of 40% PEG 200.

The spectroscopic characterization of the TMPyP4 interaction was also carried out for the duplex, in both dilute solution and under molecular crowding conditions. The CD spectra of the duplex-forming oligonucleotide show the same profile in both conditions (Figure 3.5), suggesting that the duplex structure is not affected by the presence of the molecular crowding agent. The addition of TMPyP4 does not change the shape of the spectra, suggesting that the TMPyP4 interaction does not promote any conformational change in the duplex DNA structure.



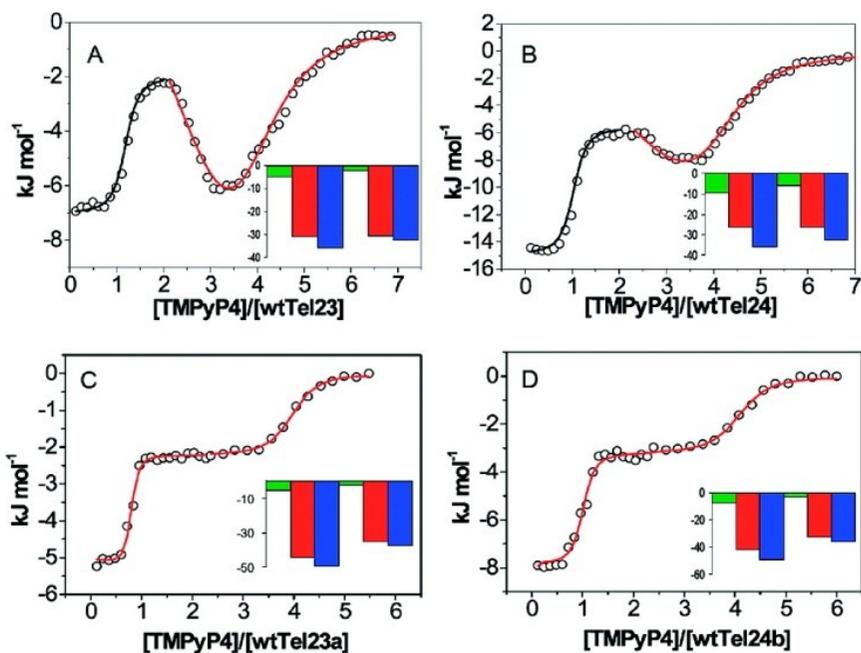
**Figure 3.5** CD spectra of the duplex in dilute solution (on the right) and under molecular crowding condition (on the left), in absence and in presence of TMPyP4.

### 3.1.1.b. ITC experiments

To better address the differences between the binding properties shown by the studied sequences, both in dilute solution and under molecular crowding condition, a calorimetric analysis of the porphyrin interaction has been carried out by means of isothermal titration experiments. In dilute solution, the results collected for the two classes of sequences,

forming the hybrid-1 or hybrid-2 structure, are quite different. The binding curves related to the sequence wtTel23 and wtTel24 show more complicated profiles compared to the sequences wtTel23a and wtTel24b. All the normalized heat binding curves have biphasic profiles, with two separate binding events centered on two different stoichiometries (Figure 3.6). In the first event, one molecule of TMPyP4 interacts with the quadruplex structure followed by the simultaneous binding of three more TMPyP4 molecules, allowing the formation of a final complex composed of one molecule of quadruplex with four molecules of TMPyP4. The calorimetric binding curves of the sequences forming the hybrid-1 structure have been analyzed assuming that the two binding events are independent of each other. On the basis of the previous assumption, the first event has been interpolated with a simple independent and equivalent-sites model (Figure 3.6 panels A and B, black line graph) whereas the second event has been analyzed with a cooperative-sites model (Figure 3.6 panels A and B, red line graph).

This analysis reveals that the first event ( $K_b = 2 \times 10^6 \text{ M}^{-1}$  for both sequences) is about ten-fold stronger than the second one ( $K_b = 5 \times 10^5 \text{ M}^{-1}$  for both sequences). The binding curves obtained for the sequences wtTel23a and wtTel24b (Figure 3.6, panels C and D) are well fitted using a multiple-sites model and the analysis suggests that the first binding event ( $K_b = 5 \times 10^8 \text{ M}^{-1}$  for wtTel23a and wtTel24b) is about 100-fold stronger than the second one ( $K_b = 4 \times 10^6 \text{ M}^{-1}$  and  $K_b = 2 \times 10^6 \text{ M}^{-1}$  for wtTel23a and wtTel24b, respectively). The high quality of the calorimetric measurements allowed to obtain a complete thermodynamic characterization for all the investigated systems (Table 3.2).



**Figure 3.6** Normalized heat of interaction between TMPyP4 and the quadruplex structures formed by the sequences wtTel23 (A), wtTel24 (B), wtTel23a (C) and wtTel24b (D). The open circles represent the experimental data obtained by integrating the raw data and subtracting the heat of ligand dilution into the buffer. The lines represent the best fit obtained by a non-linear least-squares procedure based on three different interpolating models: independent sites (black line, panels A and B), cooperative sites (red line, panels A and B) and multiple sites (red line, panels C and D). Each panel shows the thermodynamic signature of the corresponding interaction. The enthalpic contribution ( $\Delta H^\circ$ ) is shown in green, the entropic one ( $-T\Delta S^\circ$ ) in red and the Gibbs energy change ( $\Delta G^\circ_{298K}$ ) in blue.

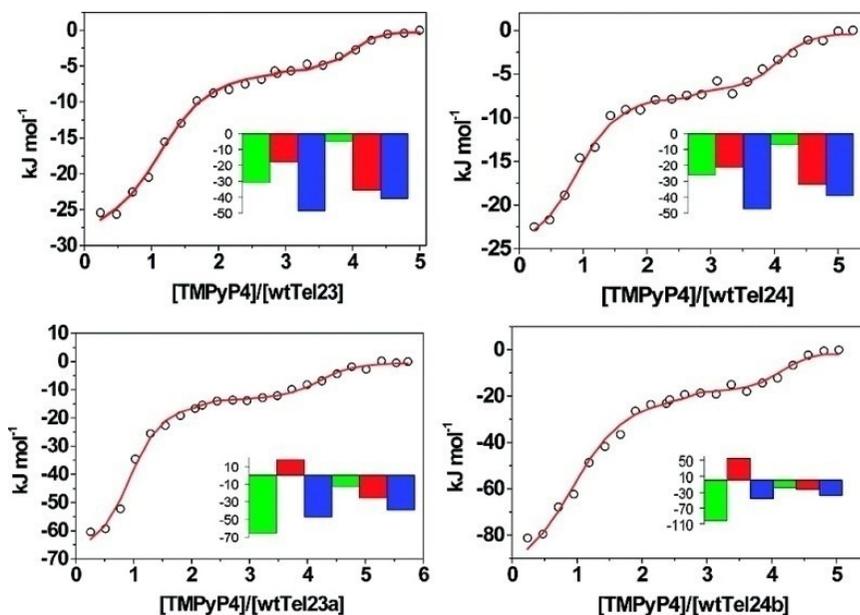
In dilute solution										
	First event					Second Event				
	$n_1$	$K_b$ ( $M^{-1}$ )	$\Delta H^\circ$ ( $kJ\ mol^{-1}$ )	$-T\Delta S^\circ$ ( $kJ\ mol^{-1}$ )	$\Delta G^\circ_{298K}$ ( $kJ\ mol^{-1}$ )	$n_2$	$K_b$ ( $M^{-1}$ )	$\Delta H^\circ$ ( $kJ\ mol^{-1}$ )	$-T\Delta S^\circ$ ( $kJ\ mol^{-1}$ )	$\Delta G^\circ_{298K}$ ( $kJ\ mol^{-1}$ )
wtTel23	1.1	$2 \times 10^6$	-5.0	-30.9	-35.9	3.0	$5 \times 10^5$	-2.0	-30.5	-32.5
wtTel24	0.9	$2 \times 10^6$	-9.4	-26.5	-35.9	3.0	$5 \times 10^5$	-6.0	-26.5	-32.5
wtTel23a	1.0	$5 \times 10^8$	-5.2	-44.4	-49.6	2.9	$4 \times 10^6$	-2.2	-35.4	-37.7
wtTel24b	0.9	$5 \times 10^8$	-7.9	-41.7	-49.6	2.9	$2 \times 10^6$	-3.3	-32.6	-35.9
duplex	1.0	$8 \times 10^5$	-54.0	20.3	-33.7					
Under molecular crowding conditions										
wtTel23	1.0	$3 \times 10^8$	-30.7	-17.9	-48.6	2.9	$1 \times 10^7$	-5.0	-35.7	-40.7
wtTel24	0.9	$2 \times 10^8$	-26.0	-21.4	-47.4	2.9	$6 \times 10^6$	-6.9	-32.0	-38.9
wtTel23a	0.9	$2 \times 10^8$	-65.2	17.8	-47.4	2.9	$4 \times 10^6$	-12.3	-25.8	-38.1
wtTel24b	1.0	$1 \times 10^8$	-102.5	55.6	-46.9	3.0	$7 \times 10^6$	-17.4	-21.8	-39.2
duplex	1.0	$6 \times 10^5$	-54.0	21.0	-33.0					

**Table 3.2** Thermodynamics of the TMPyP4 interaction (The experimental error for each thermodynamic property is < 5%.  $n_1$  and  $n_2$  are referred to the number of TMPyP4 molecules interacting in each binding event; the final stoichiometry of the complexes is  $n_{tot} = n_1 + n_2$ ).

The thermodynamic signature of the two binding events (Figure 3.6, inserted bar-graphs) for each quadruplex sample suggests that the TMPyP4 interaction is entropically driven and the small enthalpic contribution ( $\Delta H^\circ$ ) acts synergistically with the larger entropic one ( $-T\Delta S^\circ$ ). The thermodynamic characterization reveals some interesting features about the binding process. Particularly, the negative entropic contribution suggests that the TMPyP4 interaction produces a more relaxed structure with respect to the unbound quadruplex. Moreover, the small enthalpic contribution clearly indicates that the driving force of the binding process is not represented by the formation of new interactions between the binder and the target molecules, but it is principally due to the relaxation of the entire structure upon TMPyP4 interaction.

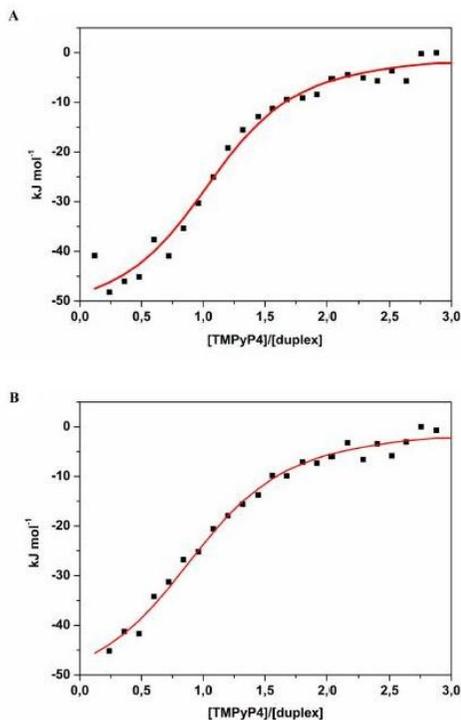
Under molecular crowding conditions, all the sequences show equivalent behaviour, producing similar binding profiles (Figure 3.7). Particularly, all the binding curves are biphasic revealing two well-defined events centered on the stoichiometry of one and four bound TMPyP4 molecules.

The interpolation procedure of the experimental data has been carried out with a multiple-sites model and the analysis suggests that the first binding event is about 100-fold stronger than the second one (Table 3.2). The thermodynamic characterization of the binding events (Table 3.2), summarized in the thermodynamic signature reported in Figure 3.7 (inserted bar-graphs), reveals that the two binding events are driven by different thermodynamic contributions. The first binding event is the strongest one and is enthalpically driven, while the second event, the weaker one, is entropically driven; those data clearly suggest that the two consecutive binding events occur by distinct mechanisms and probably they involve different regions of the quadruplex structure. In the first event the formation of new interactions is the key feature of the binding, and the opposing entropic contribution suggests that the resulting complex is more rigid with respect to the free quadruplex. On the other hand, the second binding event is centered on higher molar ratios of TMPyP4 and is entropically driven, suggesting that the resulting final complex is a more relaxed structure.



**Figure 3.7** Normalized heat of interaction between TMPyP4 and the quadruplex structures formed by the sequences wtTel23 (A), wtTel24 (B), wtTel23a (C) and wtTel24b (D) under molecular crowding condition. The open circles represent the experimental data obtained by integrating the raw data and subtracting the heat of ligand dilution into the buffer. The lines represent the best fit obtained by a non-linear least-squares procedure based on a multiple sites binding model (red line). Each panel shows the thermodynamic signature of the corresponding interaction. The enthalpic contribution ( $\Delta H^\circ$ ) is shown in green, the entropic one ( $-T\Delta S^\circ$ ) in red and the Gibbs energy change ( $\Delta G^\circ_{298K}$ ) in blue.

The ITC measurements have also been carried out on the duplex in both solution conditions (Figure 3.8). In both cases, the results of the calorimetric titrations show an exothermic interaction ( $\Delta H^\circ = -54 \text{ kJ mol}^{-1}$ ) with a stoichiometry of one molecule of TMPyP4 per duplex and similar binding constants ( $K_b = 8 \times 10^5 \text{ M}^{-1}$  and  $K_b = 6 \times 10^5 \text{ M}^{-1}$  in dilute solution and under molecular crowding condition, respectively).



**Figure 3.8** Normalized heat of interaction between TMPyP4 and duplex molecule in both A) dilute solution and B) under molecular crowding condition. The filled square represent the experimental data and the red line represents the best fit obtained by non linear least square based on the independent site model.

### 3.1.2. Discussion and Conclusions

Four different human telomeric truncated sequences have been chosen to characterize the interaction between the human quadruplex structures and TMPyP4, in dilute solution and under molecular crowding condition. All the oligonucleotide molecules contain the core sequence 5'-AGGG(TTAGGG)<sub>3</sub>-3', and are able to fold into a quadruplex structure, with differing numbers of flanking thymines either at the 5' or 3' terminal or at both termini. These oligonucleotides (Table 3.1) have been previously studied from a structural point of view and in dilute K<sup>+</sup> solution they form a mixture of at least, two conformations of intramolecular quadruplexes: the hybrid-1 and the hybrid-2 conformations [60,61]. The major conformation appears to be flanking nucleotide-dependent and intriguingly, the presence of thymines at the 3' end

region seems to drive their structures toward hybrid-2 structure formation [118]. The differing features between those structures are mainly due to the different loop arrangement and to the presence of two different cap-regions in the two structures. Particularly, the hybrid-1 conformation comprises G-quadruplex scaffold of three G-tetrad planes, with three interconnecting loops and a cap containing three hydrogen-bonded adenines that stacks on the 5' terminal G-tetrad plane [56]. This structure could contain an additional flanking nucleotide tail at the 5' terminal region, depending on the presence of thymines at the beginning of the sequences. The hybrid-2 structure is quite similar but the loop arrangement is different and possesses a distinct cap with two thymines and one hydrogen-bonded adenine that stacks on the 3' terminal G-tetrad plane of the quadruplex scaffold [56]. This structure, in contrast to the hybrid-1 one, could possess up to two flanking nucleotide tails at the terminal regions, depending on sequence. The stability of the two structures is due to a combination of the presence of the basic structural features (quadruplex scaffold, loops and cap region) of each structure, plus the effect coming from the presence of possible flanking nucleotide tails. Since the adenine cap has greater stability compared to the analogous one found in the hybrid-2 structure, the hybrid-1 fold is more stable than the hybrid-2 one. The results of several biophysical studies do not exclude the presence of additional structures in solution [137,118]. Indeed, another possible folding of these oligonucleotides seems to be represented by the basket-type conformation that has been recently found as the major structure in  $K^+$  solution for the sequence  $d[(GGGTTA)_3GGGT]$  [57], which may be more stable than either hybrid forms. This conformation, termed Form-3, is an antiparallel structure, possessing only two G-tetrad planes, with novel interconnecting loops. The existence of the Form-3 conformation shows that the polymorphism of the human telomeric DNA sequences in dilute solution is yet more complex than was previously assumed, with the implication that it may be yet more challenging to understand the basic principles governing drug-quadruplex interaction processes. On the other hand, under molecular crowding condition the major conformation seems to be represented by the parallel structure and to date, this topology is the only one clearly identified under this type of condition, which may be more representative of a cellular environment. It seems that the polymorphism of quadruplex formed by telomeric DNA sequences is diminished under molecular

crowding condition and the parallel structure is the favoured one, probably due to its particular structural features that enable optimal adaptation to that environment, especially the large and accessible G-tetrad surface areas at both 5' and 3' ends.

The combination of spectroscopic and calorimetric results obtained both in dilute solution and under molecular crowding condition, enables to shed light on the recognition processes between the quadruplex structures formed by the human sequences and TMPyP4. The CD and ITC experiments indicate that, in agreement with the previously reported study of Wei *et al.* [134], the final stoichiometry of the complexes is four. The spectroscopic results in dilute solution indicate that the porphyrin molecules, interacting with the hybrid structures, promote a major conformational rearrangement converting the hybrid structures into a basket one. Since the CD spectra of the complexes between the sequences investigated here and TMPyP4, in solution, share similarities with the CD spectrum of the recently studied Form-3 structure, it is plausible to speculate that Form-3 could be involved in the TMPyP4 interaction, though yet other forms cannot be excluded. Probably, the Form-3 conformation possesses the structural features to efficiently bind up to four molecules of TMPyP4. The Form-3 structure differs from the basket conformation found in Na<sup>+</sup> solution not only in its smaller number of G-tetrad planes but also in the presence of the distinct loop sequences, GTTA and GTTAG rather than the classical telomeric DNA loop sequence TTA. These longer loop sequences; together with the reduced number of G-tetrad planes could make the global structure more relaxed. The greater flexibility of Form-3 with respect to the hybrid structures could be the key feature that explains its affinity for TMPyP4 molecules. On the other hand, the calorimetric analysis in dilute solution reveals that the four molecules of TMPyP4 do not simultaneously bind the hybrid structures, but the interaction is sequential and composed of two distinct binding events. In the first event one molecule of TMPyP4 interacts with the quadruplex structure and after that, in the second event, three other molecules interact with the structure. Both events cooperate to promote a conformational change from the hybrid structures to the basket-type structure, according to the spectroscopic derived binding curves. The calorimetric results show that the global Gibbs energy change for the binding of TMPyP4 to the hybrid-1 structure is lower (about -68 kJ mol<sup>-1</sup>) than that for the hybrid-2 structure (about -86 kJ mol<sup>-1</sup>) and this

difference could derive from the higher stability of the hybrid-1 structure due to its adenine cap [56]. It is plausible that the interaction of TMPyP4 molecules with the 5' terminal G-tetrad plane of the hybrid-1 structure destabilizes the adenine cap. This could be a releasing event allowing the hybrid structures to evolve toward the Form-3 conformation, producing the final complex. The disruption of the adenine cap could explain the cooperative profiles of the binding curves obtained, for the hybrid-1 structures, with both circular dichroism and isothermal titration calorimetry. The hybrid-2 conformation does not possess a structured cap on the 5' terminal G-tetrad plane, moreover, that plane is partially solvent exposed because of the absence of the flanking nucleotide tail, making this structural region an attractive target for binding molecules. Probably, the first binding event revealed by ITC could be related to a TMPyP4 molecule that binds to the hybrid structures via stacking interactions on the 5' terminal G-tetrad plane, which would be consistent with the lower binding constant values obtained for the hybrid-1 structures with respect to those obtained for the hybrid-2 structures. In particular, the lower values found for the hybrid-1 structures take into account the disruption of the adenine cap. The second binding event found by ITC shows the same thermodynamic properties for both hybrid-1 and hybrid-2 structures, suggesting that in both cases the three TMPyP4 molecules interact with similar structural elements. Probably the three porphyrin molecules involved in the second binding event interact with the quadruplex structure in the final step of the conformational change, when the basket-type structure is almost reached. It is plausible that the three interconnecting loops of the Form-3 structure could be individually the targets for binding the final three TMPyP4 molecules. Comparing the previous results with the ones obtained for the duplex in dilute solution, we found that the TMPyP4 does not have significantly higher affinity for the hybrid-1 structures with respect to the duplex. On the other hand, just considering the first binding event, the affinity of TMPyP4 for the hybrid-2 structures is about 100-fold higher than the one measured for duplex. This intriguing result has not been observed to date, probably because the majority of the previous studies were focused on the hybrid-1 forming sequences. Binding of TMPyP4 to human telomeric quadruplexes is qualitatively different in conditions where molecular crowding takes place. Our spectroscopic analysis reveals that the parallel structure is the most populated conformation in this state, and that

the interaction with the TMPyP4 does not promote any conformational change. This result suggests that the parallel structure possesses the features to bind and allocate up to four porphyrin molecules, in general accord with features found in the crystal structure of a bimolecular telomeric quadruplex-TMPyP4 complex [135]. The calorimetric analysis reported here does not reveal significant differences between the four sequences; in particular the global Gibbs energy change for the binding of each sequence to TMPyP4 are similar and average around  $-86 \text{ kJ mol}^{-1}$ . The thermodynamic analysis suggests that the molecular mechanism of the interaction with TMPyP4 is the same for all the sequences, suggesting that the conformations of the complexes are structurally homogeneous. The thermodynamic signature of the first binding event is typical of an end stacking-mode interaction, in which the enthalpic contribution is higher than the entropic one, and one of the two G-tetrad terminal planes could be the target region involved in this first event, as shown in the NMR structure of a non-telomeric (c-myc) quadruplex with bound TMPyP4 [140]. A detailed examination of this structure shows that, contrary to expectation, the porphyrin core does not actually stack onto a G-tetrad, since its dimensions preclude this from occurring [94]. Instead, the edges of the out-of-plane pyridyl groups are in contact with the G-tetrad plane: an analogous arrangement is likely with telomeric quadruplexes. The second binding event is entropically driven, suggesting that it causes a relaxation of those regions of the quadruplex structure involved in the interaction. Moreover this event occurs with three different molecules of TMPyP4 that simultaneously bind to the quadruplex structure. These three similar regions of the quadruplex structure could be represented by the three interconnecting loops, with interactions as found in the crystal structure of the TMPyP4-quadruplex complex [135].

A study of the interaction between the quadruplex structure from the sequence d[AGGG(TTAGGG)<sub>3</sub>] and TMPyP4, reported a global stoichiometry of 1.5 molecules of porphyrin per quadruplex structure under molecular crowding condition [138]. The formation of a sandwich-like complex was proposed on the basis of their results, in which one TMPyP4 molecule is placed between the external planes of two distinct quadruplexes and other two TMPyP4 molecules are bound to the grooves of the two quadruplexes. We also found a first stronger binding event followed by a weaker one, in agreement with this study but our results do not suggest the formation of sandwich-like complexes.

This could possibly be due to the presence of flanking bases situated at the 3' or 5' terminal region. These are flexible structural elements that span the external planes and could prevent dimer aggregation between two quadruplex structures via end stacking interactions. From the comparison of these results, can assessed that under molecular crowding condition, the porphyrin is able to bind the quadruplex structure with higher affinity compared to the duplex DNA structure. This finding is completely unexpected and suggests that the potential biological applications of TMPyP4 may be more selective than has been previously assumed.

The overall picture, suggests that TMPyP4 molecules could bind to the truncated human quadruplex sequences, both in dilute solution and under molecular crowding, via two distinct stacking-mode interactions: with the external G-tetrad planes and with the bases of the interconnecting loops between the G-tetrad planes. The two TMPyP4-quadruplex molecular structures available to date each show just one of the two possible binding-modes [135,140]. The nature of the complexes reported in this study could involve both mode of interaction, and they could be helpful in future chemical optimization to enhance the affinity of porphyrin-type molecules for telomeric quadruplex structures.

### **3.2. Interaction between TMPyP4 and eight-repeats human telomeric sequence**

In this paragraph have been discussed the binding proprieties of the multimeric structure formed by the sequence (TTAGGG)<sub>8</sub>TT with cationic TMPyP4 (Figure 3.1) and compared with the quadruplex forming sequence (AG<sub>3</sub>TT)<sub>4</sub> (wtTel24b, Table 3.1), through a combination of spectroscopic (CD, UV and fluorescence) and calorimetric techniques (ITC). The (TTAGGG)<sub>8</sub>TT telomeric sequence consists of two consecutive mixed-type quadruplexes, one having the Hybrid-1 conformation and the other having the Hybrid-2 conformation, this model has been validated experimentally [64]. TMPyP4 is a good starting drug to explore the difference in the binding properties of multimers in comparison with individual quadruplexes as its interactions with individual

quadruplex structures have been extensively studied. The changes of the TMPyP4 spectroscopic properties (fluorescence, induced CD, UV spectrum) upon G-quadruplex binding are known to be large making it particularly suitable as experimental probe to explore the binding properties of a new G-quadruplex structure. Further, TMPyP4 has a high solubility allowing to reach the high [drug]/[DNA] ratio required in the ITC titrations to explore the binding of longer telomeric sequences.

The (AGGGTT)<sub>4</sub>, (TTAGGG)<sub>8</sub>TT oligonucleotide sequences were studied in a buffer solution containing 20 mM phosphate, 0.1 mM EDTA, 100 mM KCl. The concentration of oligonucleotides was determined by UV adsorption measurements at 90 °C using molar extinction coefficient values  $\epsilon(260 \text{ nm})$  of 245100, 505600 M<sup>-1</sup> cm<sup>-1</sup> for (AGGGTT)<sub>4</sub>, (TTAGGG)<sub>8</sub>TT, respectively [139]. TMPyP4 solutions have been prepared in the same buffer used for the oligonucleotides, and the concentrations have been estimated by UV spectroscopy using the reported molar extinction coefficient value of  $\epsilon(424 \text{ nm}) = 226000 \text{ M}^{-1} \text{ cm}^{-1}$  [141].

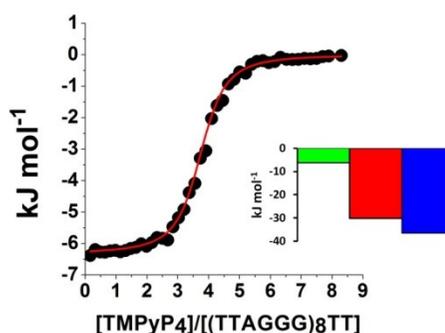
### 3.2.1. Results

#### 3.2.1.a. ITC experiments

The binding of TMPyP4 to the (TTAGGG)<sub>8</sub>TT sequence was characterized by means of isothermal titration experiments (ITC). In each titration, volumes of 5–10  $\mu\text{L}$  of a solution containing TMPyP4 at a concentration of 160–280  $\mu\text{M}$ , were injected into a DNA solution (10–20  $\mu\text{M}$ ) in the same buffer, using a computer-controlled 250  $\mu\text{L}$  microsyringe.

The ITC binding curve obtained from the integration of the heat data has a sigmoidal behaviour and was fitted with an independent and equivalent-sites model (Figure 3.9), the best-fit parameters are:  $n=4.0$ ,  $K_b = 2.4 \times 10^6 \pm 0.1 \text{ M}^{-1}$ ,  $\Delta_b H^0 = -6.2 \pm 0.3 \text{ kJmol}^{-1}$ ,  $-\text{T}\Delta_b S^0 = -30.2 \pm 0.3 \text{ kJmol}^{-1}$  and  $\Delta_b G_{298\text{K}}^0 = -36.4 \pm 0.2 \text{ kJmol}^{-1}$  ( $n$  is referred to the number of ligand molecules interacting in each binding event). In insets of Figure 3.9 the free-energy changes for the binding processes were parsed into the respective enthalpic and entropic contributions. Both the enthalpic

and the entropic terms are favourable to the binding, the largest contribution was the entropic one. The negative enthalpic contribution indicates the formation of new interactions upon binding whereas the positive entropic contribution could be due to a release of water molecules or/and a relaxation of the structure [142,143].



**Figure 3.9** Normalized heat of interaction between TMPyP4 with the structure formed by the  $(TTAGGG)_8TT$  sequence. The solid circles are the experimental data obtained by integrating the raw data and subtracting the heat of ligand dilution into the buffer. The red line represents the best fit obtained with the independent-sites model. The inset shows the thermodynamic signature of the corresponding interaction. The enthalpic contribution ( $\Delta H^0$ ) is shown in green, the entropic one ( $-T\Delta S^0$ ) in red and the Gibbs energy change ( $\Delta G^0_{298K}$ ) in blue.

As previously described, four TMPyP4 molecules bind the  $(AG_3TT)_4$  quadruplex (wtTel24b in dilute solution, Table 3.2) but in two separate events: in the first event, one TMPyP4 molecule binds the quadruplex with a  $K_b$  of  $5 \times 10^8 M^{-1}$ , in the second event, other three molecules bind the quadruplex with a  $K_b$  of  $2 \times 10^6 M^{-1}$ . Instead the thermodynamic parameters ( $K_b$ ,  $\Delta_b H^0$  and  $-T\Delta_b S^0$ ) for the binding of TMPyP4 to the  $(TTAGGG)_8TT$  structure are similar to the ones reported for the second binding event in the  $(AG_3TT)_4$  quadruplex (Table 3.2) but greatly differ (especially the  $K_b$  and the  $-T\Delta S$  term) from the parameters characterizing the first binding event. The total stoichiometry (sum of the two binding events) of 4:1 reported for the  $(AG_3TT)_4$  quadruplex is the same of the stoichiometry observed for the two-quadruplexes forming sequence

(TTAGGG)<sub>8</sub>TT. This result reveals a decrease of the number of binding sites for each quadruplex unit in the (TTAGGG)<sub>8</sub>TT structure (two binding sites per quadruplex unit) in comparison with the (AG<sub>3</sub>TT)<sub>4</sub> quadruplex (four binding sites per quadruplex).

### 3.2.1.b. CD experiments

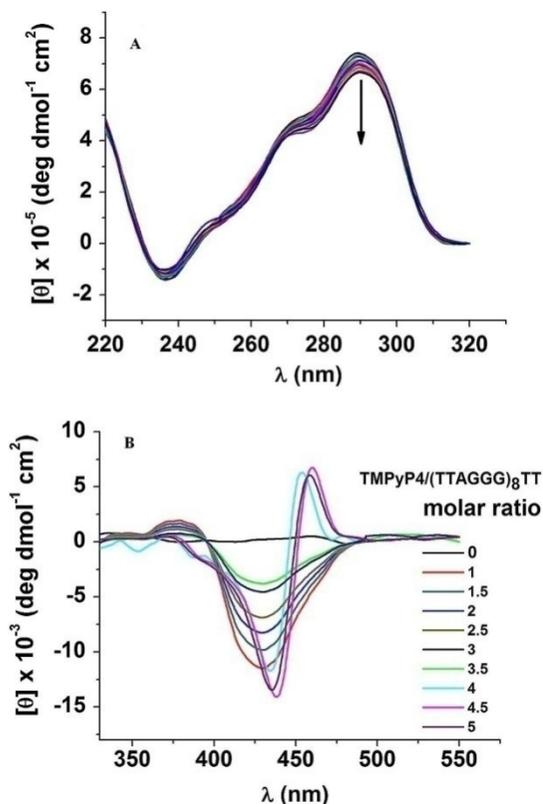
To get more information on the conformation of the multimer–drug complexes, CD measurements were performed. Figure 3.10A shows the changes in the CD spectrum of (TTAGGG)<sub>8</sub>TT on increasing the amount of TMPyP4.

Cell with 0.1 and 1.0 cm path length and oligonucleotide concentrations in the range 2–40  $\mu$ M were used to record CD spectra between 220 and 320 nm.

The CD spectrum of the (TTAGGG)<sub>8</sub>TT in absence of ligand shows the characteristic shape of a mixed-type (Hybrid-1 and Hybrid-2) G-quadruplex structures and its intensity is consistent with the formation of two-quadruplex units [75]. There are little changes of the (TTAGGG)<sub>8</sub>TT spectrum on increasing the TMPyP4 concentration indicating that the binding of the ligand does not cause drastic conformational changes in the overall multimer structure (Figure 3.10A). Figure 3.10B shows the induced CD of the TMPyP4 on increasing (TTAGGG)<sub>8</sub>TT concentration.

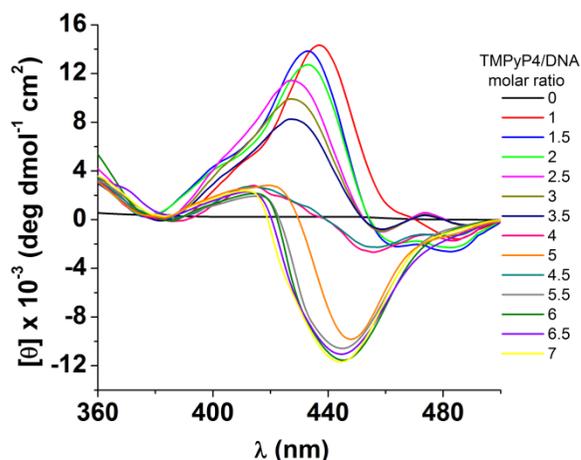
Then Induced CD titrations were carried out by stepwise addition of DNA solutions to a cell containing drug solution. The concentration of TMPyP4 was 40  $\mu$ M. The DNA concentration was varied in the range 0–60  $\mu$ M and the spectra were acquired in the wavelength range 300–600.

ICD of TMPyP4 has a complex behaviour on changing [TMPyP4]/[DNA] ratio, particularly, it has a negative value in the 440 nm region for a [TMPyP4]/[DNA] binding ratio <3.5 but it shows both positive and negative peaks for a binding ratio >3.5. Negative ICD in the Soret region of TMPyP4 is consistent with an end-stacking binding mode whereas positive ICD bands are diagnostic of an external binding mode [144,145,146]. This result suggests that two different binding modes are present that contribute to the ICD signal: an end-stacking mode at low [drug]/[DNA] ratio with a negative ICD and an external binding at higher binding ratio that gives a positive contribution to the total ICD signal.



**Figure 3.10** (A) CD spectra of  $(TTAGGG)_8TT$  (2  $\mu M$ ) on increasing TMPyP4 concentration in the range 0–16  $\mu M$  (B) Induced CD spectra of TMPyP4 (40  $\mu M$ ) on changing the  $[TMPyP4]/[(TTAGGG)_8TT]$  molar ratio.

The coexistence of these two types of binding modes in the quadruplex-TMPyP4 complexes has been previously suggested by many authors [119,145], ICD of TMPyP4 in presence of  $(AG_3TT)_4$  is positive at low  $[drug]/[DNA]$  molar ratio and decreases in magnitude on increasing the DNA concentration becoming negative for  $[drug]/[DNA]>4$  (Figure 3.11). This behaviour is different from the one observed in presence of  $(TTAGGG)_8TT$  suggesting the presence of significant difference in the binding of TMPyP4 to the two DNA sequences.



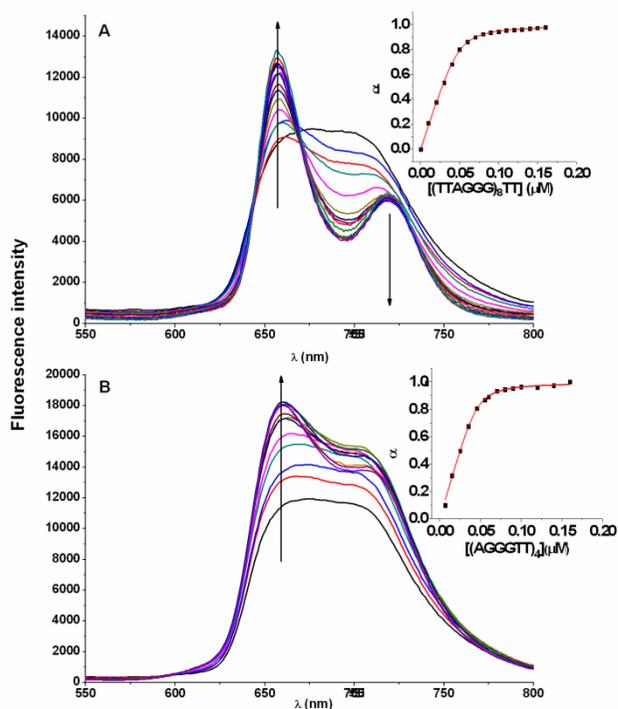
**Figure 3.11** ICD spectra are for TMPyP4 (40  $\mu$ M) on changing the  $[TMPyP4]/[(AGGGTT)_4]$  molar ratio.

### 3.2.1.c. Fluorescence titration experiments

The fluorescence properties of TMPyP4 were studied to get more information on their binding to the DNA telomeric sequences. TMPyP4 was excited at 435 nm and emission spectrum was recorded in the range 550–800 nm. The association processes between the quadruplex structures and drugs were studied by following the change in fluorescence on changing DNA concentration. The titrations were carried out by stepwise addition of DNA solutions to a cell containing drug solution. The concentration of TMPyP4 was 0.2  $\mu$ M.

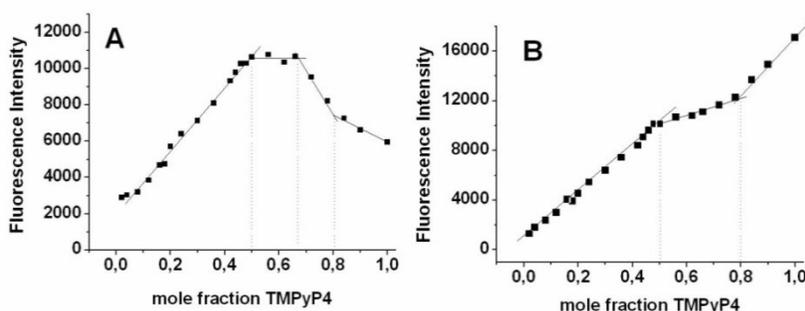
Figure 3.12 shows the changes in the fluorescence spectrum of TMPyP4 on increasing the amount of  $(TTAGGG)_8TT$  and  $(AG_3TT)_4$  concentration. Adding  $(TTAGGG)_8TT$  to the TMPyP4 solution, the structureless emission spectrum of the free drug changes into a two-band spectrum and its maxima and the minimum become more marked on increasing DNA concentration (Figure 3.12A). In particular, a significant increase of the emission intensity at 660 nm, a large decrease at 690 nm with an associated much smaller decrease of the emission intensity at 720 nm, was observed. A similar behaviour has been previously reported for the TMPyP4 binding to other G-quadruplex structures [138,128,147]. The inset of Figure 3.12A depicts the binding curve obtained by

plotting the fraction of bound ligand ( $\alpha$ ) against the [(TTAGGG)<sub>8</sub>TT] concentration (equation 2, chapter 2, paragraph 2.3). The binding curves were fitted with the equation 4 (chapter 2 paragraph 2.3) [148,119]. The binding constant obtained is  $5.0 \times 10^6 \text{ M}^{-1}$  and a stoichiometry of 4:1 in agreement with the results derived by ITC method.



**Figure 3.12** (A) Fluorescence spectra of TMPyP4 in absence and presence of successive additions of (TTAGGG)<sub>8</sub>TT at 25 °C. (B) Fluorescence titration spectra of in absence and presence of successive additions of (AG<sub>3</sub>TT)<sub>4</sub> at 25 °C. Arrows indicates the increasing DNA concentrations. The concentrations of TMPyP4 were 8 μM 0.2 μM. The insets are the plot of the fraction of TMPyP4 or azatrux bound ( $\alpha$ ) vs the [DNA] concentration, the black square represents the experimental data and the red line is the theoretical curve obtained on the basis of independent and equivalent-sites model. The excitation wavelength is 435 nm.

Examination of the Job plot (Figure 3.13A) shows several slope changes between approximately linear regions. The intersection points are at TMPyP4 mole fractions of 0.5, 0.67 and 0.8 corresponding to stoichiometric ratios of 1:1, 2:1 and 4:1. Although, the final TMPyP4:(TTAGGG)<sub>8</sub>TT stoichiometry is consistent with the overall stoichiometry determined by ITC and previous fitting of the fluorescence binding curve, this analysis gives more information on the binding process showing the presence of three separate binding events: the first two involve one TMPyP4 molecule and the last binding involves two TMPyP4 molecules. This analysis suggests the presence of different binding modes for TMPyP4 in agreement with the ICD results.



**Figure 3.13** Job plots analysis of fluorescence binding data for the binding of TMPyP4 to (TTAGGG)<sub>8</sub>TT (A) and to (AGGGTT)<sub>4</sub> (B). In all the experiments, the total molar concentration ( $[ligand] + [DNA]$ ) was kept constant at 2  $\mu$ M. The excitation wavelength is 435 nm.

Figure 3.12B shows the change in fluorescence spectrum of TMPyP4 on increasing (AG<sub>3</sub>TT)<sub>4</sub> concentration. An increase of the emission intensity at all the wavelength associated with a more marked increase of the emission intensity at 660 nm resulting in final spectrum with a peak centered at 660 nm and a shoulder at 720 nm, was observed. The fitting of the binding curve (inset of Figure 3.12B) results in a TMPyP4:DNA stoichiometry of 4:1 and a binding constant of  $1.0 \times 10^7 \text{ M}^{-1}$ . Also in this case, the Job plot analysis allowed us to separate the total stoichiometry in two separate binding events at TMPyP4 mole fractions of 0.5 and 0.8 (Figure 3.13B), corresponding to stoichiometric ratios of 1:1 and

4:1. In the first binding event one TMPyP4 molecule binds to  $(AG_3TT)_4$  quadruplex, in the second binding event other three TMPyP4 molecules bind the  $(AG_3TT)_4$  quadruplex in agreement with the reported ITC results.

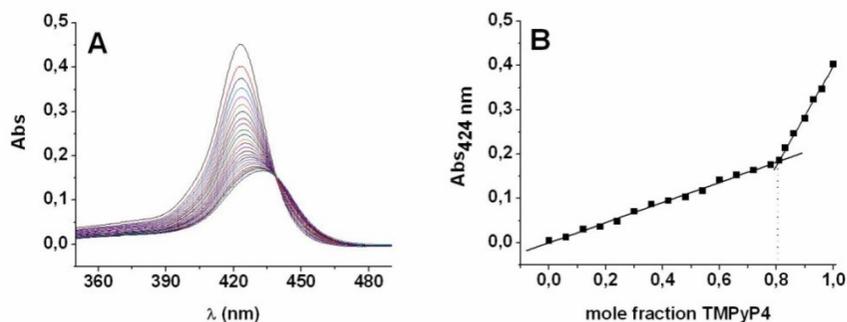
### 3.2.1.d. UV-vis absorption titration experiments

Binding of TMPyP4 to the multimeric  $(TTAGGG)_8TT$  quadruplex structure was investigated also by UV-Vis titration experiments.

The spectra were recorded in the 350–500 nm range at 25°C.

Figure 3.14A shows the absorbance of TMPyP4 on increasing  $(TTAGGG)_8TT$  concentration. An isosbestic point at 435 nm was observed. The presence of an isosbestic points, indicative of equilibrium between the bound and free ligand, has been already reported for UV titration experiments of TMPyP4 to G-quadruplex [119,138].

UV-Vis spectra of TMPyP4 on increasing DNA concentration show a red shift in the absorbance maximum from 422 nm to 434 nm and a hypochromicity of ~35%. Comparable values of red shift and hypochromicities have been previously observed for other quadruplexes and suggested to be diagnostic of an end-stacking binding mode of TMPyP4 [119,138]. Job plot of TMPyP4 (Figure 3.14B) shows a break at 0.8 (mole fraction TMPyP4) corresponding to a 4:1 binding stoichiometry.



**Figure 3.14** (A) UV absorption spectra of TMPyP4 in the absence and presence of  $(TTAGGG)_8TT$  at 25°C (B) Job plot for the binding of TMPyP4 to  $(TTAGGG)_8TT$ . The total  $[TMPyP4+DNA]$  concentration was 2  $\mu M$ .

### 3.2.2. Discussion and Conclusions

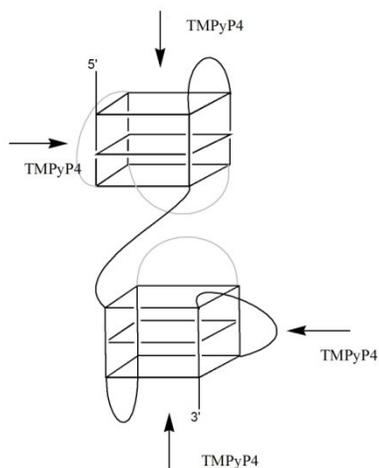
TMPyP4 has been chosen to explore the binding properties of higher-order human telomeric multimer in comparison with the single quadruplex structure.  $(TTAGGG)_8TT$  was used as multimer forming sequence, this sequence have been previously characterized and it forms a two-quadruplex structure. In particular, a high-resolution structural model have been previously suggested for this sequence consisting of two mixed-type quadruplex structures with a characteristic quadruplex–quadruplex interface stabilized by stacking interactions [64]. As single quadruplex, was used the  $(AG_3TT)_4$  quadruplex which has been previously studied from a structural point of view and it forms the mixed-type Hybrid-2 intramolecular quadruplex [56,118]. The ITC and fluorescence experiments show that four TMPyP4 molecules are able to bind the  $(TTAGGG)_8TT$  dimeric structure, this corresponds to two TMPyP4 molecules bound for each quadruplex unit in the dimer. The comparison of this value, with the 4:1 stoichiometry, reported for the binding to the single  $(AG_3TT)_4$  quadruplex, suggests that some binding sites available in the isolated quadruplex structure are not available for the binding when the quadruplex is followed by an adjacent quadruplex unit in the  $(TTAGGG)_8TT$  structure. This observation demonstrates that, at least with respect to the binding properties, the two quadruplex units in the  $(TTAGGG)_8TT$  structure do not behave as completely independent units. Interestingly, two binding events are reported for the  $(AG_3TT)_4$  quadruplex involving one and three TMPyP4 molecules whereas only one binding event was observed for  $(TTAGGG)_8TT$ . The binding constants and thermodynamic parameters observed for the TMPyP4 binding to the  $(TTAGGG)_8TT$  are similar to the ones observed for the lower affinity binding in  $(AG_3TT)_4$  and no higher affinity binding site was observed. This observation, together with the observed decrease of binding sites per quadruplex unit, suggests that the presence of an adjacent quadruplex in the  $(TTAGGG)_8TT$  structure negatively affects the binding of the TMPyP4 to each quadruplex subunit. This effect could result from the steric hindrance due to the presence of the adjacent quadruplex. In this hypothesis, the quadruplex–quadruplex interface is not involved directly in the binding process and its effect on the binding is to reduce the number of available binding sites in the quadruplex subunits. This hypothesis is consistent with the absence of CD changes upon TMPyP4 binding suggesting that the binding

does not involve major conformational changes of the (TTAGGG)<sub>8</sub>TT structure as it could be expected for a binding involving the quadruplex–quadruplex interface. In a previous study have been suggested a (TTAGGG)<sub>8</sub>TT structural model consisting of two adjacent mixed-type quadruplex structures [64]. It was shown that the formation of the quadruplex–quadruplex interface in the (TTAGGG)<sub>8</sub>TT structure involves one quadruplex end and a lateral loop for each quadruplex subunit (Figure 3.15). These structural elements are available binding sites for TMPyP4 in the monomeric quadruplex but could be not available to the binding (they represent the lost binding sites) in (TTAGGG)<sub>8</sub>TT due to their involvement in the quadruplex–quadruplex interface. On the other hand, inspection of the (TTAGGG)<sub>8</sub>TT structural model (Figure 3.15) reveals several possible binding sites not involved in the quadruplex–quadruplex interface and that can remain unaffected by the presence of the adjacent quadruplex: the two propeller loops (one for each quadruplex unit), the quadruplex grooves and the G-tetrads at the 5' and 3' ends of the (TTAGGG)<sub>8</sub>TT structure. The ability of TMPyP4 to bind externally through stacking interaction with the TTA loop bases (external binding mode) has been proved by crystallographic methods whereas the terminal G-tetrads have been proposed as possible binding sites (end-stacking binding mode) on the basis of structural, spectroscopic, calorimetric and computational studies [119,134,149,135]. The binding of TMPyP4 in these sites can happen without any steric hindrance between bound ligands as the two propeller loops, and the G-tetrads at the 5' and 3' ends, are situated on opposite sides in the (TTAGGG)<sub>8</sub>TT structural model (Figure 3.15).

This picture is consistent with the CD results showing that the binding occurs without any structural change in the (TTAGGG)<sub>8</sub>TT structure. Further, the complex behaviour of the ICD of bound TMPyP4 on varying the [TMPyP4]/[DNA] ratios is consistent with the presence of an end-stacking binding mode (negative ICD) and an external binding mode (positive ICD) [150,144,145,146].

The external binding mode is probably characterized by lower affinity as its contribution to the ICD appears at higher [TMPyP4]/[DNA] ratio. The observed ICD at higher [TMPyP4]/[DNA] ratio could be also consistent with a TMPyP4 binding to the quadruplex grooves. The presence of an end-stacking binding of TMPyP4, wherein the aromatic planar rings of TMPyP4 stack over the terminal G-tetrads of the (TTAGGG)<sub>8</sub>TT structure,

is also suggested by the high hypochromicities and large bathochromic shift observed in the UV titration experiments [134].



**Figure 3.15** Scheme showing  $(TTAGGG)_8TT$  structure along with the proposed *TMPyP4* binding sites. The schematic dimer structure was built following the  $(TTAGGG)_8TT$  structural model previously reported [64].

In summary, there is a decrease of the binding site available for each quadruplex unit and the remaining ones are unaffected by the presence of the adjacent quadruplex. Although has not been observed an improved selectivity when targeting with *TMPyP4* to the  $(TTAGGG)_8TT$  multimeric quadruplex instead of the monomeric one. The results show that i) the quadruplex units in the multimeric structure do not behave as independent with respect to their binding properties, ii) the presence of adjacent quadruplexes can result in a diverse binding ability not predictable from single quadruplex binding studies. It should be noted that *TMPyP4* is known as monomeric quadruplex ligand, not appositely designed to bind a multimeric structure. A drug design process based on a multimeric structure as target structure may provide a new route to enhance both affinity and selectivity for DNA telomeric quadruplexes.

## **CHAPTER 4 A hydrophilic three side-chained triazatruxene as a new G-quadruplex ligand**

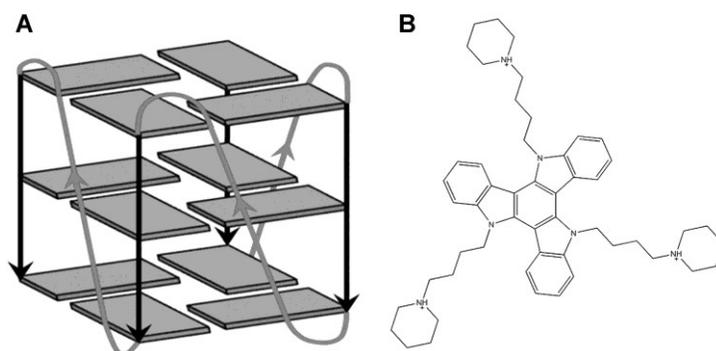
### **4.1. Interaction between triazatruxene derivative (azatrux) and human telomeric G-quadruplex under molecular crowding conditions.**

A combination of experimental and computational techniques have been employed here to fully characterize the binding of a triazatruxene based G-quadruplex ligand (azatrux) to the parallel G-quadruplex formed by the d[AGGG(TTAGGG)<sub>3</sub>TT] human telomeric DNA sequence under molecular crowding conditions. 40% PEG solution has been used to simulate crowding conditions and to induce the parallel G-quadruplex conformation to form [63]. Azatrux is a triazatruxene derivative with three side-chains (Figure 4.1B), a highly hydrosoluble compound, that has been previously designed to bind to the parallel G-quadruplex topology [151].

The mode of azatrux binding to the parallel human telomeric quadruplex was investigated by CD, UV, fluorescence spectroscopy and the energetics of the binding was characterized by means of isothermal titration calorimetry (ITC). The dynamic and structural features of the azatrux-DNA complex were explored by means of molecular modeling methods. Further, the binding of azatrux to the parallel human telomeric G-quadruplex was compared with binding to the “standard” parallel quadruplex formed by the d(TGGGGT) sequence. Finally, the selectivity of azatrux for the human telomeric G-quadruplex relative to another biologically relevant G-quadruplex (c-Kit87up) and to duplex DNA was also evaluated under molecular crowding conditions.

Quadruplex and duplex samples were prepared by dissolving the lyophilised compound in a buffer solution containing 20 mM phosphate with 70 mM KCl, 0.1 mM EDTA, 40% (w/v) of PEG 200, at pH 7.0. The d(TGGGGT) quadruplex samples were also prepared in the same buffer described above but in the absence of PEG 200. The resulting solutions were annealed by heating at 95 °C for 5 min. The solutions were then slowly cooled to room

temperature and equilibrated for 1 day at 4 °C. The concentration of oligonucleotides was determined by UV adsorption measurements at 90 °C using molar extinction coefficient values  $\epsilon_{(260\text{nm})}$  of 245100, 57800, 110700 and 226700  $\text{M}^{-1} \text{cm}^{-1}$  for d[AGGG(TTAGGG)<sub>3</sub>TT], d(TGGGGT), d(CGCGAATTCGCG) and d(AGGGAGGGCGCTGGGAGGAGGG), respectively. The molar extinction coefficients were calculated by the nearest neighbour model [139]. 5,10,15-tris [4(1-piperidino) butyl] diindolo [3,2-a:3',2'-c] carbazole (azatrux) was synthesized as reported in literature [151] and solutions were prepared in the same buffer as used for the oligonucleotides. The concentration has been estimated by UV spectroscopy using the calculated extinction coefficient value  $\epsilon_{(313\text{nm})}$  of 48,000  $\text{M}^{-1} \text{cm}^{-1}$ .



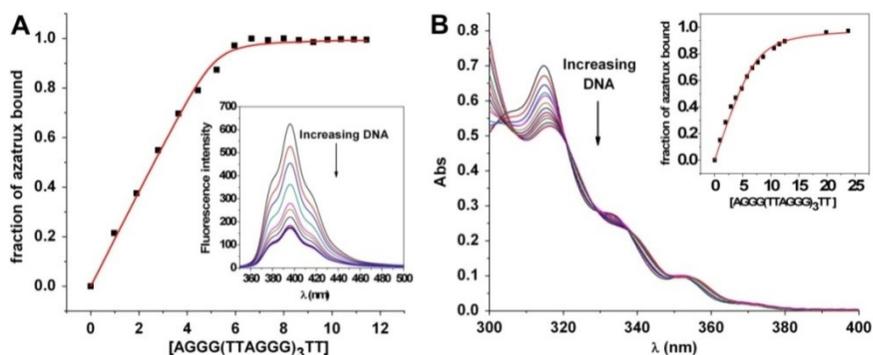
**Figure 4.1** (A) Schematic representation of the unimolecular parallel G-quadruplex structure assumed by the DNA human telomeric sequence in crowding condition. (B) Azatrux chemical structure.

#### 4.1.1. Results

##### 4.1.1.a. Binding properties of azatrux to human telomeric quadruplex under molecular crowding conditions

Fluorescence and UV spectroscopic methods were used to investigate the binding of azatrux to the parallel G-quadruplex

structure formed by the d[AGGG(TTAGGG)<sub>3</sub>TT] human telomeric DNA sequence under molecular crowding conditions. Figure 4.2A shows the fluorescence spectra of azatrux with increasing [DNA]/[azatrux] molar ratio. Free azatrux has a fluorescence maximum at 394 nm when excited at 321 nm. A decrease in fluorescence on increasing DNA concentration, but no change in spectral shape, was observed. A binding curve was obtained by plotting the fraction of ligand bound as a function of the DNA concentration (Figure 4.2A). The curve was fitted using a multiple and independent binding site model, by means of nonlinear regression algorithm, giving a 1:1 stoichiometry and a binding constant of  $8 \times 10^6 \text{ M}^{-1}$ .

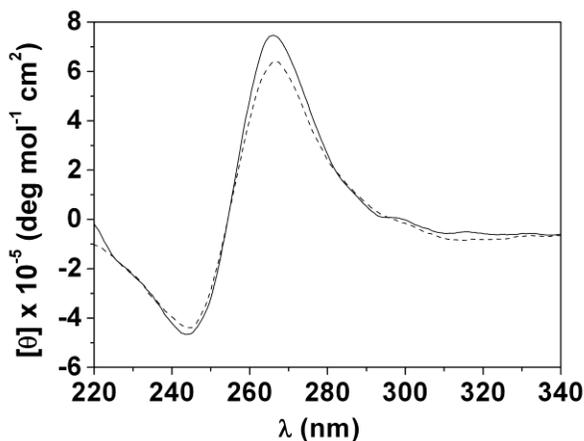


**Figure 4.2** (A) Fluorescence emission spectra of 8 μM azatrux solution in the absence and presence of successive additions of telomeric quadruplex (inset), and plot of the fraction of azatrux bound vs. the DNA concentration. (B) UV-Vis absorption titration spectra of 7 μM azatrux solution in the absence and presence of successive additions of telomeric quadruplex. The DNA concentration was varied between 0 and 20 μM, and plot of the fraction of azatrux bound vs. the DNA concentration (inset). The black squares represent the experimental data, the solid line is the best fit obtained with an independent sites model. All the titration experiments were performed at 25 °C in buffer solution containing 20 mM phosphate with 70 mM KCl, 0.1 mM EDTA and 40% (w/v) of PEG 200 at pH 7.0.

UV-Vis spectra of azatrux with increasing DNA concentration show a red shift in the absorbance maximum from 313 nm to 317 nm and a hypochromicity of ~30% (Figure 4.2B). The observation of isosbestic points at 321 nm, 329 and 337 nm clearly indicates the

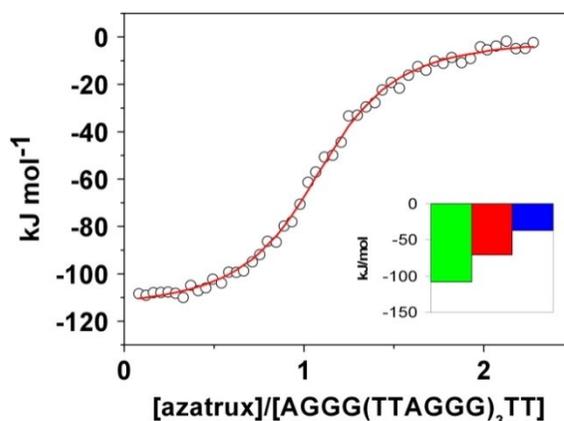
existence of equilibrium in the binding with the presence of both free and bound ligand in solution. The strong hypochromicity observed suggests a mode of binding that involves a stacking interaction between the azatrux aromatic core and the G-quartets of the quadruplex. The binding curve was obtained by plotting the fraction of ligand bound as a function of the DNA concentration (Figure 4.2B, inset). Analysis of the curve with a multiple and independent binding sites model resulted in a 1:1 stoichiometry and a binding constant of  $1 \times 10^6 \text{ M}^{-1}$ .

In the presence of 40% PEG 200, the CD spectrum shows the characteristic shape attributed to the parallel fold with a maximum around 264 nm and a minimum at 241 nm (Figure 4.3). This result is consistent with earlier reports, indicating that PEG 200 is able to induce the parallel fold in a number of quadruplex-forming oligonucleotide [96,63]. It was found that the addition of azatrux to the quadruplex has a small effect on the CD spectrum, as demonstrated by the comparison between the CD spectrum of the free quadruplex and of the 1:1 complex with azatrux (Figure 4.3). This result shows that the binding of azatrux does not significantly alter quadruplex topology.



**Figure 4.3** CD spectra of the human telomeric quadruplex in crowding conditions (solid line), together with its 1:1 complex with azatrux (dashed line).

ITC experiments have been performed in order to further characterize the thermodynamics of azatrux binding to the parallel human G-quadruplex structure [152]. In each titration, volumes of 5–10  $\mu\text{L}$  of a solution containing 160–200  $\mu\text{M}$  azatrux were injected into a solution of quadruplex DNA (10–20  $\mu\text{M}$ ) in the same buffer, using a computer-controlled 250  $\mu\text{L}$  micro-syringe. The resulting binding curve was monophasic, revealing a single binding event at 1:1 stoichiometry (Figure 4.4). The experimental data were fitted to a single-site model. The thermodynamic parameters for the binding are shown in Table 4.1. The thermodynamic signature of the binding (inset in Figure 4.4) reveals that the binding is enthalpically driven, suggesting that the complex is stabilized by the formation of favorable interactions between azatrux and G-quadruplex. The value of the ITC binding constant is in excellent agreement with the values determined by fluorescence and UV–Vis spectroscopy. The negative entropic contribution found for the binding suggests that the resulting complex is more rigid than the free quadruplex.



**Figure 4.4** Binding isotherm for the titration of  $d[\text{AGGG}(\text{TTAGGG})_3\text{TT}]$  human telomeric quadruplex with azatrux obtained by ITC at 25  $^{\circ}\text{C}$  in buffer solution containing 20 mM phosphate with 70 mM KCl, 0.1 mM EDTA and 40% (w/v) of PEG 200 at pH 7.0. The open circles represent the experimental data obtained by integrating the raw data and subtracting the heat of ligand dilution into the buffer. The red line represents the best fit obtained by a nonlinear least-squares procedure based on a single site model. The inset represents the thermodynamic signature of the corresponding interaction. The enthalpic contribution ( $\Delta H^{\circ}$ ) is shown in green, the entropic one ( $T\Delta S^{\circ}$ ) in red, and the Gibbs energy change ( $\Delta G^{\circ}_{(298\text{K})}$ ) in blue.

Quadruplex	<i>n</i>	<i>K<sub>b</sub></i> (M <sup>-1</sup> )	$\Delta_b H^\circ$ (kJ mol <sup>-1</sup> )	$T\Delta_b S^\circ$ (kJ mol <sup>-1</sup> )	$\Delta_b G^\circ$ (298K) (kJ mol <sup>-1</sup> )
<i>Under molecular crowding conditions</i>					
d[AGGG(TTAGGG) <sub>3</sub> TT]	1.1	3.1 × 10 <sup>6</sup>	-108	-71	-37
d[(TGGGGT) <sub>4</sub> ]	1.9	3.6 × 10 <sup>6</sup>	-48	-11	-37
<i>In dilute conditions</i>					
d[(TGGGGT) <sub>4</sub> ]	2.0	2.0 × 10 <sup>5</sup>	-1	28	-29

**Table 4.1** Thermodynamic parameters for the interaction of azatrux with G-quadruplex structures determined by ITC at 25 °C. (The experimental error for each thermodynamic property is <8%)

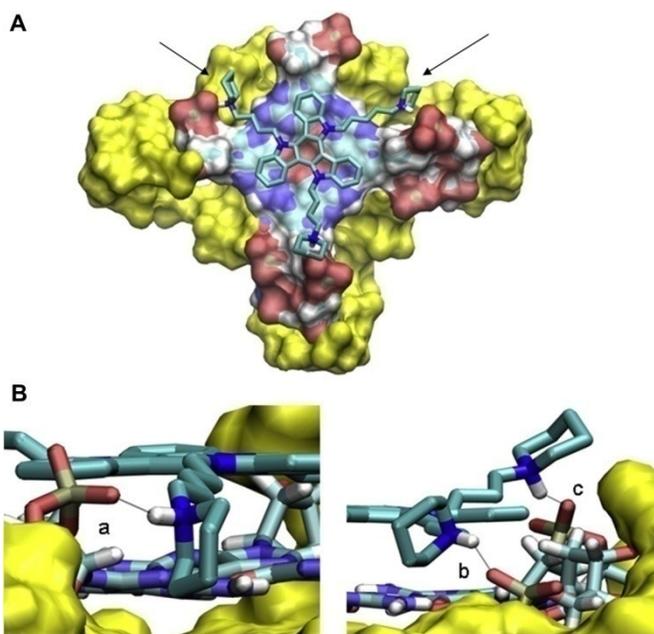
#### 4.1.1.b. Molecular model of the complex between azatrux and the human telomeric quadruplex

Molecular modeling methods were used to examine the molecular basis for the quadruplex binding of azatrux. Molecular modelling studies were performed in collaboration with Prof. Neidle at the School of Pharmacy of London during a research period in London, with the supervision of Dr. Shozeb Haider.

The crystal structure of the 22-mer human telomeric G-quadruplex DNA d[AGGG(TTAGGG)<sub>3</sub>] (PDB code 1KF1) was used as the template for modeling studies. The parallel topology of the structure results in propeller TTA loops, with the G-quartets at both 5' and 3' ends being potential binding sites for azatrux. The terminal A at the 5'-end was removed using the BUILDER module of INSIGHT II package ([www.accelrys.com](http://www.accelrys.com)) in order to generate a 21-mer [67]. K<sup>+</sup> ions are required to stabilize a quadruplex and are located in the crystal structure between the G-quartets, in a vertical alignment along the long axis in the central electronegative channel of the quadruplex. The binding sites in the model were defined by creating two subsets, one each at the 3' and 5' ends of the quadruplex. A molecular model of azatrux was built using the Builder module in the INSIGHT II molecular modeling program suite and partial charges calculated semi-empirically using MOPAC [67].

The ligand was minimized and docked into both binding sites using the CCF (Cornell Force Field) parameter set and then subjected to molecular-mechanics energy minimization (1000 steps) via the Discover program in the INSIGHT II suite. The minimized ligand was then docked in the binding sites using the Affinity docking module in INSIGHT II. The potentials were set for the ligand and the DNA using the CCF force field. The simulated annealing docking was run with 500 iteration steps, allowing ten structures to be generated and one accepted. The most probable conformation for the ligand in each structure was chosen and ranked, based on detailed considerations of optimal binding states. Initially, the azatrux conformation with the highest interaction energy was chosen as a starting point. This chosen conformation was then qualitatively analyzed based on the position and orientation of the docked ligand in relation to the terminal G-quartets. If the chosen conformation of the ligand did not result in a good stacking arrangement with the G-quartet the docked complex was discarded and the conformation with the next highest interaction energy value was analyzed. This procedure was repeated until a stable conformation with optimal G-quartet stacking was identified. The G-quartets were kept frozen in their original conformation throughout the docking procedure. This is a valid approximation because the G-quartet platform has been previously shown in a number of experimental and simulation studies to be exceptionally stable and structurally rigid [62,91,153]. The most favourable complex was selected from the docked structures on the basis of the calculated binding energies after energy minimization and corresponds to the azatrux docked on the 3'-end of the quadruplex structure. This result is probably due to two reasons: (a) firstly the orientation of the TTA loops create grooves that are deeper and accessible to ligand substituents when bound at the 3' end and (b) the 5'- 3' strand polarity generated by the propeller loop topology exposes the phosphodiester backbone on this side and thus making the 3' end more polar and favourable to interact with ligands with positively-charged side chains. The most favourable complex between azatrux and G-quadruplex DNA was chosen on the basis of highest interaction energy and orientation of the ligand around the terminal G-quartet (Figure 4.5A). This model was then used for further analysis. The model shows that the azatrux core makes extensive  $\pi$ - $\pi$ -stacking interactions with the terminal G-quartet. Furthermore, two of the three *N*-butyl-piperidine side-chain substituents of azatrux lie in the grooves formed by the TTA loops.

The positively-charged nitrogen atoms of the piperidine rings can interact with phosphate groups in the phosphodiester backbone. The length and in particular the geometric relationship of the three side-chains prevent the third one from fully accessing the grooves. This side-chain instead sits on top of the terminal G-quartet and also makes favourable interactions with the phosphodiester backbone (Figure 4.5A).



**Figure 4.5** A) Top view of azatrux making  $\pi$ - $\pi$  stacking interactions with the terminal G-quartet, with colour based on electrostatic potential, red (negative) and blue (positive). Two ligand side chains are shown positioned in the grooves created by the TTA loops (indicated by arrows) and the third positioned over the phosphodiester backbone from the terminal G-quartet. (B) The charged nitrogen atoms of the piperidine side chain end-groups (a and b) are in close hydrogen-bonding/electrostatic contact with backbone phosphate groups in two grooves, while the third is positioned on top of the sugar-phosphate backbone on the terminal G-quartet

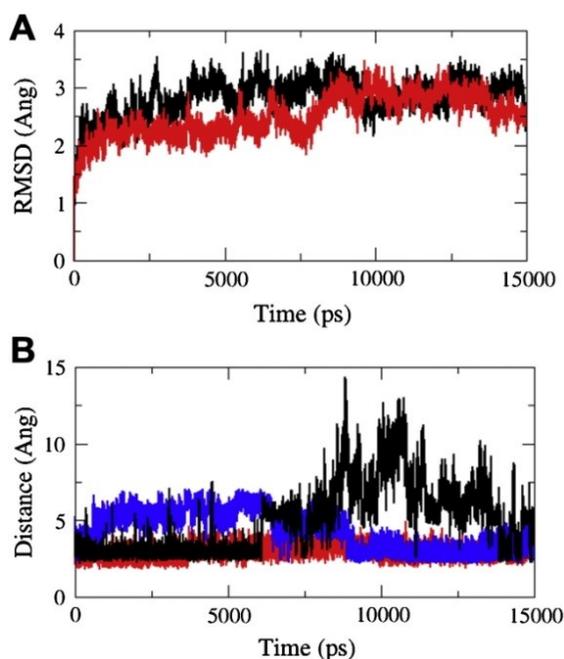
15 ns molecular dynamics simulations of the propeller-type parallel-stranded G-quadruplex DNA and the azatrux-quadruplex

complex have been performed to assess the stability contributed by the binding of azatrux.

The 21-mer native and G-quadruplex complexes docked with azatrux on the 3' end were subjected to molecular dynamics simulation using the Amber parmbsc0 force field ported to the Gromacs package. The models were solvated in a periodic TIP3P water box of dimensions whose boundaries extended at least 10 Å from any solute atom. Additional positively-charged K<sup>+</sup> counterions were included in the system to neutralize the charge on the DNA backbone. These maintained consistency with the crystallization conditions and hence simulated a uniform K<sup>+</sup> ionic environment. Each system was equilibrated with explicit solvent molecules by 1000 steps of minimization and 0.5 ns of molecular dynamics at 300 K. The entire systems were kept constrained, allowing the ions and the solvent molecules to equilibrate. They were then subjected to a series of dynamics calculations in which the constraints were gradually relaxed, until no constraints at all were applied. The final two production runs were performed for 15 ns without any restraints on the systems and co-ordinates were saved after every 5 ps for analysis of their trajectories [154]. These were analyzed using the Gromacs analysis tools, visualized by means of the VMD program [155] and graphs plotted using the Xmgrace program (<http://plasma-gate.weizmann.ac.il/Grace/>).

The root mean square deviation (RMSD) calculated from the simulation is a good indicator of the conformational stability of the simulated models. The result of these simulations, as revealed from the RMSD plots (Figure 4.6A), indicate that the azatrux-G-quadruplex complex is more stable than the G-quadruplex alone, with RMSDs of 2.3 Å and 2.7 Å, respectively. The role of the positively-charged side chains in the stabilization of the azatrux-G-quadruplex complex was also explored. The interactions between the piperidine side chains and the negatively-charged phosphodiester backbone were sampled over the course of the 15 ns simulation time. Figure 4.6B shows the distances between the positively-charged nitrogen atoms in the *N*-butyl piperidine side chains of the ligand and the negatively-charged oxygen atoms in the phosphodiester backbone of the quadruplex, calculated over the course of the 15 ns simulation. The two side chains that are positioned in the groove maintain stable interactions during the greater part of the simulation. This is due to the lack of rotation around the piperidine group in the side chain. However, the third side chain that is partially positioned over the G-quartet is highly

flexible. The molecular dynamics simulations highlight the ability of these side-chains to make multiple interactions with the phosphodiester backbone during the course of the simulation. These interactions impart rigidity to the interacting sugar-phosphate backbone and also provide a structural rationale for the stabilization of a propeller-type parallel stranded topology by azatrux.



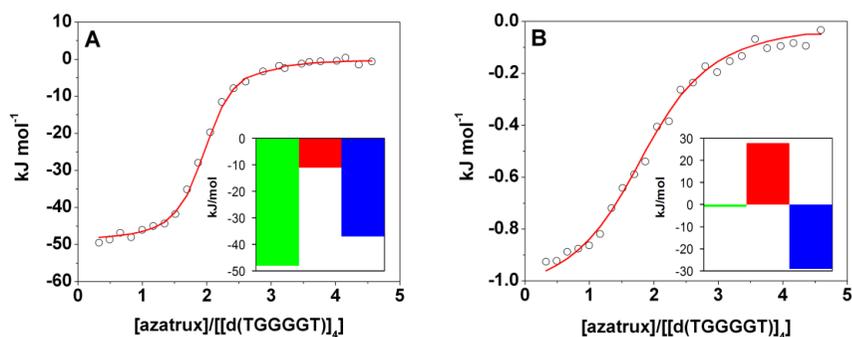
**Figure 4.6** (A) Root mean square deviation (RMSD) plot comparing the stability of the parallel-stranded G-quadruplex DNA (black) and azatrux-G-quadruplex complex (red). Analysis of the trajectories suggests that the azatrux-G-quadruplex complex structure is more stable than the native structure. (B) The distance calculated between the positively charged nitrogen atom in piperidine ring at the terminus of each ligand side chain of the ligand and the negatively-charged phosphate oxygen atoms in the phosphodiester backbone of the quadruplex, calculated over the course of the 15 ns simulation time. The three different colours represent the interactions for three different side chains of the ligand with the quadruplex. Side chains that fit in the groove (red and blue) make stable interactions while the side chain that sits on top of the phosphodiester backbone is relatively flexible.

#### 4.1.1.c. *Binding of azatrux to the parallel [d(TGGGGT)]<sub>4</sub> quadruplex*

To test the ability of azatrux to bind with an end-stacking mode to a parallel G-quadruplex topology, azatrux binding to a “standard” parallel G-quadruplex, the one formed by the d(TGGGGT) sequence in K<sup>+</sup> solution, has been examined [156]. The [d(TGGGGT)]<sub>4</sub> quadruplex has all-*anti* G-quartets, as in the parallel intramolecular 22-mer human telomeric quadruplex. The advantage of using the d(TGGGGT) sequence is that it forms a parallel tetramolecular quadruplex in dilute conditions (in the absence of PEG 200). This enables NMR titration experiments performed (NMR experiments are not straightforward in molecular crowding conditions) to determine structural features of the binding mode of azatrux to parallel quadruplexes. NMR experiments were performed by Prof. A. Randazzo at the School of Pharmacy of University of Naples.

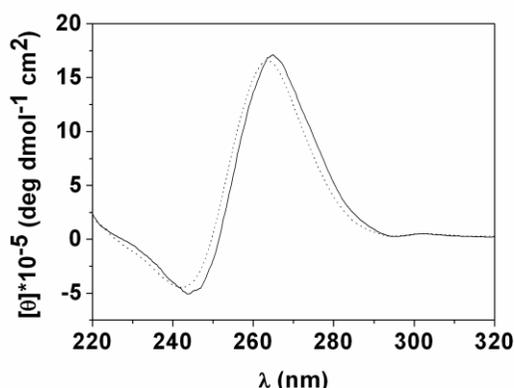
To compare the thermodynamic parameters derived for the binding of azatrux to the human telomeric sequence with the ones for the binding to [d(TGGGGT)]<sub>4</sub> quadruplex, ITC experiments were performed with [d(TGGGGT)]<sub>4</sub> in the presence and absence of PEG. ITC binding curves are monophasic and very similar in both instances (Figure 4.7).

The interpolation procedure with the experimental data has been carried out with a single set of identical sites model. In both crowding and dilute conditions, the stoichiometry of the azatrux-DNA complex is 2:1, confirming the stoichiometry obtained by NMR titration experiments. Interestingly, the thermodynamic parameters and the binding constant are different in the absence and presence of PEG. Particularly, the binding event is enthalpically driven in the presence of PEG and entropically driven in its absence, as shown by the thermodynamic signature of the processes (Table 4.1 and Figure 4.7). In the presence of PEG, the binding free energy change is more favourable and the corresponding binding constant is an order of magnitude higher than in the absence of PEG, indicating that crowding conditions increase ligand affinity for the quadruplex structure.



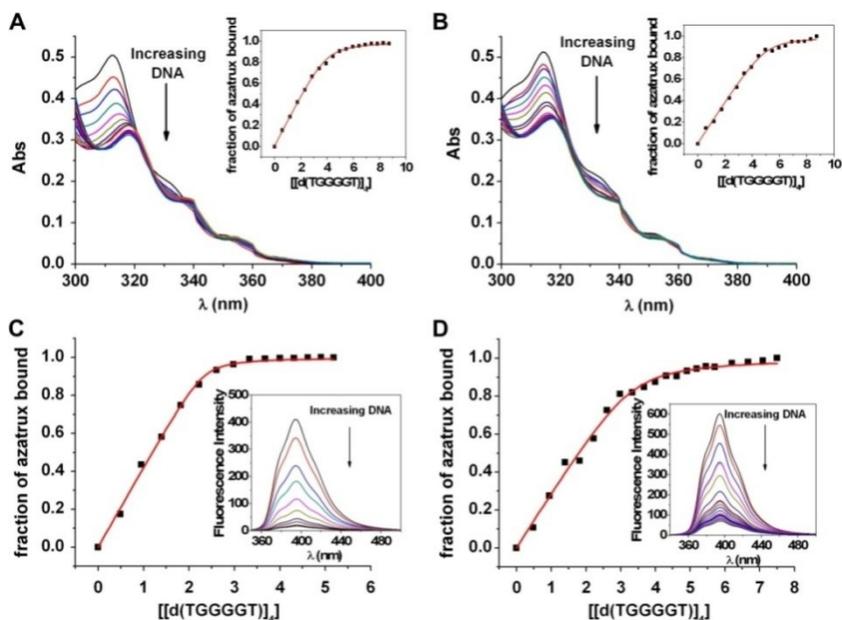
**Figure 4.7** Normalized heat of interaction between azatrux and the [d(TGGGGT)]<sub>4</sub> quadruplex obtained by ITC at 25 °C in a buffer solution containing 20 mM phosphate with 70 mM KCl, 0.1 mM EDTA, at pH 7.0, (A) in the presence of 40% PEG 200 or (B) in the absence of PEG. The open circles represent the experimental data obtained by integrating the raw data and subtracting the heat of ligand dilution into the buffer. The red line represents the best fit obtained by a nonlinear least-squares procedure based on a multiple and independent sites model. The insets represent the thermodynamic signature of the corresponding interaction: the enthalpic contribution ( $\Delta H^\circ$ ) is shown in green, the entropic one ( $T\Delta S^\circ$ ) in red, and the Gibbs energy change ( $\Delta G^\circ_{298K}$ ) in blue.

Since crowding conditions do not affect the parallel fold of the [d(TGGGGT)]<sub>4</sub> quadruplex, as shown by the CD spectrum (Figure 4.8), this effect can be directly correlated with the reduced water activity in PEG solution. The comparison of the binding constants and free energies (Table 4.1) obtained in crowding conditions for the d[AGGG(TTAGGG)<sub>3</sub>TT] and [d(TGGGGT)]<sub>4</sub> quadruplexes, reveals that the azatrux binds these two quadruplex structures with similar affinity, thus suggesting the same binding mode.



**Figure 4.8** CD spectra of the  $[d(TGGGGT)]_4$  quadruplex in dilute solution (dotted line) and in crowding conditions (solid line).

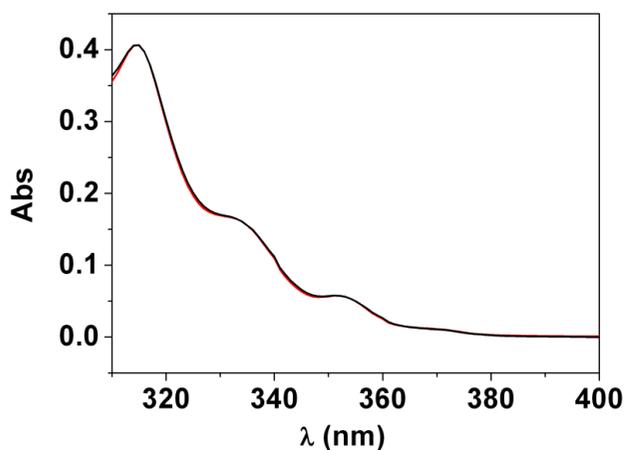
Finally, the same set of UV–Vis and fluorescence titration experiments performed with the human telomeric quadruplex was performed on the  $[d(TGGGGT)]_4$  quadruplex both in presence and absence of PEG. Figure 4.9 shows the UV and fluorescence titration experiments and the corresponding binding curves in both presence and absence of PEG. In all cases, the observed changes in the UV–Vis and fluorescence spectra of azatrux upon addition of  $[d(TGGGGT)]_4$  were very similar to the ones observed for the human telomeric sequence. Fitting of the binding curves with an independent and equivalent binding sites model confirms the 2:1 stoichiometry value found in the ITC experiments. The binding constants were:  $4 \times 10^6 \text{ M}^{-1}$  (UV in absence of PEG),  $4 \times 10^6 \text{ M}^{-1}$  (UV in presence of PEG) and  $9 \times 10^6 \text{ M}^{-1}$  (fluorescence in absence of PEG) and  $4 \times 10^6 \text{ M}^{-1}$  (fluorescence in presence of PEG).



**Figure 4.9** (A, B) UV-Vis absorption titration spectra of azatrux solutions in the absence and presence of successive additions of  $[d(TGGGGT)]_4$  quadruplex. The insets show the plot of the fraction of azatrux bound vs. the DNA concentration. (C, D) Fluorescence emission spectra (insets) of azatrux solutions in the absence and presence of successive additions of  $[d(TGGGGT)]_4$  quadruplex, and plot of the fraction of azatrux bound vs. the DNA concentration. The black squares represent the experimental data, the solid line is the best fit obtained with an independent sites model. All the titration experiments were performed at 25 °C in buffer solution containing 20 mM phosphate with 70 mM KCl, 0.1 mM EDTA and 40% (w/v) of PEG 200 at pH 7.0.

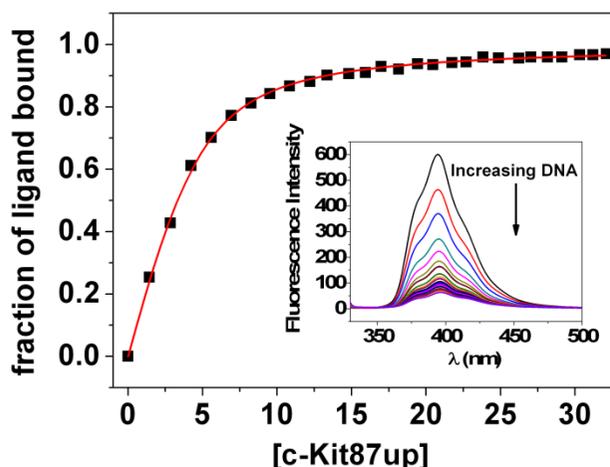
#### 4.1.1.d. Binding of azatrux to other biologically relevant DNA structures

The selectivity of azatrux for DNA quadruplex over DNA duplex was explored studying its affinity for the DNA duplex formed by the d(CGCGAATTCGCG) oligonucleotide. The UV spectrum of azatrux does not change on increasing DNA duplex concentration (Figure 4.10), suggesting that azatrux does not interact with duplex DNA.



**Figure 4.10** *UV-vis absorption spectra of (black line) 4 μM azatrux solution and (red line) of azatrux in the presence of an excess of duplex DNA (10 μM). The two spectra are superimposable.*

To evaluate the selectivity of azatrux for telomeric quadruplex over other biologically relevant quadruplex structures, its affinity for the c-Kit87up quadruplex was also explored. Fluorescence spectra of azatrux were recorded on increasing c-Kit87up concentration and the corresponding binding curves were fitted by means of an independent and equivalent site model (Figure 4.11). A stoichiometry of 2:1 and a binding constant of  $5 \times 10^5 \text{ M}^{-1}$  were found, revealing a weaker interaction with respect to that found for the telomeric quadruplex.



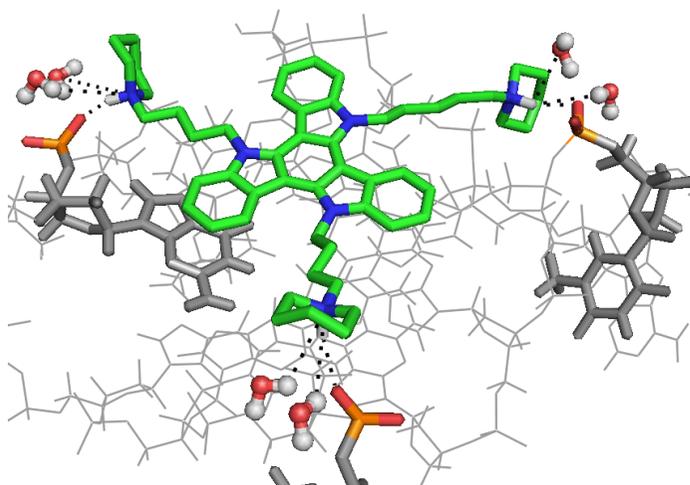
**Figure 4.11** Fluorescence emission spectra (inset) of 8  $\mu\text{M}$  azatrux solution in the absence and presence of successive additions of *c-Kit87up* quadruplex, in a buffer solution containing 20 mM phosphate with 70 mM KCl, 0.1 mM EDTA, 40% (w/v) of PEG 200, at pH 7.0, and plot of the fraction of azatrux bound vs. the concentration of DNA. The black squares represent the experimental data, the red line is the best fit obtained with an independent sites model.

#### 4.1.2. Discussion and Conclusions

Since the parallel fold is likely to be the dominant structure for the human telomeric DNA in physiological crowding conditions, it is important to understand the structural and energetic features governing the binding to this particular G-quadruplex topology. In order to induce the parallel topology in the d[AGGG(TTAGGG)<sub>3</sub>TT] human telomeric sequence, crowding condition were simulated by adding PEG 200 to the DNA solution. Fluorescence and UV titration experiments are consistent with the binding of one azatrux molecule to the DNA quadruplex. The CD spectrum of the d[AGGG(TTAGGG)<sub>3</sub>TT] in crowding condition is characteristic of the parallel topology and the binding of azatrux does not cause conformational change in the quadruplex. This observation is consistent with a model in which azatrux binds the parallel topology with an “external” binding mode, rather than an intercalative binding mode. Further, intercalation between G-quartets would be sterically prevented by the presence of the

propeller loops in the parallel G-quadruplex. However, the large hypochromic shift observed in the azatrux UV spectrum upon DNA addition suggests a mode of binding that involves extended stacking interactions between the azatrux aromatic core and DNA bases. The formation of extensive interactions between azatrux and DNA is also supported by the ITC data showing a large enthalpy change ( $-108$  kJ/mol) upon azatrux binding. A reasonable hypothesis thus involves an external binding mode with azatrux stacked on the terminal G-quartet plane of the G-quadruplex structure. Terminal stacking of small ligands has been observed in all crystal structures of quadruplex–ligand complexes [157,91,158], and is consistent with other biophysical studies [96,150]. The molecular modeling results support this hypothesis in showing that the aromatic core of azatrux can efficiently stack on the 3' terminal G-quartet without causing any major deformation of the quadruplex structure. This result is consistent with the CD results showing that azatrux binding does not change the quadruplex conformation. The molecular dynamic study enables the role of the *N*-butyl-piperidine side-chains in the binding to the parallel human quadruplex to be explored and indicates that the positively-charged nitrogen atoms of two piperidine side-chains interact with the negatively-charged phosphodiester backbone. These interactions, together with the stacking interaction between the azatrux core and the G-quartet surface, are consistent with the large enthalpy change observed by ITC for the formation of the azatrux–quadruplex complex. On the other hand, as shown by the comparison of the RMSD plots from the molecular dynamics simulations, these interactions impart rigidity to the interacting sugar-phosphate backbone in the complex, compared to the free quadruplex. This greater rigidity in the azatrux–quadruplex complex is consistent with the negative entropic contribution observed by ITC for the azatrux binding process. To further confirm the ability of azatrux to bind to the parallel quadruplex topology the binding of azatrux to the parallel-stranded [d(TGGGGT)]<sub>4</sub> quadruplex was examined. This quadruplex has all glycosidic angles populating *anti* conformations, as in the parallel human telomeric sequence under crowding conditions, but in contrast forms the same structure in the absence of crowding conditions. ITC experiments, performed in both the presence and absence of crowding conditions, show that azatrux binds the [d(TGGGGT)]<sub>4</sub> quadruplex with 2:1 stoichiometry. NMR data, obtained in the absence of crowding conditions, provide evidence for an end-stacking binding mode for

the azatrux molecules to each of the two terminal G-quartets. The thermodynamic signature of this binding in the presence of PEG 200 is in agreement with that found for the human telomeric sequence, also revealing an enthalpically driven binding process, thus suggesting a closely similar binding mode to the two quadruplexes. Interestingly, azatrux binds with a 2:1 stoichiometry the  $[d(TGGGGT)]_4$  quadruplex, but with a 1:1 stoichiometry the human telomeric quadruplex. A plausible reason for this discrepancy could be the different role played by the thymine bases in the two quadruplexes. In the  $[d(TGGGGT)]_4$ , the thymine bases are not structured and very flexible and they probably do not participate in the binding of azatrux, making the two terminal G-tetrads equivalent binding sites for the azatrux molecule. On the contrary, molecular modeling shows that, in the human telomeric sequence, the presence of the structured loop bases determines significant differences in the geometries of the two quadruplex extremities, making them not equivalent and allowing the binding of azatrux only to the 3'-end of the G-quadruplex.



**Figure 4.12** Schematic representation of water molecules associated with the side chains of azatrux. There are two solvent molecules associated per chain. There is a direct interaction between protonated nitrogen of azatrux and the phosphate oxygen atoms. At no point during the course of the simulation a water-mediated contact is observed. The dashed lines indicate the interactions between the solvent/quadruplex and the ligand. Quadruplex is coloured grey with the phosphate backbone and the contributing base in stick representation. The water molecules are represented as ball-and-sticks.

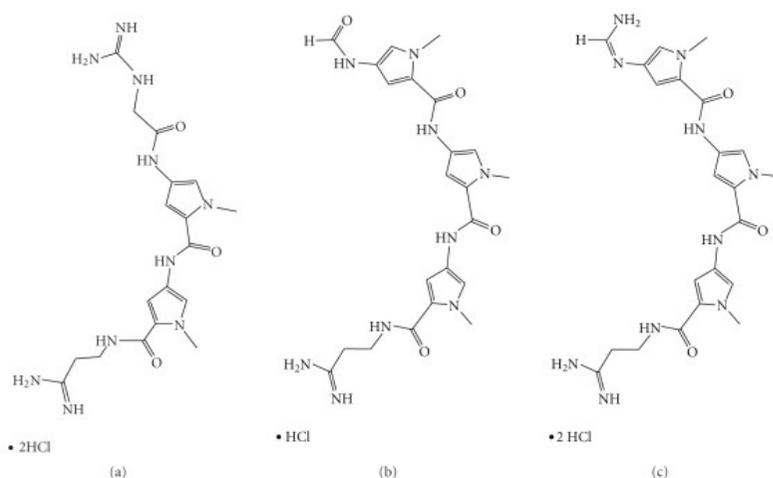
In conclusion, the experimental and computational results presented here support the hypothesis that azatrux preferably binds to parallel G-quadruplex structures, by end-stacking, and shows that the presence of a planar aromatic surface, as well as the side chains are important structural elements in optimizing this binding mode. The modeling also suggests that the length of two of the five-carbon linker chains in the azatrux molecule is optimal for ensuring effective interactions in the grooves. The third side-chain cannot reside in a groove due to the incompatibility of the three-fold pseudo-symmetry of the azatrux molecule with the four-fold symmetry of the quadruplex surface. This suggests that studies on alternative positions for this third substituent would be able to enhance binding and potentially selectivity as well. Further, molecular modeling showed the relevance of the electrostatic interactions in stabilizing the azatrux–quadruplex complex. This suggests that the polarity of the side chains is another critical parameter to vary in future studies to further modulate the G-quadruplex ligand binding properties.

## CHAPTER 5 Novel classes of potential anti-cancer agents binding to the grooves of G Quadruplex Structures

### 5.1. Selective Binding of Distamycin A Derivative to G-Quadruplex Structure (TGGGGT)<sub>4</sub>

In the past decade, a growing number of G-quadruplex binding agents have been proposed, most of them interact with the outer G-tetrads of the structures through  $\pi$ - $\pi$  stacking interactions [94]. The only groove binder experimentally proven to date is the distamycin A that interacts in a groove-binding mode with the quadruplex [d(TGGGGT)]<sub>4</sub>. This finding, along with the observation that distamycin derivatives could be effective inhibitors of the human telomerase [159], has stimulated other investigations. The importance of the crescent shape extension has been investigated by varying the pyrrole units number in distamycin A [108]. Particularly, it has been studied the interaction of two carbamoyl analogues of distamycin A, containing four and five pyrrole units, respectively. Experiments revealed that the presence of one additional pyrrole unit affects the affinity as well as the stoichiometry of the binding, whereas the addition of two pyrrole units leads to a total loss of interaction between the derivative and the [d(TGGGGT)]<sub>4</sub>.

The effect of a second cationic group, placed at the end of the molecule, on the interaction with DNA molecules, was evaluated. In particular, the interaction between the [d(TGGGGT)]<sub>4</sub> quadruplex and a new distamycin A derivative (compound **1**, Figure 5.1), where the initial formamide group is replaced by a charged N-formimidoil moiety have been studied.



**Figure 5.1** (a) Chemical structures of netropsin, (b) distamycin A, and (c) compound 1.

In addition, since compound **1** can also be considered as an analogue of netropsin because it presents one pyrrole unit more than netropsin (three instead of two), but two cationic ends like it (even if different groups), was compared the binding of compound **1** to  $[d(\text{TGGGGT})]_4$  with the binding of netropsin to the same target.

Finally, to investigate the selectivity of compound **1** for the G-quadruplex relative to duplex, was also performed a study of the interaction between the drug and the self-complementary DNA duplex  $d(\text{CGCGAATTCGCG})_2$ . This symmetric dodecamer was chosen because (a) it contains the central AATT core, which is considered the specific binding site for distamycin and netropsin; (b) the interaction with netropsin is well characterized in literature, from both structural and thermodynamic point of view.

Quadruplex and duplex samples were prepared by dissolving the lyophilised compound in a buffer solution containing 20 mM phosphate with 70 mM KCl, 0.1 mM EDTA at pH 7.0. The concentration of oligonucleotides was determined by UV adsorption measurements at 90°C using molar extinction coefficient values  $\epsilon_{(260\text{ nm})}$  of 57800 and 110700  $\text{M}^{-1}\text{ cm}^{-1}$  for  $d(\text{TGGGGT})$  and  $d(\text{CGCGAATTCGCG})$ , respectively.

Drug solutions have been prepared in the same buffer used for the oligonucleotides, and the concentration has been estimated by UV spectroscopy using the calculated extinction coefficient value  $\varepsilon_{(297\text{ nm})}$  of  $30000\text{ M}^{-1}\text{ cm}^{-1}$  for compound **1** and the reported extinction coefficient value  $\varepsilon_{(296\text{ nm})}$  of  $21500\text{ M}^{-1}\text{ cm}^{-1}$  for netropsin [160].

### 5.1.1. Results

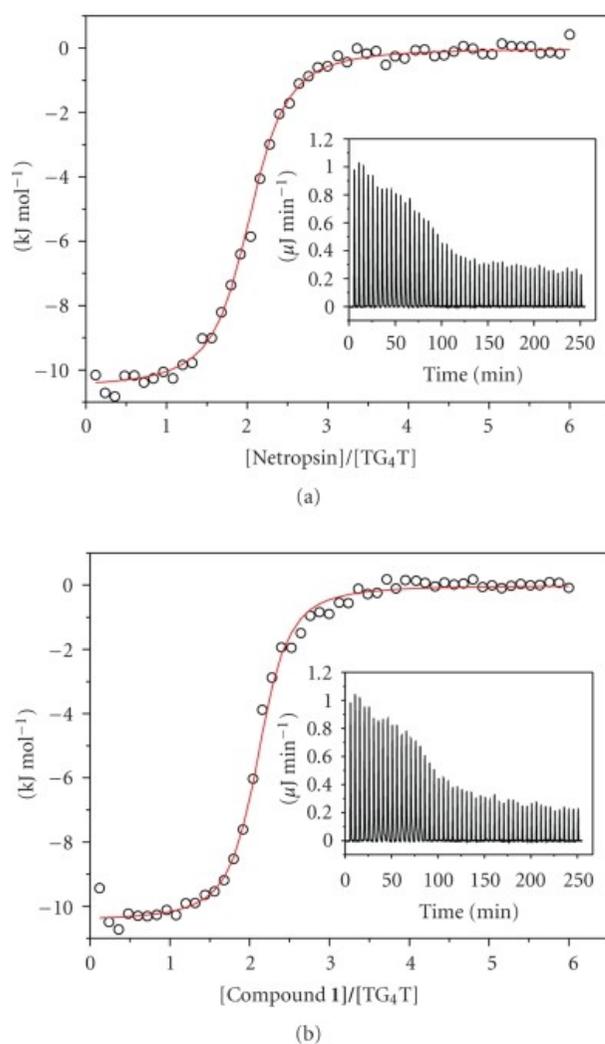
#### 5.1.1.a. ITC experiments

In Figure 5.2 are reported examples of raw ITC and integrated heat data for the titration of  $[\text{d}(\text{TGGGGT})_4]$  quadruplex with netropsin and compound **1** are shown. In each titration, volumes of  $5\text{--}10\ \mu\text{L}$  of a solution containing compound **1** or netropsin at a concentration of  $600\text{--}700\ \mu\text{M}$  were injected into a solution of quadruplex or duplex DNA ( $30\ \mu\text{M}$ ) in the same buffer, using a computer-controlled  $250\ \mu\text{L}$  microsyringe. The ITC data for drugs binding to  $[\text{d}(\text{TGGGGT})_4]$  indicate, in both cases, the formation of a 2:1 (drug:quadruplex) complex with good affinity.

The raw data for the titration of drugs with  $[\text{d}(\text{TGGGGT})_4]$  (insets in Figure 5.2) indicate an exothermic interaction, based on the positive values observed for the peaks. With each injection of ligand, less and less heat release was observed until constant values were obtained, thus reflecting, in both cases, a saturable process. The thermodynamic results, obtained from fitting the ITC data for netropsin and compound **1** binding to  $[\text{d}(\text{TGGGGT})_4]$ , are given in Table 5.1.

Ligand	$n$	$K_b$ ( $M^{-1}$ )	$\Delta_b H^\circ$ ( $kJ\ mol^{-1}$ )	$T\Delta_b S^\circ$ ( $kJ\ mol^{-1}$ )	$\Delta_b G^\circ_{(298K)}$ ( $kJ\ mol^{-1}$ )
Compound <b>1</b>	$2.0 \pm 0.1$	$1.9 (\pm 0.2) \times 10^6$	$-11.0 \pm 2.0$	$25 \pm 2$	$-36 \pm 2$
Netropsin	$2.0 \pm 0.1$	$1.2 (\pm 0.1) \times 10^6$	$-10.6 \pm 1.0$	$23 \pm 2$	$-34 \pm 2$

**Table 5.1** Thermodynamic parameters for the interaction of compound **1** and netropsin with  $[d(TGGGGT)]_4$  determined by ITC at 25°C and pH 7.0.

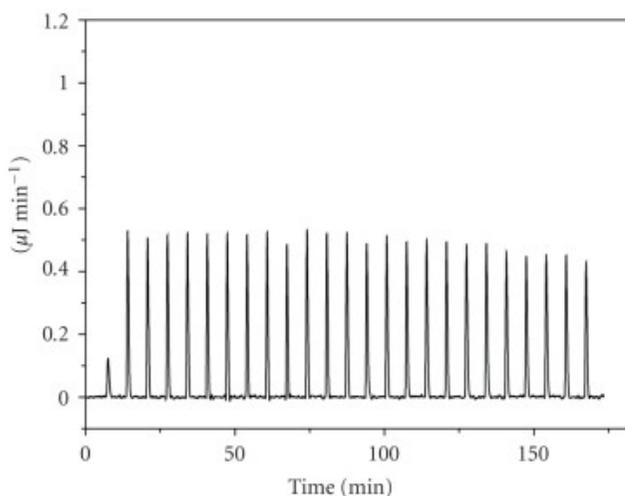


**Figure 5.2** Raw ITC data (insets) and binding isotherms for titration of  $[d(TGGGGT)]_4$  with netropsin (a) and with compound **1** (b).

The values of the binding constants and the Gibbs energy changes ( $-34 \text{ kJ mol}^{-1}$  and  $-36 \text{ kJ mol}^{-1}$  for netropsin and compound **1**, resp.) indicate that, from a thermodynamic point of view, the interactions with the quadruplex molecule are favoured at  $25^\circ\text{C}$ .

The values of  $\Delta H^\circ$  and  $T\Delta S^\circ$  show that in both cases the interactions are associated with a favourable binding enthalpy ( $-10.6 \text{ kJ mol}^{-1}$  and  $-11.0 \text{ kJ mol}^{-1}$  for netropsin and compound **1**, respectively), however, the binding processes are always entropically driven ( $23 \text{ kJ mol}^{-1}$  and  $25 \text{ kJ mol}^{-1}$  for netropsin and compound **1**, respectively).

Finally, experiments on the interaction of compound **1** with the DNA duplex  $d(\text{CGCGAATTCGCG})_2$  containing the central "AATT" core, which is considered the specific binding site for distamycin and netropsin, to evaluate if the structural modifications of compound **1** influence the binding, were performed. An example of the raw ITC data for the titration of the  $d(\text{CGCGAATTCGCG})_2$  duplex with compound **1** is shown in Figure 5.3. Resolvable binding isotherm was never obtained for the interaction of compound **1** with duplex using any combination of reactant concentrations, thus suggesting low affinity of the molecule for the investigated duplex.



**Figure 5.3** Raw ITC data for titration of  $d(\text{CGCGAATTCGCG})_2$  duplex with compound **1**. With each injection of ligand, constant heat release was observed due only to ligand dilution.

### 5.1.2. Discussion and Conclusions

Distamycin and netropsin have been recognized for decades as non intercalative DNA binding ligands that show specificity for the minor groove of dA•dT base pairs [160,161]. The binding of the drugs to the duplex DNA involves an electrostatic component from the cationic ends, hydrogen bonds from the amide NH groups, and van der Waals interactions with the wall of the groove. Some years ago, NMR studies indicated that, depending on DNA sequence, some binding sites can accommodate two distamycin molecules side-by-side in an antiparallel orientation [162]. In this 2:1 complex each ligand molecule preserves all the molecular recognition elements of groove binders. In contrast to distamycin, the netropsin binds only as a single molecule per binding site, suggesting that the side-by-side arrangement of two molecules is inhibited by charge repulsions. Both drugs have been shown little or no affinity for single-stranded DNA or RNA or for double-stranded RNA or DNA-RNA hybrids as well as they do not bind to the A helix or the left-handed Z-DNA [161,163].

Surprisingly, distamycin A has recently been shown to interact also with four-stranded parallel DNA quadruplexes [95,164,165]. Particularly, it was demonstrated, by using NMR and ITC methodology, that distamycin is able to interact with the quadruplex [d(TGGGGT)]<sub>4</sub>. Four ligand molecules bind as antiparallel dimers to the quadruplex in two opposite grooves, establishing hydrogen bonds with the guanine bases and strong coulombic interactions between the positively charged amidinium moiety of the ligand and the negative backbone phosphate groups of the quadruplex [95]. On the other hand, netropsin turned out to possess a lower affinity (NMR data) towards the quadruplex [d(TGGGGT)]<sub>4</sub> [164], even if, till now, this has never been confirmed by ITC. Interestingly, it seems that netropsin is not able to bind the quadruplex in dimeric form, most probably due to the doubly charged nature of molecule that prevents a side-by-side arrangement into the grooves.

In order to evaluate the binding properties of derivative **1**, and to perform a direct comparison with the binding behaviour of distamycin A and netropsin, [d(TGGGGT)]<sub>4</sub> has been titrated with **1** at the same experimental conditions (buffer, temperature, DNA concentration) used for distamycin A [164].

Isothermal titration calorimetry experiments were performed, indeed ITC is the only technique that directly measures the binding

enthalpy change for the formation of a complex, allowing the free energy change to be dissected into the enthalpic and entropic contribution to the association process. This reveals the nature of the forces that drive the binding and can provide insight into the nature of the intermolecular contacts formed and even into changes in solvation [166]. The understanding of those factors can be helpful in both screening among various drugs and optimizing the drug-target interactions, to direct the design of new drugs.

ITC experiments reveal that both compound **1** and netropsin bind to the investigated quadruplex. The thermodynamic profiles of the two drug-quadruplex interactions are qualitatively similar and in both cases the stoichiometry observed is 2:1 (drug:quadruplex). Interestingly, the thermodynamic parameters determined by ITC (Table 5.1) indicate that the association reactions of both ligands with  $[d(TGGGGT)]_4$  are entropically driven processes, even if the direct ITC measurements of the binding enthalpy change indicate that also the enthalpic contribution favours the associations. In the case of the interaction of distamycin with  $[d(TGGGGT)]_4$ , similar results were obtained (entropically driven process with a small favourable enthalpic contribution), except for stoichiometry and magnitude of binding constant.

As reported in the literature many times, a groove binding is generally characterized by a large favourable increase of entropy and a small favourable or unfavourable enthalpy change [167]. This could be mainly due to the fact that the association reaction is driven in large part by a hydrophobic effect and the favourable entropy is derived from the release of bound water molecules from the DNA and drug upon complex formation as well as release of counterions upon binding of cationic ligands. Interestingly, the entropically driven interactions of the compound **1** and of netropsin suggest a groove-binding mechanism also in these cases, as observed for distamycin A. Additionally, in contrast to many other groove binders, netropsin binding to duplex DNA is accompanied on average by a favourable enthalpy change. This is the result of both electrostatic interactions and hydrogen bonds formation between netropsin molecule and DNA. A favourable enthalpy change is observed by ITC for netropsin binding to  $[d(TGGGGT)]_4$  as well as for compound **1**, suggesting once again a similar groove binding mode.

As stated before, the stoichiometry obtained for the association of the investigated ligands to  $[d(TGGGGT)]_4$  is 2:1 (drug: quadruplex), while for distamycin a stoichiometry of 4:1 was

found. This finding is consistent with the observation that the distamycin molecules are able to bind as dimers, while netropsin (and probably also compound 1) is not able to dimerise because of charge repulsions. Notably, in the distamycin-[d(TGGGGT)]<sub>4</sub> complex, each distamycin dimer expands its bounded groove (similarly to that observed with duplex DNA), with concomitant reduction of the size of the adjacent ones, preventing a further interaction with other ligand molecules. In the binding of netropsin to duplex DNA the minor groove is widened by 0.5–2.0 Å by the entry of the drug [168]. The binding of the two netropsin molecules to [d(TGGGGT)]<sub>4</sub> could similarly expand the bounded grooves, reducing the size of the adjacent ones and preventing more interactions.

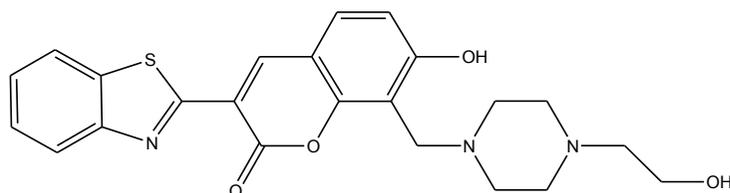
The experiments on the interaction of compound 1 with the DNA duplex d(CGCGAATTCGCG)<sub>2</sub> show that compound 1 has a poor affinity for the investigated duplex, while it is well known that both distamycin and netropsin have a high binding constant (between 10<sup>6</sup> and 10<sup>8</sup> M<sup>-1</sup>) for that target containing the specific binding site. Netropsin and compound 1 are able to bind to [d(TGGGGT)]<sub>4</sub> with good affinity, forming, in both cases, a 2:1 (drug : quadruplex) complex.

Very interestingly, this study shows that the structural modifications of compound 1 do not influence, compared to netropsin, the interaction with the quadruplex, but such modifications decrease the affinity of the ligand toward the duplex, thus enhancing the selectivity.

## **5.2. Virtual Screening-derived compounds and the G-Quadruplex Structure (TGGGGT)<sub>4</sub>: an ITC study**

As has been stated before, the evidence that the grooves of a quadruplex structure can be recognised by an organic molecule first appeared approximately three years ago [95]. An NMR study revealed that, distamycin A can bind in a dimeric form to the two opposite grooves of the quadruplex [d(TGGGGT)]<sub>4</sub> [164]. Whereas

the end-stacker ligands are the major part of all known quadruplex binders and their number is growing each day, quadruplex groove binders represent a quite unexplored and valuable field. In fact, grooves in duplex and quadruplex DNA are chemically and conformationally different, and since groove dimensions differ according to the type of quadruplex, groove binders can in principle be selective for a particular quadruplex topology. Thus, with the aim of finding brand new molecular scaffolds able to interact with the groove, in a previous study was performed an extensive structure-based virtual screening (VS) campaign [169], starting from the quadruplex structure found in complex with distamycin A. As a result, six molecules were found to be somehow groove binding agents by NMR and isothermal titration calorimetry (ITC) measurements. The study revealed that one of the six compounds is more effective in groove binding with respect to distamycin A (Figure 5.4) [170].

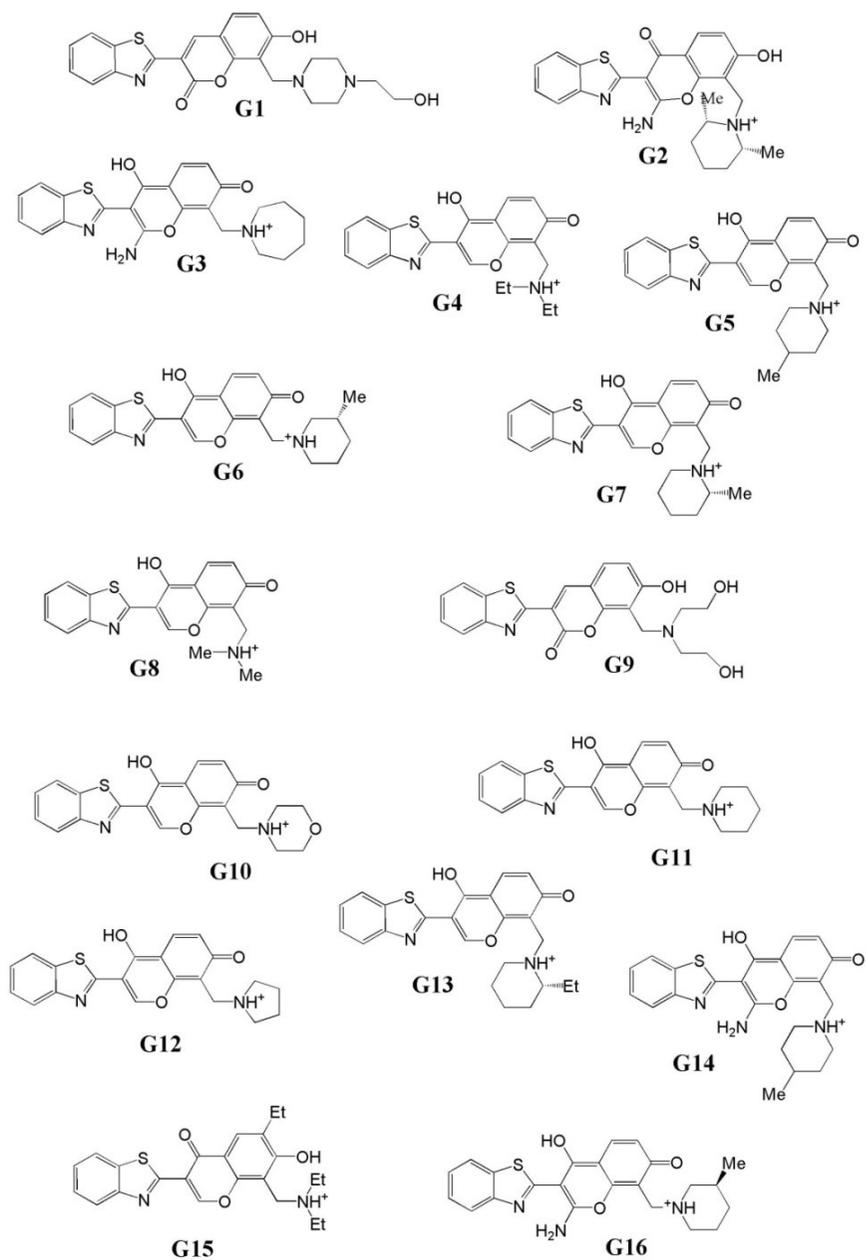


**Figure 5.4** Structure of the newly identified quadruplex binder as resulted from the virtual screening experiment and ITC experiments.

Starting from these results, was performed a thermodynamic study of the interaction between G-quadruplex structure  $[d(\text{TGGGGT})_4]$  and fifteen derivatives (**G2-G16**) of the compound **G1** selected in the previous study (Figure 5.5). The interaction has been studied by isothermal titration calorimetry (ITC), analysing the displacement of these ligands from the grooves of DNA-quadruplex by Distamycin A, the best groove binder identified until now. The  $d(\text{TGGGGT})$  and  $d(\text{CGCGAATTCGCG})$  oligonucleotide sequences used for this study were prepared by dissolving the lyophilised compound in 10 mM phosphate buffer

with 70 mM KCl, 0.2 mM EDTA, pH 7. The concentration of oligonucleotides was determined by UV adsorption measurements at 90°C using molar extinction coefficient values  $\epsilon_{(260\text{ nm})}$  of 57800 and 110700  $\text{M}^{-1}\text{ cm}^{-1}$  for d(TGGGGT) and d(CGCGAATTCGCG), respectively. The molar extinction coefficients were calculated by the nearest neighbour model [139]. Stock solutions of the sixteen drugs were prepared by solubilising weighted amounts in DMSO to a final concentration of 8 mM. The complexes between the quadruplex and the drugs were prepared diluting the drug stock solution into the quadruplex solution to get a final DNA:drug molar ratio of 1:4 and a final DMSO concentration of 7%. Distamycin A was solubilised in the same buffer used for the oligonucleotide containing 7% of DMSO.

The competition ITC experiments performed, represent a possible strategy to evaluate the efficiency of a groove binder compared to distamycin A, even if this strategy doesn't allow to directly measure all the thermodynamic parameters involved in the binding processes.



**Figure 5.5** Structures of the derived compounds from **G1** quadruplex binder.

## 5.2.1. Results

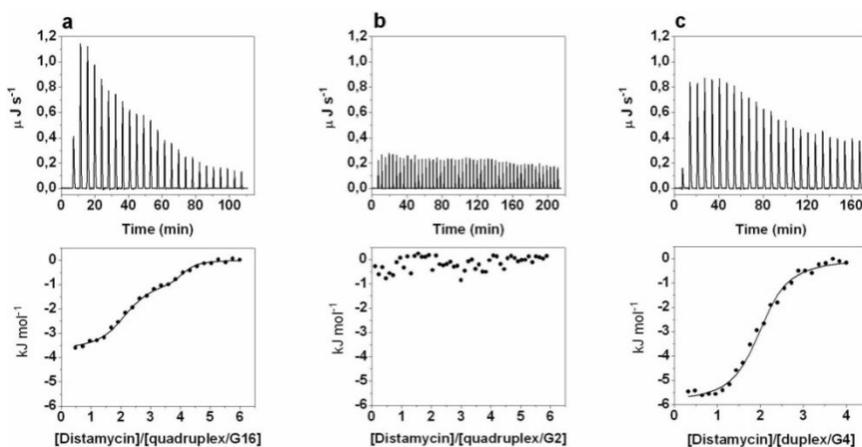
### 5.2.1.a. ITC experiments

To investigate the binding affinity of the compounds **G1-G16** for the [d(TGGGGT)]<sub>4</sub> G-quadruplex, ITC experiments were performed. Efforts to obtain direct thermodynamic information from canonical ITC experiments [152], in which a G-quadruplex solution is titrated with a ligand solution in the identical aqueous buffer, have failed due to solubility issues regarding the ligands. Instead of direct measurements, displacement experiments were effectively carried out [171], in which was analyzed the binding of distamycin A to the G-quadruplex previously saturated with a ligand. Indeed, despite the solubility issues, mixtures of G-quadruplex with each compound were successfully prepared by dissolving the ligands in DMSO and diluting them in the aqueous buffer solution containing the G-quadruplex. Once the complex is formed, the ligand will inhibit the binding of distamycin A if its affinity for the G-quadruplex is higher than the latter, conversely, it will be displaced by titration with the stronger binder. The titrations were carried out in 10 mM phosphate buffer, 70 mM KCl, 0.2 mM EDTA, 7% DMSO, pH 7, at 298 K. In each experiment, volumes of 10  $\mu$ L of a 720  $\mu$ M distamycin A solution were added into a 30  $\mu$ M solution of quadruplex-DNA complex, using a computer-controlled 250- $\mu$ L microsyringe, with a spacing of 200-400 s between each injection.

Figure 5.6a and Figure 5.6b show two representative ITC displacement experiments carried out by titrating with distamycin A the [d(TGGGGT)]<sub>4</sub> G-quadruplex saturated with compounds **G16** and **G2**, respectively. The results obtained in the two experiments are very different and clearly show two different phenomena. In the first case, the ITC profile for the titration of ligand-saturated G-quadruplex is essentially identical to the one obtained for the binding of distamycin A to the G-quadruplex alone, with two distinct binding events, each involving two ligand molecules, to give a final 4:1 stoichiometry [95,170], indicating that the presence of **G16** does not affect the interaction. On the contrary, the titration of [d(TGGGGT)]<sub>4</sub>/**G2** mixture with distamycin A (b) gives completely different results. In this case, the ITC data show constant heat release after each injection of

distamycin A, only due to ligand dilution, proving that it is no longer able to interact with the G-quadruplex.

Overall, ITC experiments showed that four out of sixteen compounds (**G1**, **G2**, **G3** and **G4**) have ability to bind to G-quadruplex DNA more tightly than distamycin A, inhibiting its interaction, whereas the other compounds bind rather weakly to the G-quadruplex. To evaluate a possible selectivity of the best ligands for the G-quadruplex over duplex, identical ITC displacement experiments by titrating with distamycin A the DNA duplex  $d(\text{CGCGAATTCGCG})_2$  saturated with compounds **G1-G4**, were performed. This self-complementary dodecamer was chosen because it contains the central AATT core, which is considered the specific binding site for distamycin A [172]. Examples of raw ITC data and binding isotherm for the titration of the  $d(\text{CGCGAATTCGCG})_2/\mathbf{G4}$  mixture with distamycin A are shown in Figure 5.6c. As can be seen, at each injection of distamycin A solution, less and less heat release was measured until constant values were obtained, implying a saturable process. The binding isotherm resulting from the integration of raw data shows a typical sigmoidal binding curve centered on a 2:1 stoichiometry and clearly show that the presence of compounds **G1-G4** in the mixture has no effect on the interaction of distamycin A with the duplex (Table 5.2).



**Figure 5.6** Raw ITC data (top panels) and normalized heat (bottom panels) for titration of  $[d(\text{TGGGGT})_4/\mathbf{G16}$  (a),  $[d(\text{TGGGGT})_4/\mathbf{G2}$  (b), and  $d(\text{CGCGAATTCGCG})_2/\mathbf{G4}$  (c) mixtures with distamycin A at 25 °C. The normalized heat for the titrations (filled circles) was obtained by integrating the raw data and subtracting the heat of the ligand dilution. The lines represent nonlinear least squares fit of the data to the appropriate binding model.

	$n_1$	$k_1$ (M <sup>-1</sup> )	$\Delta G_1$ kJ/mol	$\Delta H_1$ kJ/mol	$T\Delta S_1$ kJ/mol	$n_2$	$k_2$ M <sup>-1</sup>	$\Delta G_2$ kJ/mol	$\Delta H_2$ kJ/mol	$T\Delta S_2$ kJ/mol
<b>quadrulex/G5</b>	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
<b>quadrulex/G6</b>	1.9	5x10 <sup>7</sup>	-44	-3.0	41	2.1	4.2x10 <sup>6</sup>	-38.0	-1.0	37.0
<b>quadrulex/G7</b>	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
<b>quadrulex/G8</b>	1.9	1x10 <sup>7</sup>	-39	-8.3	31	2.0	6.0x10 <sup>7</sup>	-44.0	-4.2	40
<b>quadrulex/G9</b>	2.2	3.7x10 <sup>7</sup>	-43	-4.0	39	1.9	3.2x10 <sup>6</sup>	-37	-1.2	36
<b>quadrulex/G10</b>	2.0	2.0x10 <sup>7</sup>	-42	-4.0	38	1.9	4.0x10 <sup>6</sup>	-37	-1.7	36
<b>quadrulex/G11</b>	1.8	4.0x10 <sup>7</sup>	-43	-6.1	37	2.0	6.3x10 <sup>6</sup>	-39	-0.6	38
<b>quadrulex/G12</b>	2.0	4.7x10 <sup>6</sup>	-38	-1.5	36	2.0	1.4x10 <sup>7</sup>	-40	-4.3	36
<b>quadrulex/G13</b>	2.1	4.2x10 <sup>7</sup>	-44	-4.2	39	1.8	2.3x10 <sup>6</sup>	-36	-1.6	34
<b>quadrulex/G14</b>	1.9	3.2x10 <sup>7</sup>	-43	-6.3	37	2.2	5.5x10 <sup>6</sup>	-38	-2.0	36
<b>quadrulex/G15</b>	1.8	4.0x10 <sup>7</sup>	-44	-6.5	38	2.2	4.2x10 <sup>5</sup>	-32	-4.3	28
<b>quadrulex/G16</b>	1.9	2.5x10 <sup>7</sup>	-42	-3.7	38	2.1	1.6x10 <sup>6</sup>	-35	-1.2	34
<b>distamycin</b>	2.0	9.4x10 <sup>6</sup>	-40	-3.9	36	2.0	6.0x10 <sup>5</sup>	-33	-1.2	32
<b>Duplex/G1</b>	ND	ND	ND	ND	ND					
<b>Duplex/G2</b>	ND	ND	ND	ND	ND					
<b>Duplex/G3</b>	1.8	9.0x10 <sup>5</sup>	-34	-12	22					
<b>Duplex/G4</b>	2.0	6.0x10 <sup>5</sup>	-33	-6	27					

**Table 5.2** Thermodynamic parameters for the interaction of Distamycin A with [d(TGGGGT)<sub>n</sub>] and its complexes formed with the G1-G4 are not reported since they abolish the binding of Distamycin A; Thermodynamic parameters for the interaction of Distamycin A with duplex and its complexes formed with the G5-G16 are not reported since they abolish the binding of Distamycin A. The experimental error for each thermodynamic property is <7%

## 5.2.2. Discussion and Conclusions

A previous structure-based virtual screening (VS) campaign [169] and a subsequent NMR and ITC study [170] resulted in the identification of a molecule (compound **G1**) able to interact with the groove of the quadruplex [d(TGGGGT)]<sub>4</sub> stronger than distamycin A.

**G1** derivatives were then synthesized by Prof. Novellino from the department of Pharmaceutical Chemistry and Toxicology of University of Naples “Federico II”. In this context the aim was to evaluate the interaction of the ligands for the quadruplex [d(TGGGGT)]<sub>4</sub> by means of isothermal titration calorimetry (ITC), analysing their displacement from the grooves of DNA-quadruplex by distamycin A, the best groove binder identified until now. The competition ITC experiments represent a possible strategy to evaluate the efficiency of a groove binder compared to distamycin A.

The ITC experiments allowed to select four molecules (**G1-G4**) that bind [d(TGGGGT)]<sub>4</sub> with greater affinity than distamycin A. As it is widely accepted that specificity for DNA G-quadruplexes is a fundamental requirement for the quadruplex binder to become a drug, binding tests towards DNA duplex have been performed. Moreover to test the activity and toxicity of ligands have been performed biological assays by Dr. Biroccio from the National Cancer Institute “Regina Elena”.

The ITC study combined with biological data allowed to select four ligands that are more effective in groove binding with respect to distamycin A. Compounds **G3** and **G4** bind to the quadruplex structures with higher affinity compared to the duplex DNA. Additionally, the biological characterization of the new ligands demonstrates their ability to induce selective DNA damage at telomeric level and the induction of apoptosis and senescence on tumour cells. These results finally demonstrate that induction of tumour cell death can be efficiently achieved by compounds able to selectively target the G-quadruplex grooves. These results reinforce the notion that G-quadruplex binding compounds can act as broad inhibitors of telomere-related processes and could represent potential selective antineoplastic drugs. This project was designed and developed in collaboration with the department of Pharmaceutical Chemistry and Toxicology of University of

Naples “Federico II” and the National Cancer Institute “Regina Elena” of Rome.

## Concluding remarks

In conclusion, targeting G-quadruplex-DNA represents a high scientific challenge since this particular DNA arrangement has been found in biologically relevant area of genome and is less abundant as compared to canonical duplex-DNA. In addition, ligands able to bind selectively to G-quadruplexes may have important biological effects for therapeutic purposes. In particular, the results obtained quite convincingly pave the way for the exploitation of G-quadruplex ligands as tools to evaluate the therapeutic potential of telomeres.

Indeed, over the past decade, the G-quadruplex ligand field has developed exponentially. A glimpse at the advances made in the design and the synthesis of G-quadruplex ligands leave us convinced that the development of compounds able to discriminate not only G-quadruplex from duplex-DNA, but between the various structures of G-quadruplexes is imminent. In this perspective, it is important that investigators aim at a deeper understanding of the structural aspects of human telomere folding and the energetic aspects of quadruplex-drug molecular recognition. Many classic or newly adopted physico-chemical techniques can be very useful towards this direction, in this field ITC is clearly the method of choice to measure the free energy, enthalpy, and stoichiometry of a quadruplex-drug binding reaction.

Overall research in the field of quadruplex ligands is getting more and more exciting as knowledge increases. In the future, G-quadruplex chemists and biologists acting in conjunction could very well provide new molecular principles that may find applications both in molecular and cellular biology and may contribute to the development of novel anti-cancer therapies. Through the course of my studies as a Ph.D. student, I hope to have contributed a small piece of a much larger puzzle which in turn may help the future of cancer research.

## References

---

- [1] Blackburn E.H. and Gall J.G. **1978**, *J. Mol. Biol.* 120, 33–53.
- [2] Meyerson M., **2000**, *J. Clin. Oncol.*, 18, 2626–2634.
- [3] Kim N. W., Piatyszek M.A., Prowse K.R., Harley C.B., West M.D., Ho P.L., Coviello G.M., Wright W.E., Weinrich S.L. & Shay J.W., **1994**, *Science*, 266, 2011–2015.
- [4] Shay J.W. & Wright W.E., **2006**, *Nat Rev Drug Discov.*, 5, 577–584
- [5] Xu Y., **2011**, *Chem. Soc. Rev.*, 40, 2719–2740
- [6] Neidle S.; Parkinson G.N. **2003**, *Curr. Opin. Struct. Biol.* 13, 275
- [7] Zahler A.M, Williamson J.R, Cech T.R & Prescott D.M., **1991**, *Nature* 350, 718–720.
- [8] Zhong Z., Shiue L., Kaplan S. and de Lange T., **1992**, *Mol. Cell. Biol.*, 12, 4834.
- [9] Bianchi A., Smith S., Chong L., Elias P. and de Lange T., **1997**, *EMBO J.*, , 16, 1785.
- [10] Baumann P. and Cech T. R., **2001**, *Science*, 292, 117.
- [11] Zaug A.J., Podell E.R. and Cech T.R., 2005, *Proc. Natl. Acad. Sci. U. S. A.*, 102, 10864
- [12] Houghtaling B.R., Cuttonaro L., Chang W. and Smith S., **2004**, *Curr. Biol.*, 14, 1621.
- [13] Ye J.Z., Hockemeyer D., Krutchinsky A.N., Loayza D., Hooper S.M., Chait B.T. and de Lange T., **2004**, *Genes Dev.*, 18, 1649.
- [14] Li B., Oestreich S. and de Lange T., **2000**, *Cell (Cambridge, Mass.)*, 101, 471.
- [15] Zhu X.D., Kuster B., Mann M., Petrini J.H and de Lange T., **2000**, *Nat. Genet.*, 25, 347.
- [16] Liu D., Safari A., O’Connor M. S., Chan D. W., Laegeler A., Jun Q. and Zhou S. Y., **2004**, *Nat. Cell Biol.*, 6, 673.
- [17] Hockemeyer D., Palm W., Else T., Daniels J.P., Takai K.K., Ye J.Z., Keegan C.E., de Lange T. and Hammer G.D., **2007**, *Nat. Struct. Mol. Biol.*, 14, 754.
- [18] Loayza D. and de Lange T., **2003**, *Nature*, 424, 1013.
- [19] Lei M., Podell E.R. and Cech T.R., **2004**, *Nat. Struct. Mol. Biol.*, 11, 1223.
- [20] Giraldo R., Suzuki M., Chapman L. and Rhodes D., **1994**, *Proc. Natl. Acad. Sci. U. S. A.*, 91, 7658.

- 
- [21] Harley, C.B.; Futcher, A.B.; Greider, C.W. **1990**, *Nature*, *345*, 458-460.
- [22] Kim, N.W. **1997**, *Eur J Cancer*, *33*, 781-786.
- [23] Weinrich S.L., Pruzan R., Ma L., Ouellette M., Tesmer V.M., Holt S.E., Bodnar, A.G., Lichtsteiner S., Kim N.W., Trager J. B., Taylor R.D., Carlos R., Andrews W.H., Wright W.E., Shay J.W., Harley C.B., Morin G.B., **1997**, *Nat Genet*, *17*, 498-502.
- [24] Shay J.W., Wright W.E., **2002**, *Cancer Cell*, *2*, 257-265.
- [25] Feng J., Funk W.D., Wang S.S., Weinrich S.L., Avilion A.A., Chiu C.P., Adams R.R., Chang E., Allsopp R. C., Yu J., Le S. Y., West M.D., Harley C.B., Andrews W.H., Greider C.W., Villeponteau B., **1995**, *Science*, *269*, 1236-1241.
- [26] Nakamura T.M., Morin G.B., Chapman K.B., Weinrich S.L., Andrews W.H., Lingner J., Harley C.B., Cech T.R., **1997**, *Science*, *277*, 955-959.
- [27] Beattie T.L., Zhou W., Robinson M. O., Harrington L., **2001**, *Mol Cell Biol*, *21*, 6151-6160.
- [28] Blasco M.A. **2005**, *Nat. Rev. Genet.*, *6*, 611-622.
- [29] Blackburn E.H. **2005**, *FEBS Lett.*, *579*, 859-862.
- [30] Weinrich S.L., Pruzan R., Ma L., Ouellette M., Tesmer V.M., Holt, S.E.; Bodnar A.G., Lichtsteiner S., Kim N. W., Trager J.B., Taylor R. D., Carlos R., Andrews W. H., Wright W. E., Shay J. W., Harley C. B., Morin G. B., **1997**, *Nat Genet*, *17*, 498-502.
- [31] Beattie T. L., Zhou W., Robinson M.O., Harrington L., **1998**, *Curr Biol.*, *8*, 177-180.
- [32] Nugent C.I., Lundblad V., **1998**, *Genes Dev*, *12*, 1073-1085.
- [33] Kim N.W., Piatyszek M.A., Prowse K.R., Harley C.B., West M.D., Ho P.L., Coviello G.M., Wright W.E., Weinrich S.L., Shay J.W., **1994**, *Science*, *266*, 2011-2015.
- [34] Shay J.W., Gazdar A.F., **1997**, *J ClinPathol*, *50*, 106-109.
- [35] Sharma S., Raymond E., Soda H., Sun D., Hilsenbeck S.G., Sharma A., Izbicka E., Windle B., Von Hoff D. D, **1997**, *Ann Oncol.*, *8*, 1063-1074.
- [36] Mergny J.L., Helene C., **1998**, *Nat. Med.*, *4*, 1366-1367.
- [37] Neidle S., Kelland L.R., **1999**, *Anticancer Drug Des*, *14*, 341-347.
- [38] Gellert G.C., Jackson S.R., Dikmen Z.G., Wright W.E., Shay J.W., **2005**, *Drug Discovery Today: Disease Mechanism*, *2*, 159-164.

- 
- [39] Granotier C., Pennarun G., Riou L., Hoffschir F., Gauthier LR., De Cian A., Gomez D., Mandine E., Riou J.F., Mergny J.L., Mailliet P., Dutrillaux B., Boussin F.D., **2005**, *Nucleic Acids Res.*, 33:4182–4190.
- [40] Chang C.C., Wu J.Y., Chien C.W., Wu W.S., Liu H, Kang C.C., Yu L.J., Chang T.C. **2003**, *Anal Chem.*, 75:6177–6183.
- [41] Gellert M., Lipsett M.N., Davies D.R., **1962**, *Proc. Natl. Acad. Sci. USA*, 48, 2013-2018.
- [42] Simonsson T., **2001**, *Bio lChem*, 382, 621-628.
- [43] Burge S., Parkinson G.N., Hazel P., Todd A.K., Neidle S., **2006**, *Nucleic Acids Res.*, 34, 5402-5415.
- [44] Phan A.T., **2010** *FEBS J.*, 277(5):1107-17 Review
- [45] Sasisekharan V., Zimmerman S., Davies D.R. J., **1975**, *Mol. Biol.*, 92, 171-174.
- [46] Pinnavaia T.J., Marshall C.L., Mettler C.M., Fisk C.L., Miles H.T., Becker E.D., **1978**, *J. Am. Chem. Soc.*, 100, 3625-3627.
- [47] Simonsson T., Sjoback R.J., **1999**, *Biol. Chem.*, 274, 17379-17383.
- [48] Laughlan G., Murchie A.I., Norman D. G., Moore M.H., Moody P.C., Lilley D.M., Luisi B., **1994**, *Science*, 265, 520-524.
- [49] Chen F.M., **1992**, *Biochemistry*, 31, 3769-3776.
- [50] Guschlbauer W., Chantot J.F., Thiele D., **1990**, *J. Biomol. Struct. Dyn.*, 8, 491-511.
- [51] Strahan G.D., Keniry M.A., Shafer R.H., **1998**, *Biophys. J.*, 75, 968-981.
- [52] Kerwin S.M., **2000**, *Curr. Pharm. Des.*, 6, 441-478.
- [53] Wang Y., Patel D.J., **1995**, *J. Mol. Biol.*, 251, 76-94.
- [54] Kang C., Zhang X., Ratliff R., Moyzis R., Rich A., **1992**, *Nature*, 356, 126-131.
- [55] Wang Y & Patel DJ, **1993**, *Structure*, 1, 263–282.
- [56] Dai J., Punchihewa C., Ambrus A., Chen D., Jones R.A., Yang D., **2007**, *Nucleic Acids Res.*, 35, 2440.
- [57] Lim K.W., Amrane S., Bouaziz S., Xu W., Mu Y. Patel D.J., Luu K.N. Phan, A.T. **2009**, *J. Am. Chem. Soc.*, 131, 4301
- [58] Ambrus A., Chen D., Dai J., Bialis T., Jones R.A., Yang D., **2006**, *Nucleic Acids Res.*, 34, 2723.
- [59] Phan A.T., Kuryavyi V., Luu K.N., Patel D.J., **2007**, *Nucleic Acids Res.*, 35, 6517.

- 
- [60] Phan A.T., Luu K.N., Patel D.J., **2006**, *Nucleic Acids Res.*,34, 5715.
- [61] Dai J., Carver M., Punchihewa C., Jones R.A., Yang D., **2007**, *Nucleic Acids Res.*,35, 4927.
- [62] Parkinson GN., Lee MPH & Neidle S.,**2002**, *Nature*,417, 876–880.
- [63] Xue Y, Kan Z.Y., Wang Q, Yao Y, Liu J, HaoY.H. & Tan Z., **2007**, *J Am ChemSoc*, 129, 11185–11191.
- [64] Petraccone L, Trent J.O. & Chaires J.B., **2008**, *J Am Chem Soc*, 130, 16530–16532.
- [65] Wang Y & Patel D.J.,**1993**, *Structure*, 1, 263–282.
- [66] Campbell N.H., Patel M, Tofa A.B., Ghosh R, Parkinson G.N. & Neidle S., **2009**, *Biochemistry*, 48, 1675–1680.
- [67] Haider S, Parkinson G N & Neidle S., 2008, *Biophys J*, 95, 296–311.
- [68] Pomerantz A.K., Moerner W.E. & Kool E.T., **2008**, *J Phys Chem B*, 112, 13184–13187.
- [69] Yu H.Q., Miyoshi D. & Sugimoto N., **2006**, *J Am ChemSoc*, 128, 15461–15468.
- [70] Vorlickova M, Chladkova J, Kejnovska I, Fialova M &Kypř J, **2005**, *Nucleic Acids Res*, 33, 5851–5860.
- [71] Wang Y & Patel D.J., **1992**,*Biochemistry*, 31, 8112–8119.
- [72] Phan A.T. & Patel D.J., **2003**, *J Am Chem Soc*, 125, 15021–15027.
- [73] Zhang N, Phan A.T. & Patel D.J., **2005**, *J Am ChemSoc*,127, 17277–17285.
- [74] Petraccone L., Trent J.O. and Chaires J.B., **2008**, *J.Am. Chem. Soc.*, 130, 16530–16532.
- [75] Petraccone L., Garbett N.C., Chaires J.B. and Trent J.O., **2010**, *Biopolymers*, 93, 533–548.
- [76] Xu Y., Ishizuka T., Kurabayashi K. and Komiyama M., **2009**, *Angew. Chem. Int. Ed. Engl.*, 48,7833–7836.
- [77] Collie G.W., Parkinson G.N., Neidle S., Rosu F., De Pauw E. and Gabelica V., **2010**, *J. Am. Chem. Soc.*, 132,9328–9334
- [78] Gray, R. D.; Li, J.; Chaires, J. B. **2009**, *J. Phys. Chem. B*, 113, 2676
- [79] Petraccone L, Spink C.H., Trent J.O, Garbett N.C., Mekmaysy C.S., Giancola C, Chaires J.B. *J Am Chem Soc.* (Just Accepted November **2011**)
- [80] Minton A.P., Wilf J., **1981**, *Biochemistry*, 20:4821–4826.
- [81] Allen P. Minton, **2001**, *J Biol Chem.*, 276 (14):10577-80.

- 
- [82] Heddi B, Phan A.T., **2011**, *J Am Chem Soc.*, 133(25):9824-33.
- [83] Miyoshi D., Nakao A., Sugimoto N., **2002**, *Biochemistry*, 41, 15017–15024.
- [84] Zhou J., Wei C., Jia G., Wang X., Tang Q., Feng Z., Li C., **2008**, *Biophys Chem.*, 136(2-3):124-7
- [85] Minton A.P., **1997**, *Curr. Opin. Biotech.*, 8, 65–69.
- [86] Arola A. and Vilar R., *Curr. Top. Med. Chem.*, 2008, **8**, 1405–1415
- [87] Monchaud D. and Teulade-Fichou M. P., **2008**, *Org. Biomol. Chem.*, **6**, 627–636
- [88] Georgiades S. N., Abd Karim N. H., Suntharalingam K. and Vilar R., **2010**, *Angew. Chem., Int. Ed.*, 49, 4020–4034
- [89] Parkinson G.N., Ghosh R. & Neidle S., **2007**, *Biochemistry*, 46, 2390–2397.
- [90] Campbell N.H., Parkinson G.N., Reszka A.P. & Neidle S., **2008**, *J Am Chem Soc.*, 130, 6722–6724.
- [91] Haider S.M., Parkinson G.N. & Neidle S., **2003**, *J Mol Biol*, 326, 117–125
- [92] Phan A.T., Kuryavyi V, Gaw H.Y. & Patel D.J., **2005**, *Nat Chem Biol*, 1, 167–173.
- [93] Dixon I.M., Lopez F, Tejera A.M., Este've J.P., Blasco M.A., Pratviel G. & Meunier B., **2007**, *J Am Chem Soc*, 129, 1502–1503.
- [94] Neidle S., **2009**, *Current Opinion in Structural Biology*, 19, (3), 239–250.
- [95] Martino L, Virno A, Pagano B, Virgilio A, Di Micco S, Galeone A, Giancola C, Bifulco G, Mayol L & Randazzo A, **2007**, *J Am Chem Soc*, 129, 16048–16056.
- [96] Martino L, Pagano B, Fotticchia I, Neidle S, Giancola C., **2009**, *J Phys Chem B.*, 113(44):14779-86.
- [97] Cummaro A, Fotticchia I, Franceschin M, Giancola C, Petraccone L., **2011**, *Biochemie*, 93(9):1392-400.
- [98] Ginnari-Satriani L., Casagrande V., Bianco A., Ortaggi G., Franceschin M., **2009**, *Org Biomol Chem.*, 7(12):2513-6.
- [99] Petraccone L., Fotticchia I., Cummaro A., Pagano B., Ginnari-Satriani L., Haider S., Randazzo A., Novellino E., Neidle S., Giancola C., **2011**, *Biochemie*, 93(8):1318-27
- [100] Ababou A., Ladbury J. E., **2006**, *J Mol Recognit*, 19, 79-89.
- [101] Doyle M. L, **1997**, *Curr Opin Biotechnol*, 8, 31-35.
- [102] Haq I., Ladbury J., **2000**, *J Mol Recognit*, 13, 188-197.

- 
- [103] Pierce M. M., Raman C. S., Nall B. T., **1999**, *Methods*, 19, 213-221.
- [104] Leavitt S., Freire E., **2001**, *Curr. Opin. Struct. Biol.*, 11, 560-566.
- [105] O'Brien R., Ladbury J.E., Chowdry B.Z., **2001**, *Oxford University Press: Oxford*, 263-286.
- [106] Wiseman T., Williston S., Brants J.F., Lin L.N., **1989**, *Anal. Biochem.*, 179, 131-137.
- [107] Buurma N.J., Haq I., **2007**, *Methods*, 42(2):162-72.
- [108] Pagano B., Virno A., Mattia C.A., Mayol L., Randazzo A., and Giancola C., **2008**, *Biochimie*, 90 (8), 1224–1232
- [109] Ranjbar B., Pooria G., **2009**, *Chem Biol Drug Des.*, 74, 101-120.
- [110] Warshaw M.M., Cantor C.R. **1970**, *Biopolymers*;9:1079–1103.
- [111] Cantor C.R., Schimmel P.R., **1980**, *Biophysical Chemistry*, W. H. Freeman: San Francisco, CA, Vol. 2
- [112] Bloomfield V.A., Crothers D.M., Tinoco I.Jr., with contribution from, Hearst J.E., Wemmer D.E., Kollman P.A., Turner D.H. **2000**, *Nucleic Acids: Structures, Properties, and Functions*. Sausalito, California, USA: University Science Books.
- [113] Ranjbar B., Gill P., **2009**, *Chem Biol Drug Des.*, 74(2):101-20
- [114] Masiero S., Trotta R., Pieraccini S., De Tito S., Perone R., Randazzo A. and Spada G. P., **2010**, *Org. Biomol. Chem.*, 8, 2683–2692
- [115] Balagurumoorthy P., Brahmachari S. K., Mohanty D., Bansal M. and Sasisekharan V., **1992**, *Nucleic Acids Res.*, 20, 4061
- [116] Gray D.M. and Bollum F.J., **1974**, *Biopolymers*, 13, 2087
- [117] Lu M., Guo Q. and Kallenbach N.R., **1993**, *Biochemistry*, 32, 598
- [118] Dai J., Carver M., Yang D., **2008**, *Biochimie*, 90, 1172.
- [119] Arora A, Balasubramanian C, Kumar N, Agrawal S, Ojha RP, Maiti S. **2008**, *FEBS J.* 275(15):3971-83
- [120] Kimura T., Kawai K., Fujitsuka M., Majima T., **2006**, *Chem. Commun.*, 401–402
- [121] Betts M. J., Sternberg M. J., **1999**, *Protein Eng*, 12, 271-283.
- [122] Totrov M., Abagyan R., **1994**, *Nat Struct Biol*, 1, 259-263.
- [123] Strynadka N. C., Eisenstein M., Katchalski-Katzir E., Shoichet B. K., Kuntz I. D., Abagyan R., Totrov M., Janin J., Cherfils J., Zimmerman F., Olson A., Duncan B., Rao

- 
- M., Jackson R., Sternberg M., James M. N., *Nat Struct Biol*, **1996**, 3, 233-239.
- [124] Pattabiraman N., Langridge R., **1985**, *J Biomol Struct Dyn.*, 2(4):683-92
- [125] Shi D.F., Wheelhouse R.T., Sun D.Y., Hurley L.H., **2001**, *J. Med. Chem.*, 44, 4509.
- [126] Izbicka E., Wheelhouse R.T., Raymond E., Davidson K.K., Lawrence R.A., Sun D., Windle, B.E., Hurley L.H., Von Hoff D.D., **1999**, *Cancer Res.*, 59, 639.
- [127] Wheelhouse R.T., Sun D., Han H., Han F.X., Hurley L.H., **1998**, *J. Am. Chem. Soc.*, 120, 3261.
- [128] Zhang H.J., Wang X.F., Wang P., Ai X.C., Zhang J.P., **2008**, *Photochem. Photobiol. Sci.*, 7, 948
- [129] Ren J., Chaires J. B., **1999**, *Biochemistry*, 38, 16067.
- [130] Wang P., Ren L., He H., Liang F., Zhou X., Tan Z., **2006**, *ChemBiochem.*, 7, 1155.
- [131] Haq I., Trent J.O., Chowdhry B.Z., Jenkins T.C., **1999**, *J. Am. Chem. Soc.*, 121, 1768.
- [132] Han H.; Langley D.R.; Rangan A.; Hurley L.H. **2001**, *J. Am. Chem. Soc.*, 123, 8902.
- [133] Mita, H.; Ohyama, T.; Tanaka, Y.; Yamamoto, Y. **2006**, *Biochemistry*, 45, 6765.
- [134] Wei C.; Jia G.; Yuan J.; Feng Z.; Li C. **2006**, *Biochemistry*, 45, 6681.
- [135] Parkinson G.N.; Ghosh R.; Neidle S. **2007**, *Biochemistry*, 46, 2390.
- [136] Han F.X.G.; Wheelhouse R.T.; Hurley L.H. **1999**, *J. Am. Chem. Soc.*, 121, 3561.
- [137] Gaynutdinov T.I., Neumann R.D., Panyutin I.G., **2008**, *Nucleic Acids Res.* 36, 4079.
- [138] Wei C., Jia G., Zhou J., Han G., Li C., **2009**, *Phys. Chem. Chem. Phys.* 11, 4025
- [139] Cantor C.R., Warshaw M.M., Shapiro H., **1970**, *Biopolymers*, 9, 1059.
- [140] Phan A.T., Kuryavyi V., Gaw H.Y., Patel D.J. **2005**, *Nat. Chem. Biol.* 1, 167.
- [141] Pasternack R.F., Gibbs E.J., Villafranca J.J., **1983**, *Biochemistry*, 22 2406-2414
- [142] Roselin L.S., Lin M.S., Lin P.H., Chang Y. and Chen W.Y., **2010**, *Biotechnol. J.*, 5, 85–98.

- 
- [143] Falconer R.J., Penkova A., Jelesarov I. and Collins B.M., **2010**, *J. Mol. Recognit*, 23,395–413.
- [144] Ghazaryan A.A., Dalyan Y.B., Haroutiunian S.G., Tikhomirova A., Taulier N., Wells J.W. and Chalikian T.V., **2006**, *J. Am. Chem. Soc.*, 128, 1914–1921
- [145] Yamashita T., Uno T. and Ishikawa Y., **2005**, *Bioorg. Med. Chem.*, 13,2423–2430.
- [146] Garbett N.C., Ragazzon P.A. and Chaires J.B., *Nat. Protoc.*, 2, **2007**, 3166–3172.
- [147] Jia G., Feng Z., Wei C., Zhou J., Wang X. and Li C., **2009**, *J. Phys. Chem. B*, 113, 16237–16245.
- [148] Zhao C., Ke-wei Z., Yu-hua H. and Zheng T., **2009**, *J. Am. Chem. Soc.*, 131, 10430–10438.
- [149] Han H., Langley D.R., Rangan A. and Hurley L.H., **2001**, *J. Am. Chem. Soc.*, 123,8902–8913.
- [150] Hudson J.S., Brooks S.C. and Graves D.E., **2009**, *Biochemistry*, 48,4440–4447.
- [151] Ginnari-Satriani L., Casagrande V., Bianco A., Ortaggi G. and Franceschin M., **2009**, *Org. Biomol. Chem.*, 7,2513–2516.
- [152] Pagano B., Mattia C.A., Giancola C. **2009**, *International Journal of Molecular Sciences*. 10(7):2935–2957.
- [153] Pagano B., Mattia C.A., Cavallo L., Uesugi S., Giancola C. and Fraternali F., **2008**, *J. Phys. Chem. B*, 112, 12115–12123.
- [154] Haider S. and Neidle S., **2010**, *Methods Mol. Biol. (Clifton, NJ)*, 608,17–37.
- [155] Humphrey W., Dalke A. and Schulten K., **1996**, *J. Mol. Graph.*, 14,33–38 27–38.
- [156] Laughlan G., Murchie A.I., Norman D.G., Moore M.H., Moody P.C., Lilley D.M. and Luisi B., **1994**, *Science (New York, N.Y)*, 265,520–524.
- [157] Neidle S. and Parkinson G.N., **2008**, *Biochimie*, 90,, 1184–1196.
- [158] Parkinson G.N., Cuenca F. and Neidle S., **2008**, *J. Mol. Biol.*,381, 1145–1156
- [159] Zaffaroni N., Lualdi S., Villa R. **2002**; *European Journal of Cancer*. 38(13):1792–1801.
- [160] Rentzeperis D., Marky L.A., Dwyer T.J., Geierstanger B.H., Pelton J.G., Wemmer D.E., **1995**; *Biochemistry*. 34(9):2937–2945

- 
- [161] Wartell R.M., Larson J.E., Wells R.D., **1974**, *The Journal of Biological Chemistry*. 249(21):6719–6731
- [162] Pelton J.G., Wemmer D.E., **1990**, *Journal of the American Chemical Society*. 112(4):1393–1399
- [163] Zimmer C. **1975**, *Progress in Nucleic Acid Research and Molecular Biology*. 15:285–318.
- [164] Randazzo A., Galeone A., Mayol L., **2001**, *1H Chemical Communications*. 11:1030–1031.
- [165] Cocco M.J., Hanakahi L.A., Huber M.D., Maizels N., **2003**, *Nucleic Acids Research*.; 31(11):2944–2951.
- [166] Haq I., **2002**, *Archives of Biochemistry and Biophysics*. 403(1):1–15.
- [167] Chaires J.B., **2006**, *Archives of Biochemistry and Biophysics*. 453(1):24–29.
- [168] Kopka ML, Yoon C, Goodsell D, Pjura P, Dickerson RE. **1985**, *Proceedings of the National Academy of Sciences of the United States of America*. 82(5):1376–1380
- [169] Cosconati S., Marinelli L., Trotta R., Virno A., Mayol L., Novellino E., A.J. Olson and Randazzo A., **2009**, *J.Am. Chem. Soc.*, 131, 16336–16337
- [170] Trotta R., De Tito S., Lauri I., La Pietra V., Marinelli L., Cosconati S., Martino L., Conte M.R., Mayol L., Novellino E., Randazzo A., **2011**, *Biochimie*, 93(8):1280-7.
- [171] Velazquez Campoy A., Freire E., **2005**, *Biophys. Chem*. 115, 115-124.
- [172] Klevit R.E., Wemmer D.E., Reid B.R., **1986**, *Biochemistry* 25, 3296-3303.

---

## Publications

1. Martino L., Pagano B., **Fotticchia I.**, Neidle S., Giancola C., **Shedding light on the interaction between TMPyP4 and human telomeric quadruplexes**, *J Phys Chem B.*, 2009,113(44):14779-86.

### Abstract

The nature of the binding mode and stoichiometry of the TMPyP4 cationic porphyrin to G-quadruplex structures continues to be controversial, with no consensus model to date, especially for intramolecular G-quadruplexes from human telomeric sequences. Those sequences possess intricate polymorphism in solution that appears to be reduced under molecular crowding conditions in which the parallel structure appears to be the most populated one. We have performed a systematic study, in dilute solution and under molecular crowding conditions, of the binding reactions between TMPyP4 and four G-quadruplexes formed by different truncations of human telomeric DNA, with 5'- or 3'-flanking bases, using isothermal titration calorimetry and circular dichroism. The results clearly indicate that all of these G-quadruplexes are able to bind up to four TMPyP4 molecules. CD studies show that interaction with TMPyP4 promotes the conversion of the hybrid structures to an antiparallel conformation in dilute solution, while under molecular crowding conditions the interaction does not promote any conformational change. ITC reveals in both cases that the binding process comprises two sequential events, a first in which one molecule of TMPyP4 interacts with the quadruplex structures and a second in which three other molecules bind to the structures. The selectivity of TMPyP4 for the quadruplex relative to duplex DNA was also investigated under molecular crowding conditions showing that TMPyP4 has enhanced selectivity for quadruplex DNA compared to the duplex structure. This finding reinforces the potential applications of TMPyP4.

- 
2. Pagano B., **Fotticchia I.**, De Tito S., Mattia C.A., Mayol L., Novellino E., Randazzo A., Giancola C., **Selective Binding of Distamycin A Derivative to G-Quadruplex Structure [d(TGGGGT)]<sub>4</sub>**, *J Nucleic Acids.*, 2010, 30.

### **Abstract**

Guanine-rich nucleic acid sequences can adopt G-quadruplex structures stabilized by layers of four Hoogsteen-paired guanine residues. Quadruplex-prone sequences are found in many regions of human genome and in the telomeres of all eukaryotic organisms. Since small molecules that target G-quadruplexes have been found to be effective telomerase inhibitors, the identification of new specific ligands for G-quadruplexes is emerging as a promising approach to develop new anticancer drugs. Distamycin A is known to bind to AT-rich sequences of duplex DNA, but it has recently been shown to interact also with G-quadruplexes. Here, isothermal titration calorimetry (ITC) and NMR techniques have been employed to characterize the interaction between a dicationic derivative of distamycin A (compound 1) and the [d(TGGGGT)]<sub>4</sub> quadruplex. Additionally, to compare the binding behaviour of netropsin and compound 1 to the same target, a calorimetric study of the interaction between netropsin and [d(TGGGGT)]<sub>4</sub> has been performed. Experiments show that netropsin and compound 1 are able to bind to [d(TGGGGT)]<sub>4</sub> with good affinity and comparable thermodynamic profiles. In both cases the interactions are entropically driven processes with a small favourable enthalpic contribution. Interestingly, the structural modifications of compound 1 decrease the affinity of the ligand toward the duplex, enhancing the selectivity.

- 
3. Giancola C., Ercole C., **Fotticchia I.**, Spadaccini R., Pizzo E., D'Alessio G., Picone D., **Structure-cytotoxicity relationships in bovine seminal ribonuclease: new insights from heat and chemical denaturation studies on variants**, *FEBS J.*, 2011, 278(1):111-22.

### **Abstract**

Bovine seminal ribonuclease (BS-RNase), a homodimeric protein displaying selective cytotoxicity towards tumor cells, is isolated as a mixture of two isoforms, a dimeric form in which the chains swap their N-termini, and an unswapped dimer. In the cytosolic reducing environment, the dimeric form in which the chains swap their N-termini is converted into a noncovalent dimer (termed NCD), in which the monomers remain intertwined through their N-terminal ends. The quaternary structure renders the reduced protein resistant to the ribonuclease inhibitor, a protein that binds most ribonucleases with very high affinity. On the other hand, upon selective reduction, the unswapped dimer is converted in two monomers, which are readily bound and inactivated by the ribonuclease inhibitor. On the basis of these considerations, it has been proposed that the cytotoxic activity of BS-RNase relies on the 3D structure and stability of its NCD derivative. Here, we report a comparison of the thermodynamic and chemical stability of the NCD form of BS-RNase with that of the monomeric derivative, together with an investigation of the thermal dissociation mechanism revealing the presence of a dimeric intermediate. In addition, we report that the replacement of Arg80 by Ser significantly decreases the cytotoxic activity of BS-RNase and the stability of the NCD form with respect to the parent protein, but does not affect the ribonucleolytic activity or the dissociation mechanism. The data show the importance of Arg80 for the cytotoxicity of BS-RNase, and also support the hypothesis that the reduced derivative of BS-RNase is responsible for its cytotoxic activity.

- 
- Cummaro A., **Fotticchia I.**, Franceschin M., Giancola C., Petraccone L., **Binding properties of human telomeric quadruplex multimers: a new route for drug design.**, *Biochemie*, 2011, 93(9):1392-400.

#### **Abstract**

Human telomeric G-quadruplex structures are known to be promising targets for an anticancer therapy. In the past decade, several research groups have been focused on the design of new ligands trying to optimize the interactions between these small molecules and the G-quadruplex motif. In most of these studies, the target structures were the single quadruplex units formed by short human DNA telomeric sequences (typically 21-26 nt). However, the 3'-terminal single-stranded human telomeric DNA is actually 100-200 bases long and can form higher-order structures by clustering several consecutive quadruplex units (multimers). Despite the increasing number of structural information on longer DNA telomeric sequences, very few data are available on the binding properties of these sequences compared with the shorter DNA telomeric sequences. In this paper we use a combination of spectroscopic (CD, UV and fluorescence) and calorimetric techniques (ITC) to compare the binding properties of the (TTAGGG)(8)TT structure formed by two adjacent quadruplex units with the binding properties of the (AG(3)TT)(4) single quadruplex structure. The three side-chained triazatruxene derivative azatrux and TMPyP4 cationic porphyrin were used as quadruplex ligands. We found that, depending on the drug, the number of binding sites per quadruplex unit available in the multimer structure was smaller or greater than the one expected on the basis of the results obtained from individual quadruplex binding studies. This work suggests that the quadruplex units along a multimer structure do not behave as completely independent. The presence of adjacent quadruplexes results in a diverse binding ability not predictable from single quadruplex binding studies.

- 
5. Petraccone L., **Fotticchia I.**, Cummaro A., Pagano B., Ginnari-Satriani L., Haider S., Randazzo A., Novellino E., Neidle S., Giancola C., **The triazatruxene derivative azatrux binds to the parallel form of the human telomeric G-quadruplex under molecular crowding conditions: biophysical and molecular modeling studies**, *Biochemie*, 2011, 93(8):1318-27

**Abstract**

The present study has employed a combination of spectroscopic, calorimetric and computational methods to explore the binding of the three side-chained triazatruxene derivative, termed azatrux, to a human telomeric G-quadruplex sequence, under conditions of molecular crowding. The binding of azatrux to the tetramolecular parallel [d(TGGGGT)](4) quadruplex in the presence and absence of crowding conditions, was also characterized. The data indicate that azatrux binds in an end-stacking mode to the parallel G-quadruplex scaffold and highlights the key structural elements involved in the binding. The selectivity of azatrux for the human telomeric G-quadruplex relative to another biologically relevant G-quadruplex (c-Kit87up) and to duplex DNA was also investigated under molecular crowding conditions, showing that azatrux has good selectivity for the human telomeric G-quadruplex over the other investigated DNA structures.