Università degli Studi di Napoli "Federico II"
Facoltà di Farmacia

Dottorato di Ricerca in "*Scienza del Farmaco*" XXI Ciclo

Indirizzo
Chimica delle Sostanze Naturali Bioattive

Dipartimento di Chimica delle Sostanze Naturali

G-Quadruplex DNA: novel targets and drugs

Coordinatore
Chiar.ma Prof.ssa M. Valeria D'Auria

Tutor:
Chiar.mo Prof. Antonio Randazzo

Candidata
Dott.ssa Ada Virno

INDEX

Abstract

G-quadruplexes are higher-order DNA and RNA structures formed by G-rich sequences based on tetrads of hydrogen-bonded guanine bases.
The important role of these structures in biological systems lies mainly on three features, since they are involved in: i) the architecture of telomeres of many organisms; ii) G-rich sequences that are present within a wide range of genes; iii) the scaffolds of several oligonucleotide aptamers.
For these reason, the interest in G-quadruplexes is witnessed by hundreds of reports dealing with various aspects of these DNA secondary structures, such as the chemical nature and the topology of the backbone and the loops (where present), the conformation of the glycosydic linkage, as well as the influence of structural modifications on the stability of the complexes. The structural variability of the quadruplexes has been further increased by introducing modification on both bases and sugar-phosphate backbone. In this frame, unprecedented DNA quadruplex structures containing a 3'-3' or 5'-5' inversion of polarity site in the G-tract are presented. The quadruplexes are characterized by different elements of symmetry and glycosidic angle conformations. Concerning the base modifications, we report the effect of the incorporation of 2'-deoxy-8-(hydroxyl) adenosine on the stability of quadruplexes formed by modified human telomeric DNA and the biophysical properties of quadruplexes containing two or three 8-bromodeoxyguanosine residues. Furthermore, we have also considered the effect of the introduction of an A-residue into a quadruplex forming oligonucleotide containing a 5'-5' polarity of inversion site (see chapter 1 below).

Aptamers are nucleic acid macromolecules that bind to molecular targets, including proteins, with high affinity and specificity. The aptamer drug TBA (Thrombin Binding Aptamer) is a thrombin inhibitor in development for use as an anticoagulant during coronary artery bypass graft procedures. In order to improve the properties of TBA, a number of researches have been devoted to its structure-activity relationship to post-SELEX modifications. In this frame, we have undertaken the synthesis and the study of a mini library composed of several TBA analogues containing a 3'-3' or a 5'-5' inversion of polarity site at different positions into the sequence. Particularly, the aim of this study was to use a biological driven approach in order to improve the knowledge of the interactions between thrombin and TBA that are critical for the biological activity. We have designed and synthesized a TBA based oligodeoxynucleotide containing a 5'-5' site of polarity inversion, namely $^{3'}$GGT$^{5'}$-$^{5'}$TGGTGTGGTTGG$^{3'}$ (**1**). Here we report a detailed structural study, along with thermal stability analysis and a structure-activity relationship of this new modified TBA.

Although TBA was the first aptamer to be discovered and considering the advantageous characteristics of LNAs, no studies dealing with the effects of substitution of a regular nucleotide with an LNA residue in the TBA sequence have been reported yet. In light of the above, in this work we report the study of four new TBA based oligonucleotides containing LNA residues, namely $^{5'}$ggttggtgtggttgg$^{3'}$ (1), $^{5'}$ggTTggTGTggTTgg$^{3'}$ (2), $^{5'}$gGTTGGTGTGGTTGG$^{3'}$ (3) and $^{5'}$GGTTGGTGTGGTTGg$^{3'}$ (4), where upper case and lower case letters represent DNA and LNA residues, respectively. Moreover, the chemical-physical properties and the three-dimensional characterization, based on NMR and CD

spectroscopy, associated with molecular mechanics and dynamics calculations, of 4 are also reported (see chapter 2 below).

Telomere protects the ends of the chromosome from damage and recombination and its shortening has been implicated in cellular senescence. Telomeric DNA consists of tandem repeats of simple short sequences, rich in guanine residues. In the presence of metal ions such as K^+ or Na^+, telomeric DNA can form G-quadruplexes.

Telomerase, the enzyme which elongates the G-rich strand of telomeric DNA, is active in about 85% of tumors, leading the cancer cells to infinite lifetime. The inhibition of telomerase has become an attractive strategy for the anticancer therapy and, because telomerase requires a single-stranded telomeric primer, the formation of G-quadruplex complexes by telomeric DNA inhibits the telomerase activity. Furthermore, small molecules that stabilize G-quadruplex structures have been found to be effective telomerase inhibitors and, then, the use of drugs to target G-quadruplexes is emerging as a promising way to interfere with telomere replication in the tumors cells and to act as anticancer agents. Distamycin-A has been shown to interact with four-stranded parallel DNA quadruplex containing oligonucleotides of different sequence. The complex between distamycin A and the parallel DNA quadruplex $[d(TGGGGT)]_4$ has been studied by ^1H NMR spectroscopy and isothermal titration calorimetry (ITC). In order to unambiguously assert that distamycin A interacts with the grooves of the quadruplex $[d(TGGGGT)]_4$, we have analyzed the NMR titration profile of a modified quadruplex, namely $[d(TGG^{Me}GGT)]_4$, and we have applied the recently developed Differential-Frequency Saturation Transfer Difference (DF-STD) method, for assessing the ligand–DNA binding mode. The three-dimensional structure of the 4:1 distamycin A /

$[d(TGGGGT)]_4$ complex has been determined by an in-depth NMR study followed by dynamics and mechanics calculations. All results unequivocally indicate that distamycin molecules interact with $[d(TGGGGT)]_4$ in a 4:1 binding mode, with two antiparallel distamycin dimers that bind simultaneously two opposite grooves of the quadruplex. The affinity between Dist-A and $[d(TGGGGT)]_4$ enhances (~10-fold) when the ratio of distamycin A to the quadruplex is increased. In this paper we report the first three-dimensional structure of a groove-binder molecule complexed to a DNA quadruplex structure.

Furthermore, in this work, isothermal titration calorimetry (ITC) and ^{1}H-NMR studies have been conducted to examine the binding of distamycin A and its two carbamoyl derivatives (compounds **1** and **2**) to the target $[d(TGGGGT)]_4$ and $d[AG_3(T_2AG_3)_3]$ quadruplexes from the Tetrahymena and human telomeres, respectively. The interactions were examined using two different buffered solutions containing either K^+ or Na^+ at a fixed ionic strength, to evaluate any influence of the ions present in solution on the binding behaviour. Experiments reveal that distamycin A and compound **1** bind the investigated quadruplexes in both solution conditions, conversely compound **2** appears to have a poor affinity in any case. Moreover, these studies indicate that the presence of different cations in solution affects the stoichiometry and thermodynamics of the interactions (see chapter 3 below).

"Every day I remind myself that my inner and outer life are based on the labors of other men, living and dead, and that I must exert myself in order to give in the same measure as I have received and am still receiving."

Albert Einstein

CHAPTER 1

Hunting G-quadruplexes

1. G-quadruplexes: structures

The double helix structure of duplex DNA is well known. The two antiparallel strands are held together by complementary base pairing between adenine (A) and thymine (T), and between guanine (G) and cytosine (C), using two hydrogen bonds in AT, and three in GC. However, this is not the only base pairing arrangement that can occur between bases, and alternative base pairing leads to alternative structures, including triple-stranded and four-stranded structures, called a G-quadruplex. G-quadruplexes, as the name suggests, have a core that is made up of guanine bases only, with four guanines arranged in a rotationally symmetric manner, making hydrogen bonds from N1–O6 and N2–N7 around the edges of the resulting square (see Fig. 1).

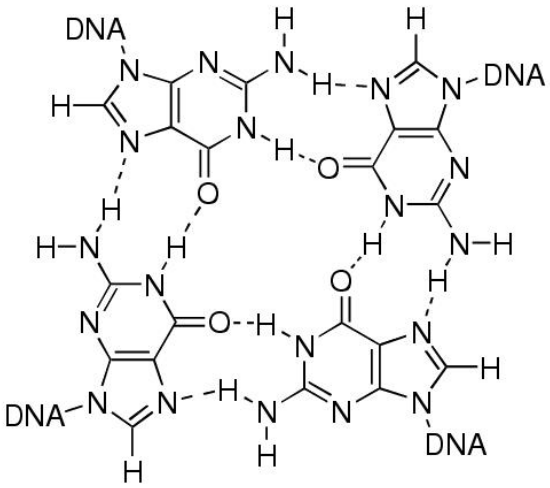

Fig. 1 Four guanines can hydrogen bond in a square arrangement to form a G-quartet. There are two hydrogen bonds on each side of the square.

These planar structures are called G-quartets, and are stabilized by monovalent cations, in particular K^+ and to a lesser extent NH_4^+ and Na^+, which interact with the lone pairs on the O6 atoms surrounding the

central core. They can form discovered in this way in 1910, though the structure was not spontaneously at sufficiently high concentrations of guanine, and indeed were determined until 1962. These G-quartets have large p-surfaces, and hence tend to stack on each other due to π–π stacking, as well as to enable cations to intercalate between the G-quartets. In particular, oligonucleotides with contiguous runs of guanine, such as d(TGGGT) can form stacked structures with the G-quartets linked by the sugar–phosphate backbone. These are called G-quadruplexes and can form from DNA or RNA strands. They are helical in nature due to the constraints of π–π stacking, although for convenience they are often depicted without the helicity, as shown in Fig. 2.

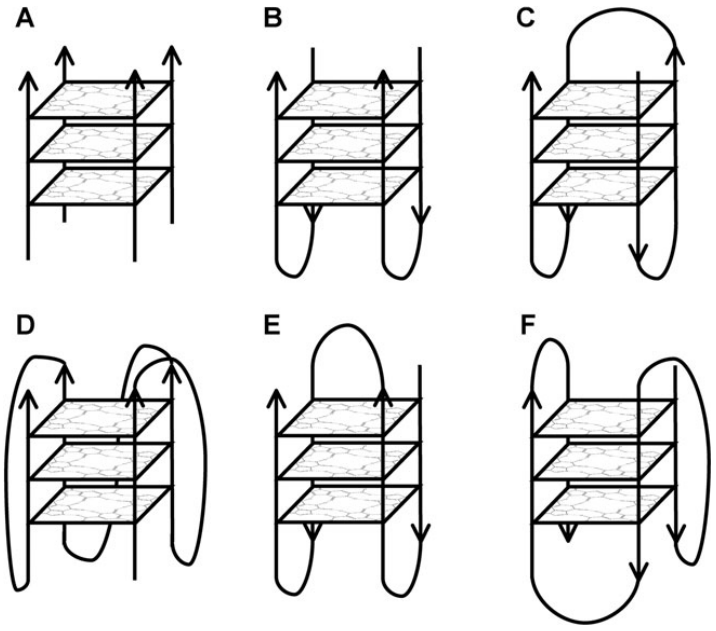

Fig. 2 G-quartets can stack on top of each other, forming G-quadruplex structures. These are held together by π-stacking and the sugar–phosphate backbone.

Since there is a directionality to the strands, customarily described as from the 50 end to the 30 end, there are topological variants possible for these four strands. All four strands may be parallel, three parallel and one

in the opposite direction (antiparallel), or there may be two in one direction and two in the other, either with the parallel pairs adjacent to each other or opposite each other. A shorthand has arisen which describes all the arrangements with at least one antiparallel strand as 'antiparallel', although this does not give a full description of the structures. These are depicted in Fig. 3.

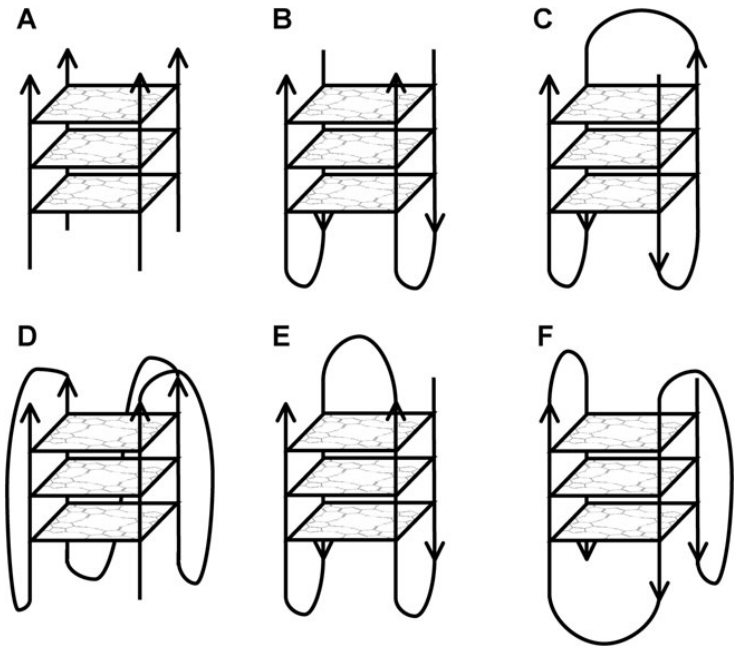

Fig. 3 G-quadruplexes can adopt a range of different stoichiometries and folding patterns. (A) Tetramolecular structure with all strands parallel; (B) bimolecular antiparallel structure with adjacent parallel strands; (C) unimolecular antiparallel structure with alternating parallel strands; (D) unimolecular parallel structure with three double chain reversal loops; (E) unimolecular antiparallel structure with adjacent parallel strands and a diagonal loop; (F) unimolecular mixed structure with three parallel and one antiparallel strands. All three structures (D), (E) and (F) have been observed for the human telomeric repeat.

At a molecular level, the different directionality of the strands relates to the conformational state of the glycosidic bond between the guanine base and the sugar. This may be either syn or anti, as depicted in Fig. 4.

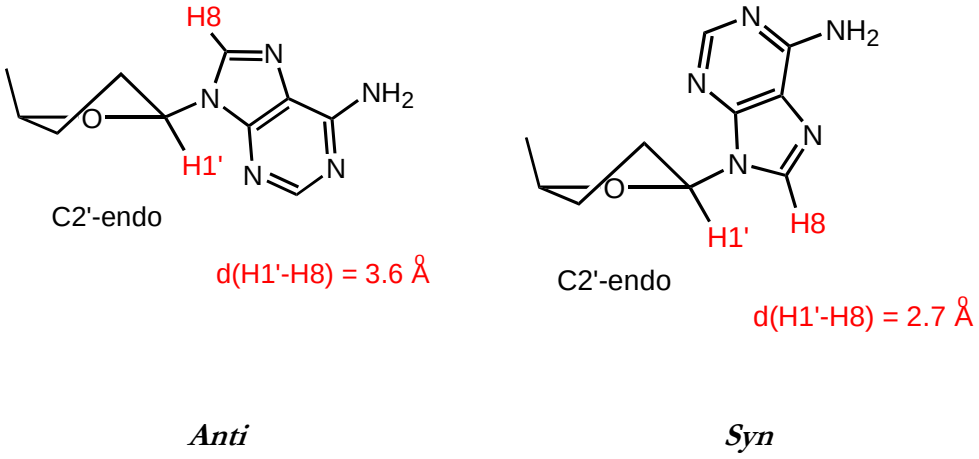

Fig 4 The bond between the base and sugar can rotate. It has two preferred conformations, *syn* and *anti*.

When all four strands are parallel, all the bases are in the anti conformation and the grooves between the backbones are all of equal size—the system is entirely C4 symmetric. When any of the strands are antiparallel, the bases must be in the syn form in order for the hydrogen bonds to be formed correctly. This then affects the orientation of the backbone relative to the G-quartets, and hence results in grooves of different sizes. When successive guanines (starting with the guanine contributing N1 and N2) are both anti or both syn, the groove is medium in size; if the first is anti and the second syn, the groove is wider, and if the first is syn and the second anti, then the groove is narrower. Thus a G-quadruplex with adjacent parallel strands will be arranged with glycosidic bonds anti–syn–syn–anti, and will have grooves that are wide, medium, narrow and medium. In contrast, a structure with alternating strands will have glycosidic bonds anti–syn–anti–syn, with grooves wide, narrow, wide and narrow. G-quadruplexes may be

comprised of four separate strands, as in the example above, forming tetramolecular G-quadruplexes, which are always found in the all-anti parallel form. Alternatively, they may be formed from two strands, each with two sets of contiguous guanines, or just from one strand, folding back on itself to form an intramolecular structure. In either of these cases, there will be loops that serve to connect the strands of the structure together. Depending on which strands are connected, these loops may cross diagonally across the top of the structure, joining diagonally opposed antiparallel strands; go across a side, linking adjacent antiparallel strands; or may loop around the side of the structure linking parallel strands and forming a double-strand reversal loop. Some examples of these are shown in Fig. 3.

2 G-quadruplexes: functions

The important role of these structures in biological systems lies mainly on three features, since they are involved in: i) the architecture of telomeres of many organisms; ii) G-rich sequences that are present within a wide range of genes; iii) the scaffolds of several oligonucleotide aptamers.

2.1 Telomeres and telomerase

A fundamental difference in the behavior of normal versus tumor cells in culture is that normal cells divide for a limited number of times (exhibit cellular senescence) whereas tumor cells usually have the ability to proliferate indefinitely (are immortal). There is substantial experimental evidence that cellular aging is dependent on cell division and that the total cellular lifespan is measured by the number of cell generations, not by chronological time.

Fig 4 Telomere length progressively shortens in somatic cells with successive rounds of cell division, leading eventually to senescence and apoptosis

This means there is an intrinsic molecular counting process occurring during cell growth that culminates in the cessation of cell division.

It is now evident that the progressive loss of the telomeric ends of chromosomes is an important timing mechanism in human cellular aging (Fig 4).

Human telomeres contain long stretches of the repetitive sequence TTAGGG which are bound by specific proteins. With each cell division, telomeres shorten by 50–200 bp, primarily because the lagging strand of DNA synthesis is unable to replicate the extreme 3' end of the chromosome (known as the end replication problem).

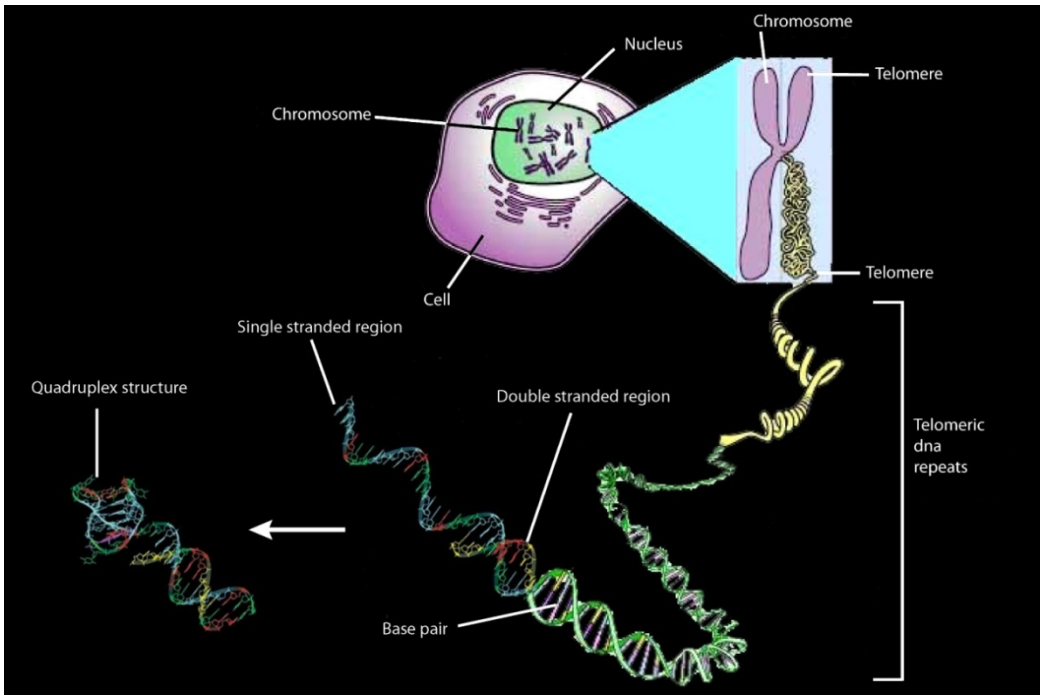

Fig 5 A telomere is the region on ends of chromosomes, composed of repeating units of 6 nucleotides (humans = TTAGGG, repeating 1500 times),without genetic value, capable of forming G-quadruplex structures.

When telomeres become sufficiently short, cells enter an irreversible growth arrest called cellular senescence. In most instances cells become senescent before they can accumulate enough mutations to become cancerous, thus the growth arrest induced by short telomeres may be a potent anti-cancer mechanism.

Telomerase, an eukaryotic ribonucleoprotein (RNP) complex , helps to stabilize telomere length in human stem cells, reproductive cells and cancer cells by adding TTAGGG repeats onto the telomeres using its intrinsic RNA as a template for reverse transcription. Telomerase activity has been found in almost all human tumors but not in adjacent normal cells. The most prominent hypothesis is that maintenance of telomere stability is required for the long-term proliferation of tumors. Thus, escape from cellular senescence and becoming immortal by activating telomerase, or an alternative mechanism to maintain telomeres, constitutes an additional step in oncogenesis that most tumors require for their ongoing proliferation. This makes telomerase a target not only for cancer diagnosis but also for the development of novel anti-cancer therapeutic agents. One approach to telomerase inhibition involves sequestering its substrate, single-stranded telomeric DNA, by inducing it to form G-quadruplex structures. There has been considerable work on developing G-quadruplex ligands, especially to target the human telomeric repeat and hence block the action of telomerase.

2.2 Transcription regulation

Gene transcription is tightly regulated, by a variety of methods. One method that is used in some cases is based on the presence of G-quadruplexes located in the promoter region of a gene, broadly speaking the kilobase upstream of the transcription start site (TSS). This model

was originally demonstrated by Hurley and co-workers for the oncogene c-myc, an important transcription factor involved in regulating around 15% of all human genes. As a result of this, overexpression of c-myc has been implicated in a wide range of cancers including colorectal cancer. Within its promoter there is a region, 115–142 basepairs upstream of the TSS, which is highly sensitive to nucleases, suggesting that it forms an accessible structure free from histone proteins. This region controls the vast majority of the transcription of the gene, and studies in vitro of the sequence d(GGGGAGGGTGGGGAGGGTGGGGAAGG) show that it is capable of forming into a family of polymorphic G-quadruplexes, using various combinations of the guanine runs underlined. It has further been shown that the G-quadruplex ligand TMPyP4 (see below) binds to this element leading to downregulation of c-myc expression. This clear proof of principle led to the proposal that this may be a general mechanism for gene regulation. The simplest form of the model (Fig. 6) proposes that there is an equilibrium between two forms of the DNA. On one side of the equilibrium is double helix DNA, and transcription occurs as normal; on the other side, one strand is separated, and has folded up into a G-quadruplex. This structure then acts as a steric block to transcription. Addition of a G-quadruplex ligand, whether a small molecule or a protein, will energetically favour the G-quadruplex form, and hence move the equilibrium towards that side and reduce the transcriptional activity. This dynamic equilibrium has been experimentally demonstrated. Although this model is presented in terms of steric blockage leading to a reduction in transcriptional activity, as was found for c-myc, it is also possible that the G-quadruplex form could be an activating domain, either because of putative protein recognition of

the G-quadruplex, or if the accessibility of the other strand leads to increased transcriptional activity.

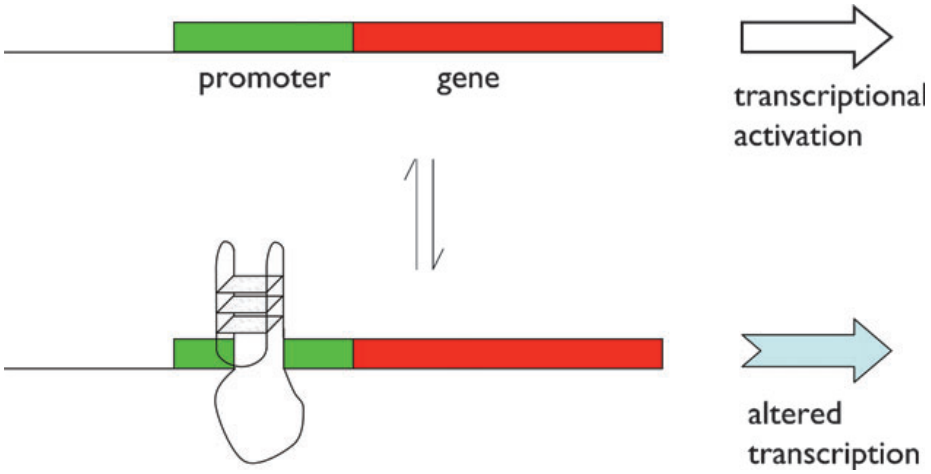

Fig. 6 The formation of a G-quadruplex in a promoter can affect the level and nature of transcription from that gene. At the simplest level, it may act as a steric block to the transcription machinery.

However, although a wide variety of different genes have now been shown to have promoter G-quadruplexes, such as VEGF, HIF-1a, Bcl-2, Ret, c-kit and KRAS, none have yet had G-quadruplex formation leading to increased transcriptional activity. Although in Fig. 6 the complementary C-rich strand is drawn as an unstructured sequence, it is possible that it could form an alternative four-stranded structure called an i-motif (Fig. 7). An i-motif has four strands forming a structure somewhat like two interleaved ladders, with pairs of strands held together by diagonal C*R*C+ bonds. These rely on the protonation of the N3 of cytosine, which has a pKa of 4.2. As a result, these structures are generally only stable under acidic conditions, but the stability will clearly be different in the context of chromosomal DNA, especially if a G-

quadruplex structure forms and holds the ends of the i-motif together in the correct orientation.

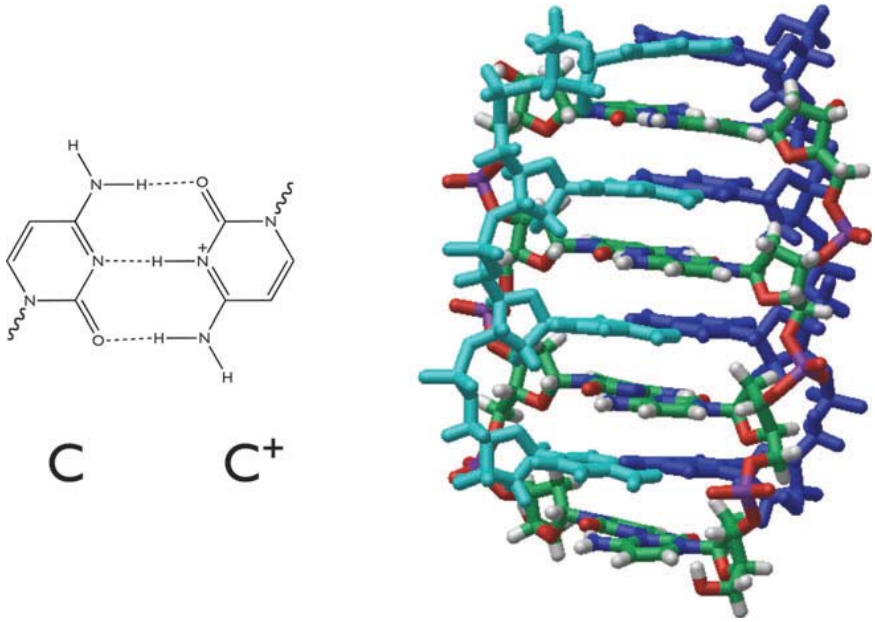

Fig. 7 Left: two cytosines can pair up, forming three hydrogen bonds, when one of them has been protonated. Right: four cytosinecontaining strands can form an i-motif, with interleaved C:C+ bonds. One pair of binding strands is shown in CPK colouring, the other in blue/cyan.

Nonetheless, there is still controversy as to the biological relevance of the i-motif structure. Using the predictive algorithm quadparser, it is possible to investigate how many human genes contain G-quadruplex motifs in their promoter regions, both to identify novel genes for experimental testing, but also to see if there is evidence of over- or under-representation of such genes. This analysis has shown that almost half of all genes (43%) contain putative G-quadruplex structures in their promoters, which is considerably more than would be expected by chance, based on the rest of the genome. The enrichment of G-quadruplex motifs occurs increasingly nearer the TSS, with the first few hundred bases seeming to be particularly important for G-quadruplex motifs to be present. Interestingly, genes involved in cancer are even

more likely to have such promoter G-quadruplexes, with 67% having such sequences. Using the online Gene Ontology database, which describes the functions of every human gene, it is possible to see if there is any general bias for types of genes that have promoter G-quadruplexes. This has revealed that genes involved in tightly-regulated processes such as development, neurogenesis and cell differentiation are more likely than other classes of genes to have promoter G-quadruplexes, whereas genes involved in processes such as protein biosynthesis, olfaction and immune response are much less likely to have promoter G-quadruplexes.[1]

[1] See Julian Leon Huppert, Chem. Soc. Rev., 2008, **37**, 1375–1384 and references cited therein.

2.3 Aptamers

Aptamers[1] are nucleic acid macromolecules that bind to molecular targets, including proteins, with high affinity and specificity. Aptamers are typically from 15 to 40 nucleotides in length and can be composed of DNA and RNA. Base composition defines aptamer secondary structure, consisting primarily of helical arms and single-stranded loops. Stable tertiary structure, resulting from combinations of these secondary structures, allows aptamers to bind to targets via van der Waals, hydrogen bonding and electrostatic interaction.

The conceptual framework and process of aptamer generation emerged from pioneering experiments published in the early 90s.[2-3] Particularly, Gold *et al.*[3] described a process of *in vitro* aptamer selection, dubbed SELEX (Systematic Evolution of Ligand by Exponential enrichment).

SELEX is a combinatorial chemistry methodology based on oligonucleotide libraries which are screened for high affinity binding to a given target. High affinity ligands can be isolated from the library using iterative rounds affinity-based enrichment, alternating with oligonucleotide amplification.

The aptamer drug TBA (Thrombin Binding Aptamer), also named ARC183 (Archemix Corp) is a consensus DNA 15-mer, namely $^{5'}$GGTTGGTGTGGTTGG$^{3'}$,[4,5] discovered with this technique (Fig. 8).

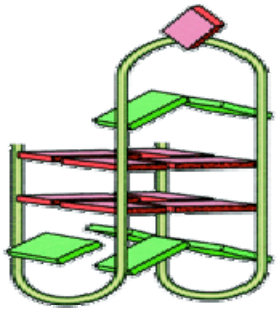

Fig.8 Pictorial moldel of TBA

The three dimensional solution structure of TBA was solved using NMR and X-ray techniques (Fig. 9).[6,7] TBA is characterized by a chair-like quadruplex structure consisting of two G-tetrads connected by two TT loops and a single TGT loop. The aptamer is a thrombin inhibitor in development for use as an anticoagulant during coronary artery bypass graft procedures. Currently, the only approved anticoagulant for coronary artery bypass graft is heparin. TBA exhibits a K_d of 2 nM for thrombin, 50 nM for prothrombin, and binding to other serum proteins or proteolytic enzymes is essentially undetectable.[2] As suggests by others,[8] TBA binds at the anion exosite I of thrombin in a conformation little changed from that of the unbounded species. TBA is a strong anticoagulant *in vitro*, and inhibits thrombin-catalyzed activation of

fibrinogen, and thrombin induced platelet aggregation. TBA has key advantages in that it avoids heparin use and the risk of associated thrombocytopenia, is a specific inhibitor with rapid onset, is effective at inhibiting clot-bound thrombin, and has a short *in vivo* half-life of approximately 2 min which allows for rapid reversal of its effects and the avoidance of dose-adjusting complications of heparin and protamine. Neither significant toxicities nor excessive bleeding intraoperatively has been observed.

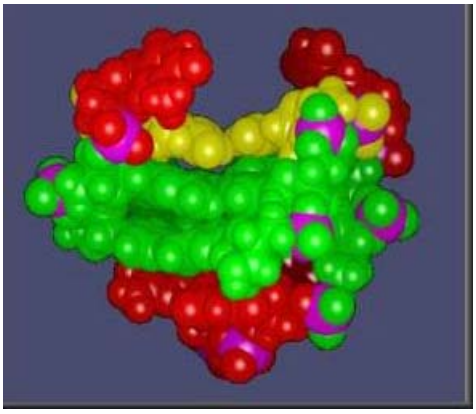

Fig. 9 Three dimensional solution structure of TBA.

Notes and references

1 Nimjee, S. M.; Rusconi, C. P.; Sullenger, B. A. *Annual Review of Medicine,* 2005, **56**, 555.

2 Block, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J. *Nature,* 1992, **355**, 564.

3 Gold, L.; Polisky, B.; Uhlenbeck O.; Yarsus, M. *Annu. Rev. Biochem.*, 1995, **64**, 763.

4 Griffin, L. C.; Tidmarsh, G. F.; Bock, L. C.; Toole, J. J.; Leung, L. K. *Bloods*, 1993, **81**, 3271.

5 Li, W. X.; Kaplan, A. V.; Grant, G. W.; Toole, J. J.; Leung, L. L., *Blood*, 1994, **83**, 677.

6 Wang, K. Y., McCurdy, S.; Shea, R. G.; Swaminathan, S.; Bolton, P. H. Biochemistry, 1993, **32**, 1899.

7 Macaya, R. F.; Schultze, P.; Smith, F. W.; Roe, J. A.; Feigon, J. Proc. Natl. Acad. Sci. USA, 1993, **90**, 3745-3749.

8 Paborsky, L. R.; McCurdy, S. N.; Griffin, L. C.; Toole J. J.; Leung, L. K. J. Biol. Chem., 1993, **268**, 20808.

3 Effect of rubidium and cesium ions on the dimeric quadruplex formed by the *Oxytricha nova* telomeric repaeat oligonucleotide d(GGGGTTTTGGGG)

In order to improve the knowledge of the G-quadruplex structures, and in particular to provide further insight into the role of monovalent cations in the folding and conformation of this kind of molecule, we have investigated by ^{1}H-NMR the effect of the addition of Rb^{+} and Cs^{+} ions to the quadruplex formed by oligonucleotide d(GGGGTTTTGGGG).

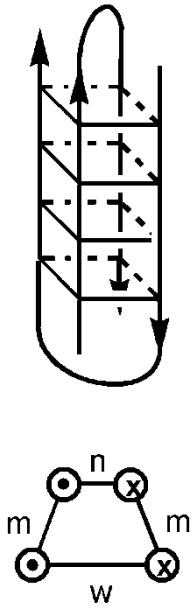

Fig. 1 Schematic illustration of [d(GGGGTTTTGGGG)]$_2$ quadruplex structure. Strand direction along each edge of the quadruplex is indicated by arrows. Top view is reported below. Types of grooves are indicated as narrow (n), medium (m) and wide (w). Strand polarity is indicated by the arrow tails (crosses) and heads (points) in circles.

The solution structure of this oligonucleotide has been determined by an in-depth NMR study.[1,2] It possesses four G-tetrads with guanosines that adopt a syn-syn-anti-anti conformations, which result in a quadruplex of alternative wide, medium, narrow, medium width grooves between strands, and possessing two TTTT loops across the diagonal at the edge of the structure (Fig. 1). Here we report the ^{1}H-NMR titration experiments of the quadruplex $[d(GGGGTTTTGGGG)]_2$ originally containing 50 mM Na^+. The titration profiles with RbCl and CsCl turned to be completely different (Fig. 2).

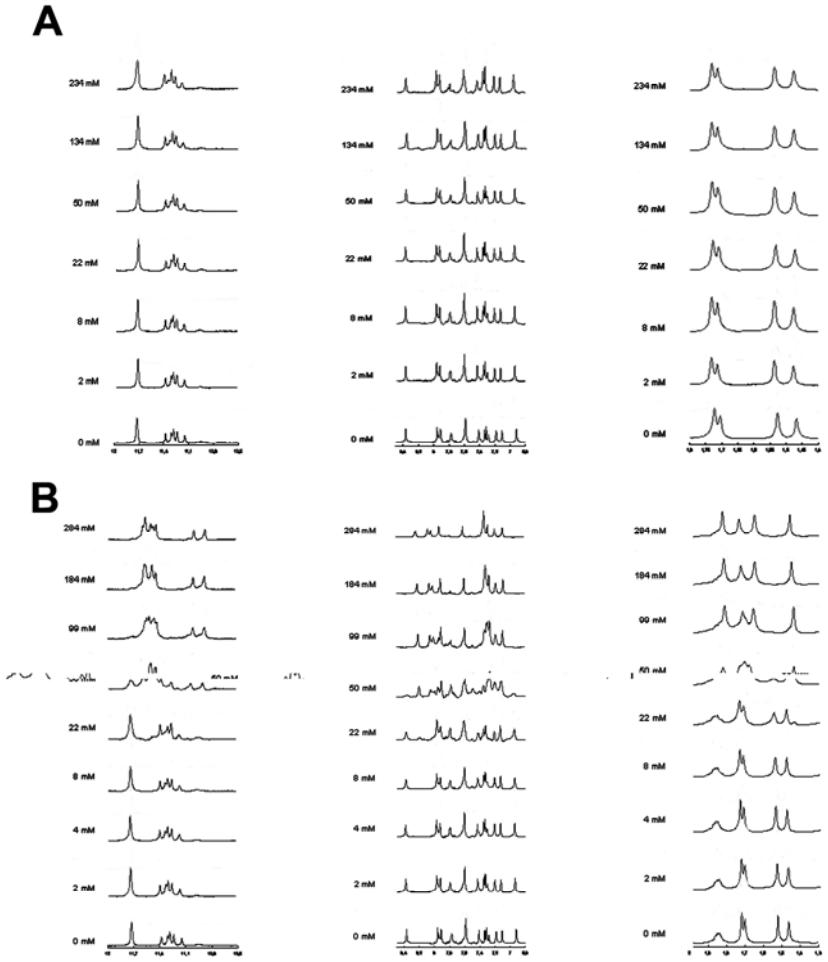

Fig. 2 Titration (T=300 K, 500 MHz) with CsCl (A) and RbCl (B) of 1.5 mM solution of $[d(GGGGTTTTGGGG)]_2$. The CsCl and RbCl millimolar concentrations are shown along the side of spectra.

The addition of CsCl caused no changes in the chemical shift of the oligonucleotide (Fig. 2A). The addition of RbCl (Fig. 2B) up to a total concentration of 15 mM did not cause any significant proton spectra change. Surprisingly, at concentration of RbCl above 15 mM, a new set of proton signals could be observed, whose intensities rose by increasing the amount of RbCl, with the concomitant falling off of the original signals which completely disappeared at RbCl concentration of 100mM. This could be interpreted assuming that the resonances of the quadruplex gradually moved from their Na^+ to their Rb^+ form chemical shift, and that, above 20mM, Rb^+ ions are more affine to the quadruplex formed by the *Oxytricha nova* telomeric repeat than the Na^+ ions.

These preliminary results could shed more light to understand the influence of the monovalent cations on the $[d(GGGGT\bar{T}\bar{T}TGGGG)]_2$ quadruplex structure and stability. The different affinities of different cations can be used , in principle, to modulate the structures of aptamers based on the scaffold possessed by $[d(GGGGT\bar{T}\bar{T}TGGGG)]_2$, and /or for the development of new structural motif to be used in the natotechnology field.

An NMR study devoted to study the structural changes due to the binding of Rb^+ is currently in progress in our laboratories.

Notes and references

1. Smith, F. W., Feigon, J. 1992. *Nature*, **356**, 164-168.
2. Smith, F. W., and J. Feigon. 1993. *Biochemistry*. **32**,8682–8692.

4 Backbone modifications

Biological relevance of DNA G-quadruplex structures mainly lies on two features: their probable presence in many important regions of genome and their fundamental role as scaffold in several aptamers provided with useful biological properties.[1] In fact, the interest in G-quadruplexes is witnessed by hundreds of reports dealing with various aspects of these DNA secondary structures, such as the chemical nature and the topology of the backbone and the loops (where present), the conformation of the glycosydic linkage, as well as the influence of structural modifications on the stability of the complexes.[1] Several studies dealing with modified sugar-phosphate backbone containing G-quadruplexes have been described, as well.[2,3] However, at the best of our knowledge, no reports describing G-quadruplexes containing 3'-3' or 5'-5' inversion of polarity included in the guanine tract have been reported yet, even though these backbone modifications are well-know to increase nuclease resistance in therapeutic oligodeoxyribonucleotides (ODNs) such as antisense ODNs[4] and aptamers.[5]

The structural variability of the quadruplexes has been further increased by introducing modification on both bases and sugar-phosphate backbone.

4.1 A New Class of DNA Quadruplexes Formed by Oligodeoxyribonucleotides Containing a 3'-3' or 5'-5' Inversion of Polarity Site

Unprecedented DNA quadruplex structures containing a 3'-3' or 5'-5' inversion of polarity site in the G-tract are presented. The quadruplexes are caracterised by different elements of symmetry and glycosidic angle conformations.

In this frame, we wish to report here the NMR and CD studies of the two quadruplexes named **Q33** and **Q55,** formed by the oligodeoxynucleotides (ODNs) $^{5'}$TGG$^{3'}$-$^{3'}$GGT$^{5'}$ and $^{3'}$TGG$^{5'}$-$^{5'}$GGT$^{3'}$, respectively, in which four strands are characterized by the presence of a 3'-3' or 5'-5' inversion of polarity site within the G-stretches.

ODNs $^{5'}$TGG$^{3'}$-$^{3'}$GGT$^{5'}$ and $^{3'}$TGG$^{5'}$-$^{5'}$GGT$^{3'}$ were synthesized by standard methods for the 3'-5' tracts, and using 5'-phosphoramidites for the 5'-3' tracts. For the former, a CPG resin linked to a thymine residue through the 5'OH function was also employed. The NMR samples of **Q33** and **Q55** were prepared at a concentration of 2.0 mM (0.6 ml, 90% H_2O / 10% D_2O) and they were studied in two different buffers: 10 mM potassium phosphate, 70 mM KCl, 0.2 mM EDTA (pH 7.0) and 10 mM sodium phosphate, 70 mM NaCl, 0.2 mM EDTA (pH 7.0). The 1D proton spectra of both samples in the two buffers are almost superimposable.

Apart from some weak resonances due to very minor conformations also present in solution (whose relative intensities turned out to be insensitive to temperature changes), the one- dimensional proton spectrum of **Q33** consists of well defined signals, suggesting that, in the conditions used here, the modified ODN mainly forms a single well-defined hydrogen-

bonded conformation. This was confidently identified as a quadruplex structure as reported below. Particularly, two sharp resonances in the region corresponding to imino protons involved in Hoogsteen hydrogen bonds are present, along with three base proton and one methyl signals at higher fields. A comparison between the ^{1}H-NMR spectra of **Q33** and that of the unmodified $[d(TGGGGT)]_4$,[1] led us to conclude that the species under investigation is a quadruplex structure, as well. Furthermore, the halved number of resonances in the spectrum of **Q33** clearly indicates it to possess further elements of symmetry, in addition to the fourfold symmetry.

The NOE pattern of **Q33** turned out to be similar to that observed for other parallel quadruplex structures.[1,2] As a matter of fact, an unbroken path of NOE connectivities between H8/H6 protons and the deoxyribose protons of the adjacent nucleoside at the 5' side is clearly observable, which is typical of right-handed helix structures.[1,2] Furthermore, the lack of strong NOEs between any G H8 and H1' of the same residue, in comparison to those observed between each G H8 and its deoxyribose H2'/H2", suggests that all G residues are in the *anti* glycosidic conformation.[1,2] It is interesting to note that, since all residues assume an *anti* conformation, all bases possess the same orientation with respect of the strand subunit orientation. Thus, the two tetrads comprising the inversion of polarity site will be characterized by an opposite (clockwise and anti-clockwise) disposal, leading to a less efficient (head-to-head) stacking between couple of Gs in each strand (Figure 1, model **II**).

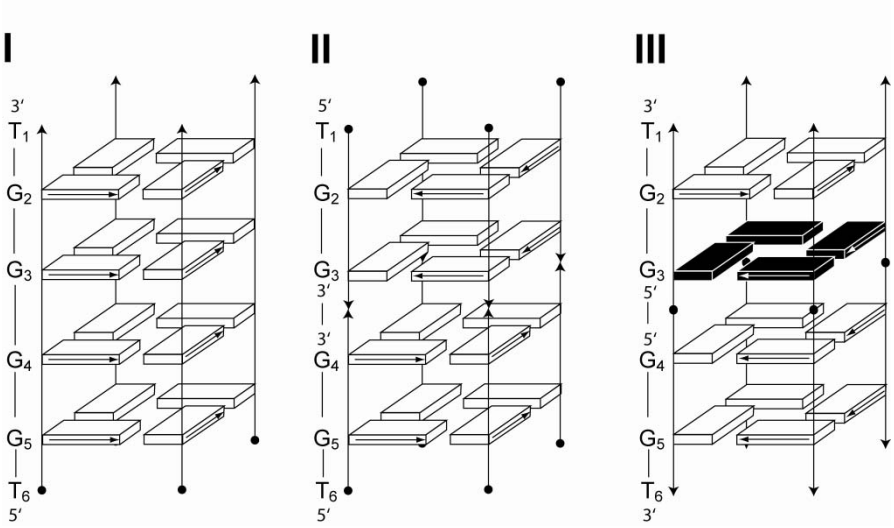

Fig. 1 Schematic illustration of the structure of $[d(TGGGGT)]_4$ (**I**), **Q33** (**II**) and **Q55** (**III**). Black arrowheads and circles indicate 3' and 5' edges of each subunit/strand, respectively. *Anti* and *syn* guanines are depicted as white and black solids, respectively. Arrows on solids indicate the direction of the proton donors and acceptors in Hoogsteen hydrogen bonds.

As for **Q55,** its ^{1}H-NMR spectrum (see supporting information) clearly indicates the presence of a unique species in solution, most likely a quadruplex structure, as suggested by a set of four resonances in the region of imino protons involved in Hoogsteen hydrogen bonds. Surprisingly, the number of signals in the spectrum of **Q55** matches that of $[d(TGGGGT)]_4$, rather than that of **Q33**. This datum points to a different degree of symmetry for the two novel complexes.

Useful information to elucidate the structure of **Q55** arose from ^{31}P-NMR spectroscopy. First, the one-dimensional proton decoupled phosphorus spectrum displays five signals. After assigning the ^{1}H resonances within each deoxyribose by a 2D TOCSY experiment, the 2D proton-detected heteronuclear ^{1}H-^{31}P COSY allowed us to correlate each phosphorus resonance to the pertinent protons. Following this procedure, we unambiguously assigned the protons of six deoxyribose

rings and five phosphorous. These data led us to hypothesize that, differently from that observed for **Q33**, no further elements of symmetry are present, whereas the four strands are equivalent to each other (fourfold symmetry). Confirmative evidences were gained from an extensive analysis of NOESY spectra of **Q55**, as described below.

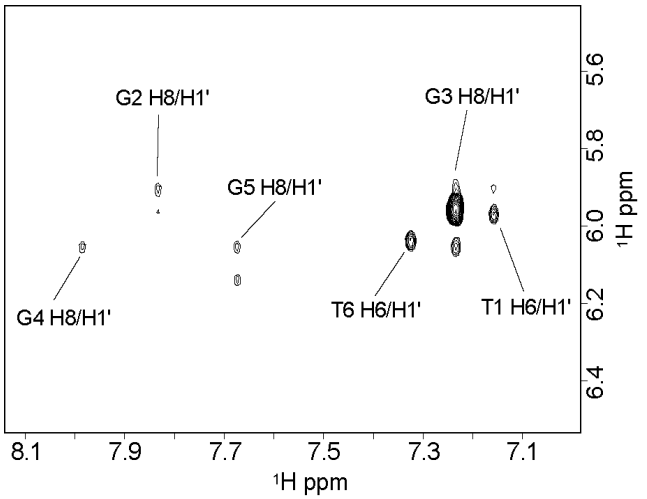

Fig. 2 2D NOESY (700 MHz, 100 ms mixing time, T=5°C) contour plots correlating base and sugar H1' protons in **Q55**.

Interestingly, only one G nucleotide shows intense H8-H1' NOE (Figure 2), diagnostic of a *syn* glycosidic conformation. In agreement with literature data,[3] the H8 resonance of this *syn* G residue is upfield shifted with respect to those of the *anti* ones. This nucleotide shows a single sequential connectivity to another G residue which, in turn, is sequentially connected to a T residue. Furthermore, sequential connectivities between H8/H6 protons and the deoxyribose protons of the adjacent nucleoside at the 5' side could be clearly observed for the remaining portion of the strand, comprised of two more G and a T residues.

The whole of the above data could be plainly interpreted assuming that the *syn* G residue is adjacent to the inversion of polarity site (Figure 1, model **III**), and that the path of NOE connectivities is broken at this level. On the other hand, the existence of NOE sequential connectivities along the two subunits of the strand suggests that the backbone of **Q55** adopts a right-handed helix conformation.[1,2]

In the NOESY spectrum of **Q55** recorded in H_2O, we observed sequential imino-imino NOEs arising from intra-strand contacts, along with the connectivities between imino protons of G2 and G5 nucleotides with the adjacent methyls of T1 and T6 residues, respectively. Furthermore, NOE connectivities G3-H8/G3-NH, G3-NH_2 and G4-H8/G4-NH, G4-NH_2, also present in the spectrum, were ascribed to contacts between protons belonging to distinct, although equivalent, bases within each tetrad, thus further confirming the C4 symmetry of the complex.

CD spectra for **Q33** and **Q55** and their natural counterpart $[d(TGGGGT)]_4$ (Figure 1, model **I**) were acquired at 20°C and reported as supporting information. All the measurements were performed at a concentration of 1 x 10^{-4} M, with both buffers used for NMR experiments. The CD spectrum of $[d(TGGGGT)]_4$ is characterized by a positive and a negative bands at 243 and 263 nm, respectively. The CD spectra of **Q33**, instead, is characterized by a profile never observed before for quadruplex structures, having two positive bands at 253 and 294 and two negative bands at 232 and 270 nm, respectively. On the other hand, **Q55** is characterized by a CD profile very similar to those of folded antiparallel quadruplexes involving deoxyguanosine alternating between *syn* and *anti* conformations about the glycosidic bond[4] (two maximum bands at 250 and 295 nm and a minimum band at 270 nm).

CD profiles of **Q33** and **Q55** do not fit into literature data regarding parallel and antiparallel quadruplexes so far reported. This is not quite surprising, considering the unprecedented structural features present in the novel complexes. Nevertheless, in order to elucidate the relationship between structure and CD profile, further studies on similar molecules are in order.

In order to estimate their thermal stability, **Q33** and **Q55** were subjected to melting and annealing CD experiments in comparison with $[d(TGGGGT)]_4$, under the same experimental conditions. All the measurements were performed at a concentration of 1 x 10^{-4} M using both Na^+ and K^+ buffers. Taking into account that the rates of quadruplex formation/dissociation are very slow, we collected the data at 10°C/h. **Q55** and $[d(TGGGGT)]_4$ were still structured even at 95°C using a K^+ buffer, while we were able to obtain good melting profiles using a Na^+ buffer for all three samples. In spite of the very slow scan rate used, an extreme hysteresis phenomenon was observed (no sigmoidal annealing profile was obtained), thus indicating that the systems were not at equilibrium. Therefore, from the melting curves, the apparent melting temperature of 65°C, 52°C and 90°C could be measured for $[d(TGGGGT)]_4$, **Q33** and **Q55**, respectively. A considerable increase (25°C) in the apparent melting temperature could be determined for **Q55** when compared with the natural counterpart, while **Q33** was the least stable complex.

The above finding concerning the relative stabilities of the three complexes is intriguing and certainly deserves further studies in depth. At the moment, we can only hypothesize that, as far as **Q33** and **Q55** are concerned, the origin of the different behavior is to be sought for in the higher flexibility of the inversion of polarity site in **Q55**, where one of

the two tetrads comprising the inversion site of polarity is allowed to arrange all residues in the *syn* glycosidic conformation. Thus, having the two subunit of each strands opposite polarity, the two tetrads have the same orientation, similar to a canonical parallel quadruplex[1,2] (Figure 1, model **I** and **III**), allowing a better stacking between the residues of the two G-quartets. The energy gain deriving from the more efficient stacking in the central, more structured part of the quadruplex exceeds the loss of stability due to the adoption of the less favorable *syn* glycosidic conformation for the G3 residues and the less efficient stacking between G2 and G3 tetrads at the edge of the complex. More importantly, the higher structural flexibility**,** along with the unprecedented arrangement of the two tetrads in the middle of **Q55,** might also favor a more efficient complexion with Na^+, thus accounting for the increased T_m value towards $[d(TGGGGT)]_4$. Molecular mechanics and dynamics calculations are currently in progress to verify these hypotheses.

Notes and references

1. F. Aboul-ela, A. I. H. Murchie, D. G. Norman, D. M. J. Lilley, *J. Mol. Biol.* 1994, **243**, 458.
2. (a)P. K. Patel, A. S. R. Koti, R. V. Hosur, *Nucleic Acids Res.* 1999, **27**, 3836; (b) C. Cheong, P. B. Moore, *Biochemistry* 1992, **31**, 8406; (c) R. Jin, B. L. Gaffney, C. Wang, R. A. Jones, K. J. Breslauer, *Proc. Natl. Acad. Sci. USA* 1992, **89**, 8832; (d) Y. Wang, D. J. Patel, *J. Mol. Biol.* 1993, **234**, 1171; (e) P. K. Patel, R. V. Hosur, *Nucleic Acids Res.* 1999, **27**, 2457.
3. (a)F. W. Smith, J. Feigon, *Biochemistry* 1993, **32**, 8683; (b) K. Y. Wang, S. McCurdy, R. G. Shea, S. Swaminathan, P. H. Bolton,

Biochemistry 1993, **32**, 1899; (c) Y. Wang, D. J. Patel, *Structure* 1993, **1**, 263.

4. (a)M. Lu, Q. Guo, N. R. Kallenbach, *Biochemistry* 1993, **32**, 598; (b) Q. Guo, M. Lu, N. R. Kallanbach, *Biochemistry* 1993, **32**, 3596; (c) I. Smirnov, R. H. Shafer, *Biochemistry* 2000, **39**, 1462.

4.2 Molecular modelling studies of four stranded quadruplexes containing a 3'-3' or 5'-5' inversion of polarity site in the G-tract

In an attempt to get the solution structures at the atomic level of **Q33** and **Q55**, restrained molecular mechanic and dynamic calculations have been performed. We firstly studied the structure adopted by **Q33**. Thus, a quantitative estimation of proton-proton distances has been done measuring the cross-peak intensities in 2D NOESY experiments acquired at 700 MHz, both at T=25°C and T=5°C (mixing time 100 ms for the experiment acquired in D_2O and 200 ms for the experiment acquired in H_2O). 656 distance restraints (164 for each strand) have been retrieved. Moreover, as suggested by the presence of G imino protons in the 1D ^{1}H-NMR spectrum, 64 supplementary distance restraints (HN1-O6, N1-O6, HN2-N7, N2-N7) for 16 hydrogen bonds (corresponding to the four G-quartets), were also incorporated during the computations. Further, in agreement with NMR data, glycosidic torsion angles were fixed in the *anti*-domain (the χ angle was kept in a range of -160°/-70°). Therefore, three-dimensional structures which satisfy NOEs were constructed by simulated annealing (SA) calculations. An initial structure of the oligonucleotide was constructed possessing a random conformation and minimized, in order to eliminate any possible source of initial bias in the folding pathway. Restrained simulations were carried out for 500 ps using the CVFF force field as implemented in Discover software (Accelrys, San Diego, USA). The restrained SA calculations started at 1000 K, and, thereafter, the temperature was decreased stepwise down to 273 K. The aim step was to energy-minimize and

refine the structures obtained by using the steepest descent followed by the quasi-Newton-Raphson (VA09A) algorithms. A total of 10 structures have been generated. To better characterize the local conformation space of the molecules, each structure was then reconsidered by adding K^+ counterions (there were no constrains for K^+ counterions), and soaking the molecule in explicit H_2O. Thus, each structure was heated to 500 K and kept there for 50 ps, and then cooled to 300 K and the system was maintained there for other 50 ps. All 10 structures obtained were subsequently energy minimized using a combination of steepest descent and quasi-Newton-Raphson (VA09A) algorithms. It was possible to obtain a good superimposition of the 10 structures with RMSD values of 0.48 ± 0.16 for heavy atoms.

Concerning **Q55**, it was subject to the same procedure described for **Q33**. In this case, 520 distances (130 for each strand) have been used in the calculations, and, as suggested by the presence of G imino protons in the 1D ^{1}H-NMR spectrum, 64 supplementary distance restraints (HN1-O6, N1-O6, HN2-N7, N2-N7) for 16 hydrogen bonds (corresponding to the four G-quartets), were also incorporated during the computations. It is interesting to note that the χ angles of twelve out of sixteen guanines involved in the formation of the four G-tetrads were fixed in the *anti*-domain (-160°/-70°), whereas a range of 10°/100° (*syn*-conformation) was used for the remaining four G residues. At the end of the calculations, also for Q55, it was possible to observe a good superimposition of the 10 structures generated, with RMSD values of 0.92 ± 0.32 for heavy atoms.

The obtained structures reveal that **Q55** and **Q33** possess a different stacking (Fig. 1) and different twist angle (and therefore different helical winding) at the inversion of polarity level (Fig. 2).

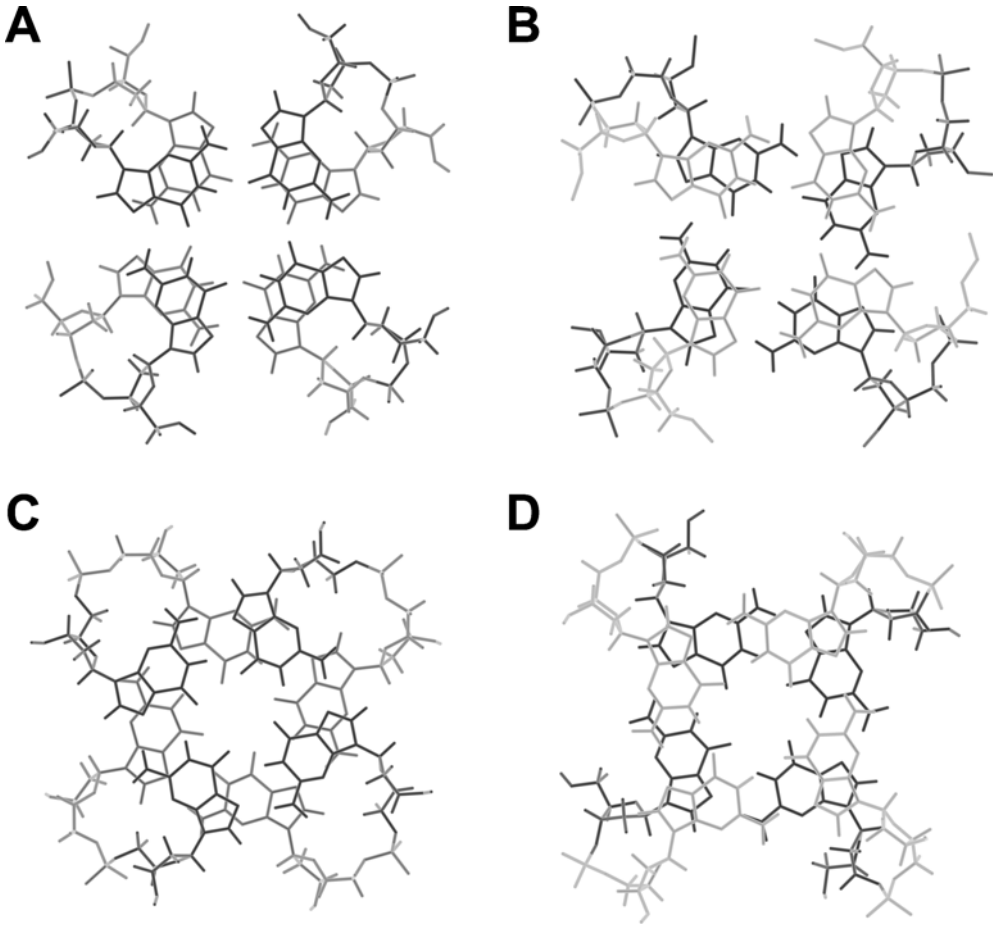

Fig.1 Top view of the stacking observed between the G-quartets (G3- and G4-quartet) comprising the inversion of polarity site of **Q33** (A) and **Q55** (C), and stacking observed between the G2- and G3- quartet of **Q33** (B) and **Q55** (D).

Surprisingly Q33 is characterized by a better stacking between the two G-quartets adjacent the inversion of polarity site. Thus, the higher thermal stability of Q55 could be ascribed to a more efficient stacking between G2- and G3-tetrads. However, an entropic contribution, closely connected to the higher flexibility of the inversion of polarity site of Q55, may also be responsible of the different melting behavior. Another possible explanation could be ascribed to the different bases orientation in the two complexes, that, in principle, could lead to a different efficiency in the coordination of K^+ ions. For this reason an in-depth thermodynamic study will be performed.

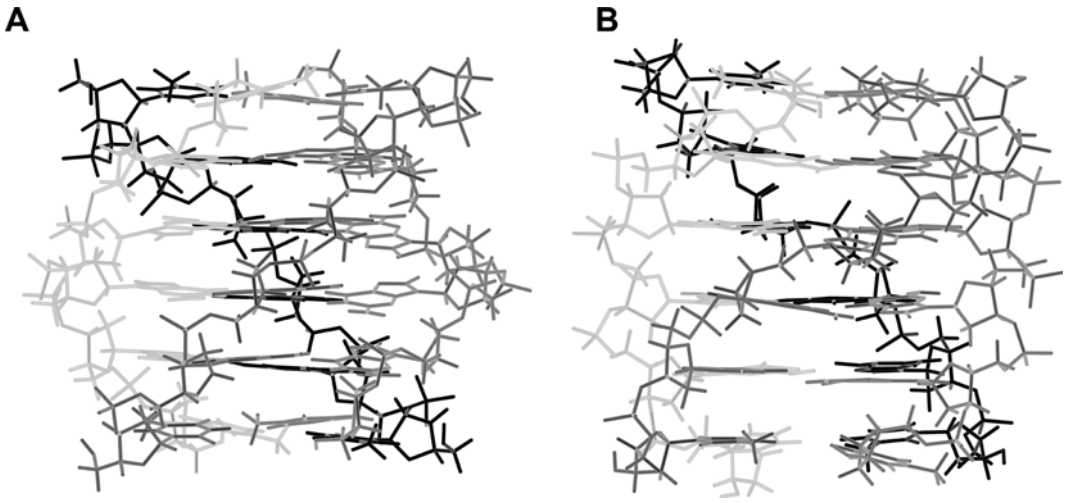

Fig. 2 Side view of average structures of the best 10 structures of **Q33** (A) and **Q55** (B) quadruplexes.

4.3 A further contribution to the extreme variability of quadruplex structures from oligodeoxyribonucleotides containing inversion of polarity sites in the G-tract

In order to get further insights about the aptitude of the inversion of polarity sites to affect the nucleotide glycosidic conformation of adjacent G-residues into quadruplex structures and their influence on the typology and stability of the resulting complexes, we have investigated the behaviour in solution of oligonucleotides $^{5'}TG^{3'}$-$^{3'}GGGT^{5'}$ (**QS33**) and $^{3'}TG^{5'}$-$^{5'}GGGT^{3'}$ (**QS55**) in which, compared to **Q33** and **Q55**, the inversion of polarity sites are one-residue shifted towards the oligonucleotide ends.

ODNs **QS33**, **QS55** and their natural counterpart d(TGGGGT) were synthesized by standard methods for the 3'-5' tracts and using 5' phosphoramidites for the 5'-3' tracts. For the former, a CPG resin linked to the thymine residue through the 5'-OH function was also employed. Main structural features of **QS33** and **QS55** were elucidated by NMR.

NMR samples were prepared at a concentration of approximatively 2 mM, in 0.6 ml (H_2O/D_2O 9:1) buffer solution having 10 mM KH_2PO_4, 70 mM KCl, 0.2 mM EDTA, pH 7.0. The simple appearance of the one-dimensional NMR spectrum of **QS33** provides strong evidence in favour of the formation of a stable quadruplex structure, showing mainly a single well-defined hydrogen-bonded conformation. In fact, the ^{1}H-NMR spectrum of **QS33** (700 MHz, T = 25°C), clearly displays the presence of four well defined signals in the region 10.5-12 ppm, attributable to imino protons involved in Hoogsteen hydrogen bonds of

four G-quartets and no additional peaks due to the potential presence of a minor conformation could be detected. Furthermore, the spectrum of **QS33** shows the same number of resonances of the natural counterpart. This datum clearly indicates that **QS33** is characterized by having four strands equivalent to each other, therefore possessing a fourfold symmetry (C4). In addition, the NOE pattern of **QS33** turned out to be similar to that of other parallel quadruplex structures. Moreover, the entire path of NOE connectivities observed for all Gs indicates that the backbone conformations resemble that of the unmodified $[d(TGGGGT)]_4$ possessing a right-handed helix structure. The lack of strong NOEs between any G H8 and H1' of the same residue, in comparison to those observed between each G H8 and its ribose H2'/H2", indicates that all G residues are in the *anti* glycosidic conformation. Only one structure for **QS33** is consistent with results from NMR investigation and the data present in literature about the relationship between the strand orientation and the glycosidic conformation of G residues[1], and it is reported in Fig. 1 (structure **I**).

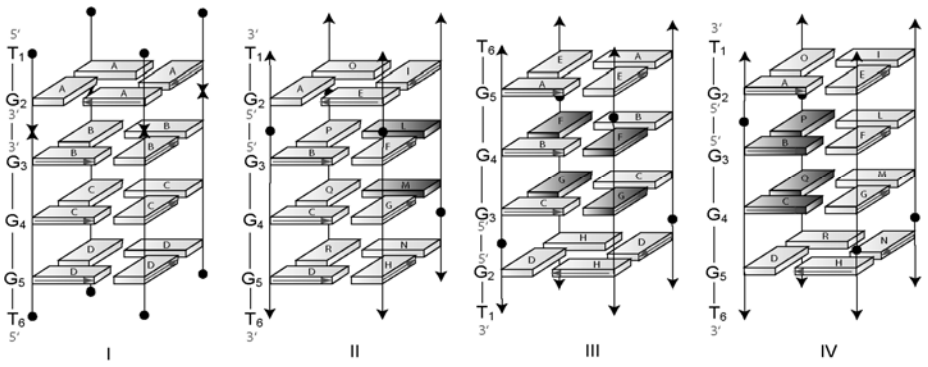

Fig. 1. Schematic illustration of the structure of **QS33** (**I**) and the theoretical strands arrangement for **QS55** (**II**, **III** and **IV**). According to NMR data, **QS55** adopts structure **III**. Black arrowheads and circles indicate 3' and 5' edges of each subunit/strand, respectively. The *anti* and *syn* guanines are depicted as light gray and dark gray solids, respectively. Arrows on solids indicate the direction of the proton donors and acceptors in Hoogsteen hydrogen bonds.

Interestingly, as observed in **Q33**, since all bases possess the same disposal (*anti*) with respect of the strand subunit orientation, the two tetrads comprising the inversion of polarity site will be characterized by an opposite (clockwise and anti-clockwise) orientation (Fig. 1, structure **I**).

As for **QS55**, inspection of its ^{1}H-NMR spectrum reveals that there are well defined signals. Particularly, the set of resonances in the region 10.5-11.5 ppm indicates the presence of eight imino protons involved in Hoogsteen hydrogen bonds. The double number of resonances in comparison with the four Gs in the sequence, clearly indicates that, in contrast with the behaviour of **QS33**, the quadruplex possesses a binary symmetry (C2). Useful information about the structure adopted by **QS55** arose from an in-depth analysis of NOESY spectrum (700 MHz, T = 25°C, mixing time = 100 ms) (Fig. 2).

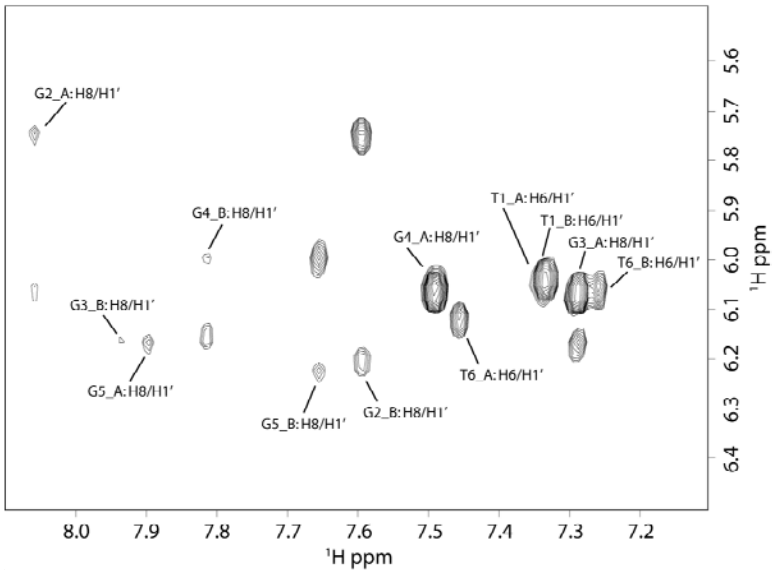

Fig. 2. 2D NOESY (700 MHz, 100 ms, T = 25°C) contour plots correlating base sugar H1' protons in **QS55**. Residues labelled as A and B belong to strands $^{3'}TG_{anti}{}^{5'}$-$^{5'}G_{syn}G_{syn}G_{anti}T^{3'}$ and $^{3'}TG^{5'}$-$^{5'}GGGT^{3'}$ (all residues *anti*), respectively.

Firstly, two intense H8-H1' NOEs (diagnostic of a *syn* glycosidic conformation) were clearly observable.

The entire set of NOE crosspeaks allowed us to assign these signals to two G nucleotides. In agreement with literature data, the H8 resonance of these *syn* G residues is upfield shifted with respect to those of the *anti* ones. Interestingly, these *syn* Gs turned out to be connected to each other, and each *syn* G residue shows a sequential connectivity to another *anti* G, which, in turn, is sequentially connected to a T residue. All this means that one out of two pairs of equivalent strands adopts the following arrangement $^{3'}$TG(*anti*)$^{5'}$-$^{5'}$G(*syn*)G(*syn*)G(*anti*)T$^{3'}$ (strand A). Consequently, the other strands' pair is characterized by all residues adopting an *anti* glycosidic conformation: $^{3'}$TG(*anti*)$^{5'}$-$^{5'}$G(*anti*)G(*anti*)G(*anti*)T$^{3'}$ (strand B). The analysis of the 2D TOCSY experiment along with the heteronuclear 2D ^{31}P-^{1}H COSY, further confirmed these hypothesis. From a theoretical point of view, **QS55** could adopt several strand arrangements in which all G residues are involved in tetrads differing each other from the relative strands orientation and from the glycosidic conformation of guanosines. However, taking into account the glycosidic arrangements of the bases within each strand, three structures can be proposed for **QS55** (**II**, **III** and **IV** in Fig. 1). Nevertheless, simple symmetry considerations allow us to easily rule out quadruplex arrangements **II** and **IV**, for which a total lack of symmetry is expected. Therefore, we propose that **QS55** adopts the quadruplex structure **III** in which equivalent strands are mutually opposed along the tetrads diagonal.

Interestingly, structure **III** is characterized by several remarkable structural features: i) the coexistence of two different types of quadruplex in the same complex. In fact, at the ends of the structure two

parallel all-*anti* tetrads are present, while the central core is formed by two antiparallel *syn-anti-syn-anti* tetrads; ii) the presence of a tetramolecular antiparallel quadruplex in the centre of the structure, that has never been described to date (although embedded between two parallel tracts); iii) the unprecedented occurring, to the best of our knowledge, of two *syn* Gs in a raw in an antiparallel alternating tract; iv) the presence of grooves variable in size along the quadruplex. Indeed, according to the literature data[1], the ends of the structure should show grooves of medium size, while the central core should possess a wide sized groove for the *anti*-Gs sides (residues B and C in Fig. 1, structure **III**) and a narrow sized groove for the *syn*-Gs sides (residues F and G in Fig. 1, structure **III**) of the quadruplex.

CD spectra for **QS33** and **QS55** and their natural counterpart $[d(TGGGGT)]_4$ were acquired at 20 °C and are reported as supporting information. All the measurements were performed at a concentration of 1×10^{-4} M, with the potassium buffer used for NMR experiments.

As already reported elsewhere,[2] the CD spectrum of $[d(TGGGGT)]_4$ is characterized by negative and positive bands at 243 and 263 nm, respectively. The CD spectrum of **QS33**, instead, is characterized by two positive bands at 254 and 274 and two negative bands at 232 and 270 nm, respectively. **QS33** CD spectrum appears very similar to that of the analogous quadruplex **Q33**,[2] similarly characterized by all G-residues in *anti* glycosidic conformations and a parallel-like strands arrangement.

As far as **QS55** is concerned, surprisingly its CD spectrum shows a negative band at 237 nm and a positive one at 260 nm. In this case the **QS55** spectrum is quite different from that showed by the analogous quadruplex **Q55**,[2] in which an all *syn*-G tetrad adjacent to the inversion of polarity site occurs in a parallel-like scaffold.

In order to estimate their thermal stability, **QS33** and **QS55** were subjected to melting and annealing CD experiments in comparison with $[d(TGGGGT)]_4$,[3] under the same experimental conditions. All the measurements were performed at a concentration of 1 x 10^{-4} M, using both Na^+ and K^+ buffers. Taking into account that the rates of quadruplex formation/dissociation are very slow, we collected the data at a 10 °C h^{-1} rate. Using a K^+ buffer, **QS55** and $[d(TGGGGT)]_4$ were still structured even at 90 °C, while we were able to obtain good melting profiles using a Na^+ buffer for all three samples. According to literature data concerning tetramolecular quadruplexes, in spite of the very slow scan rate used, an extreme hysteresis phenomenon was observed (no sigmoidal annealing profile was obtained), thus indicating that the systems were not at equilibrium. Therefore, from the melting curves, the apparent melting temperatures (T_m) of 65 °C, 47 °C and 53 °C could be measured for $[d(TGGGGT)]_4$, **QS33** and **QS55**, respectively. In conclusion, the results showed here describe an unprecedented quadruplex complex provided by peculiar structural features never reported to date. Our results complete investigation on the ability to form quadruplex of ODNs TGGGGT possessing inversion of polarity sites in the G-tract and they shed more light about the influence of 3'-3' and 5'-5' phosphodiester bonds on both the nucleosides glycosidic conformations and the strands relative orientation. The structures of **QS33** and **QS55**, along with those of **Q33** and **Q55** might inspire the design of new aptameric nucleic acids characterized by novel structural motifs hardly realizable with unmodified DNA/RNA.

In conclusion, the results described here could shed more light in the understanding of the variables involved in the formation and the stability

of quadruplex structures. Furthermore, the new features found in the two complexes **Q33** and **Q55** might guide the design of novel aptameric and catalytic nucleic acids provided with novel structural motifs for ligand-binding pockets and/or diverse molecular recognition capabilities from those possessed by native RNA/DNA sequences.

Notes and references

1. G. N. Parkinson in *Quadruplex Nucleic Acids*, ed. S. Neidle and S. Balasubramanian, RSC Publishing, London, 2006, ch. 1, pp. 1-30 and references cited therein.
2. V. Esposito, A. Virgilio, A. Randazzo, A. Galeone and L. Mayol, *Chemical Comm.*, 2005, 3953.
3. F. Aboul-ela, A. I. H. Murchie, D. G. Norman and D. M. J. Lilley, *J. Mol. Biol.*, 1994, **243**, 458.

5. Base modifications

To enhance the affinity of aptamers toward selected targets and to increase their stability in biological fluids, many modified nucleotides have been employed. Biophysical methods, such as CD spectroscopy and calorimetry, can be easily applied, allowing the determination of the influence of nucleotide modification on the conformation and thermal stability of different G-quadruplex structures.

5.1 Effect of the Incorporation of 2'-Deoxy-8-(Hydroxyl) Adenosine on the Stability of Quadruplexes Formed by Modified Human Telomeric DNA.

In this work, the adenines of the human telomeric repeat sequences d(TAGGGT) and d(AGGGT) were substituted by 2'-deoxy-8-(hydroxyl)adenosine ($A \rightarrow A^{OH}$). The oligonucleotides were synthesized on a Millipore Cyclon Plus DNA synthesizer, using solid phase β-cyanoethyl phosphoramidite chemistry and by using commercially available 5'-DMT-aminoprotected-8-(hydroxyl)adenosine-3'-phosphoramidite. The quadruplexes were formed by dissolving the purified oligonucleotides in the appropriate buffer and heating the solution at 90°C for 5 min. The solution was slowly cooled to room temperature and then equilibrated for one day at 4°C. The buffer 10 mM KH_2PO_4, 1.0 M KCl and 0.1 mM EDTA at pH 7.0 The CD spectra for all sequences were recorded at T = 20°C and at single strand concentration of 2×10^{-4} M (Fig 1). These spectra are diagnostic of parallel-stranded quadruplex structures.[1]

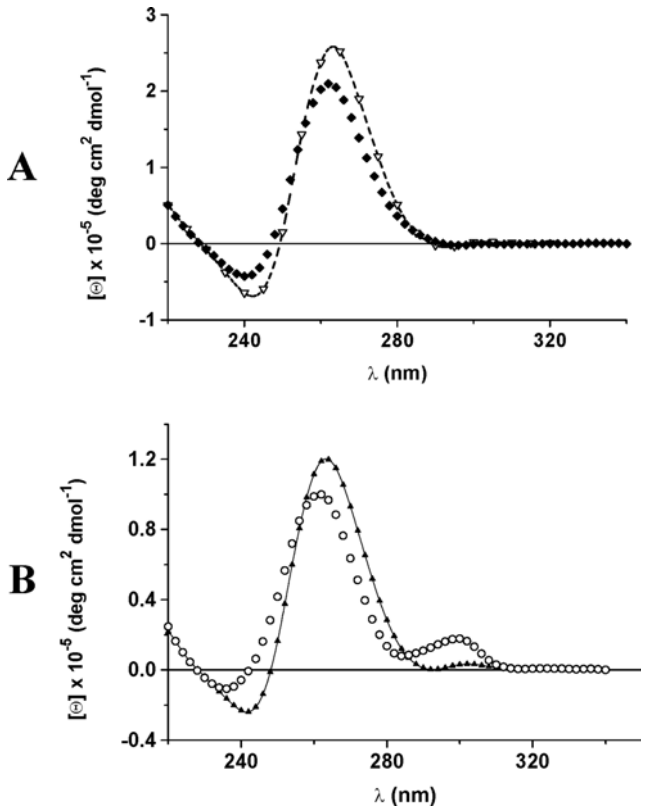

Fig. 1 CD spectra of quadruplexes at 10°C: A) (∇) $[d(TAGGGT)]_4$, (_) $[d(TA^{OH}GGGT)]_4$, B) (_) $[d(AGGGT)]_4$, (◦) $[d(AOHGGGT)]_4$.

Differential scanning calorimetry (DSC) measurements were carried out on a second generation Setaram Micro-DSC. The experiments were performed at constant single strand concentration of 8×10^{-4} M. The melting temperatures for each system were influenced by scan rate. This arises when the complexes are not at thermodynamic equilibrium during the temperature changes and it is due to slow rates of dissociation and/or association process. In this non-equilibrium conditions only the enthalpy change relative to the quadruplex dissociation process may be directly obtained by differential scanning microcalorimetry

measurements. The equilibrium melting temperatures were obtained by circular dichroism (CD) monitored denaturation experiments following a previously reported procedure.[2] The CD and DSC melting profiles for the modified and unmodified quadruplexes are shown in Fig. 2.

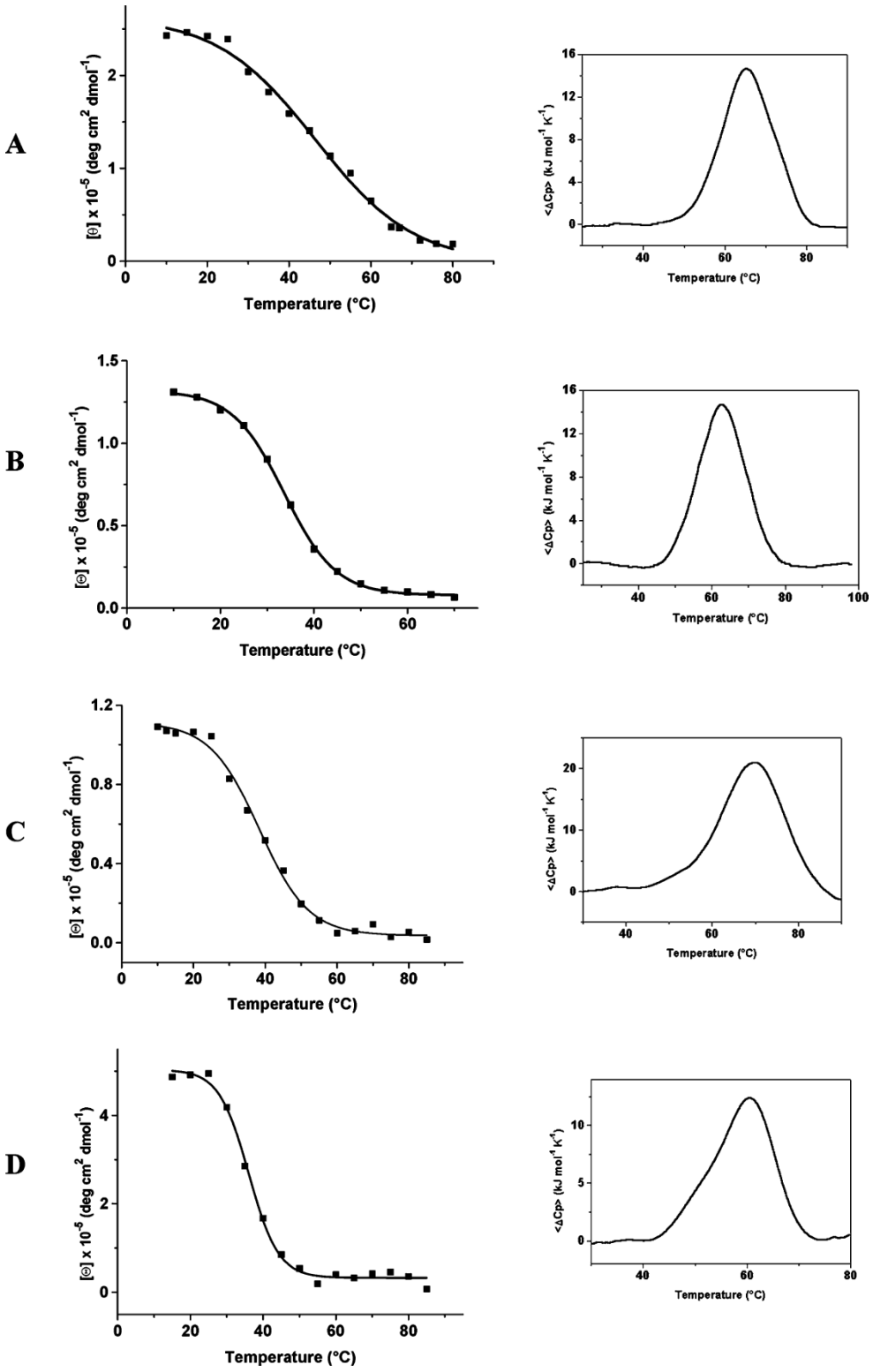

Fig 2 CD melting curves (on the left) and DSC profiles (on the right) for $[d(TAGGGT)]_4$ (A), $[d(TA^{OH}GGGT)]_4$ (B), $[d(AGGGT)]_4$ (C), and $[d(A^{OH}GGGT)]_4$ (D).

The thermodynamic parameters for the quadruplex dissociation are collected in Table 1.

Quadruplex	Tm (°C) (± 1)	ΔH° kJ mol^{-1}	ΔS° kJ mol^{-1}K^{-1}	ΔG°(298 K) kJ mol^{-1}(± 1)
[d(TAGGGT)]$_4$	51	210 ± 10	0.43 ± 0.03	81
[d(TA^{OH}GGGT)]$_4$	38	180 ± 9	0.33 ± 0.03	82
[d(AGGGT)]$_4$	37	280 ± 12	0.68 ± 0.04	77
[d(A^{OH}GGGT)]$_4$	37	280 ± 12	0.68 ± 0.04	77

Table 1 Thermodynamic parameters for the dissociation process of the quadruplexes.

The enthalpy value for the [d(AGGGT)]$_4$ is 70 kJ mol^{-1} higher than the enthalpy change reported for the [d(TAGGGT)]$_4$. This result is in perfect agreement with the NMR data that indicate the formation of an A-tetrad in the quadruplex [d(AGGGT)]$_4$ but not in the [d(TAGGGT)]$_4$ quadruplex.[3] For the modified sequences, inspection of Table 1 reveals that the introduction of the hydroxy group in position 8 of adenosine in the [d(AGGGT)]$_4$ quadruplex does not alter the thermodynamic stability of the resulting quadruplex. On the other hand, the same modification in the [d(TAGGGT)]$_4$ quadruplex causes a decrease of both the entalpic and entropic terms. Further, the melting temperature of the [d(TA^{OH}GGGT)]$_4$ quadruplex is 13°C lower than the melting temperature obtained for the unmodified quadruplex. These data seem to indicate that the presence of the A^{OH} residue in the [d(TA^{OH}GGGT)]$_4$ significantly affects the quadruplex conformation whereas the structure of the [d(A^{OH}GGGT)]$_4$ quadruplex is probably very similar to the unmodified quadruplex. To give further insight into the conformational features of the modified quadruplexes, molecular modelling studies are currently in progress in our laboratories.

Notes and references

1. Petraccone, L.; Erra, E.; Esposito, V.; Randazzo, A.; Mayol, L.; Nasti, L.; Barone, G.; Giancola, C. *Biochem* 2004, **43**, 4877–84.
2. Petraccone, L.; Erra, E.; Esposito, V.; Randazzo, A.; Galeone, A.; Barone, G.; Giancola, C. *Biopolymers* 2005, **77**, 75–85.
3. Patel, P.K.; Koti, A.S.R.; Hosur, R.V. *Nucleic Acids Res.* 1999, **27**, 3836–3843.

5.2 Biophysical Properties of Quadruplexes Containing Two or Three 8-Bromodeoxyguanosine Residues

In this frame, we report the incorporation of two or three 8-Br-dG on the same sequence. Commercially available 5'-ODMT- N-isobutyrroyl-8-bromodeoxyguanosine-3'-phosphoramidite was used to prepare the modified oligonucleotides. The pentamers d(TG^{Br}G^{Br}GT), d(TG^{Br}GG^{Br}T), d(TGG^{Br}G^{Br}T), and d(TG^{Br}G^{Br}G^{Br}T) were assembled using the standard solid phase β-cyanoethylphosphoramidite chemistry. The crude oligomers were purified by HPLC and desalted. The NMR samples were prepared at a concentration of approx. 1.0 mM (0.5 ml, 90% H_2O/10% D_2O), having 10 mM potassium phosphate, 70 mM KCl, 0.2 mM EDTA (pH 7.0) buffer. The samples were annealed for 5-10 minutes at 80°C and slowly cooled down to room temperature, then the ^{1}H-NMR spectra of them were recorded by using pulsed-field gradient WATERGATE[1] for H_2O suppression. Interestingly, only d(TG^{Br}GG^{Br}T) and d(TG^{Br}G^{Br}GT) were able to fold into a well-defined quadruplex structure. In fact, three well defined singlets were present in the region between 10.5 and 12 ppm of their spectra. These signals were

attributed to imino protons involved in the Hoogsteen hydrogen bonds of G quartets. Moreover, a total of three signals, belonging to one guanine H8 and two thymine H6 protons, were present in the region between 7 and 8.5 ppm, and this is in perfect agreement with the presence of two 8-Br-dG residues in the oligonucleotides (Fig. 1).

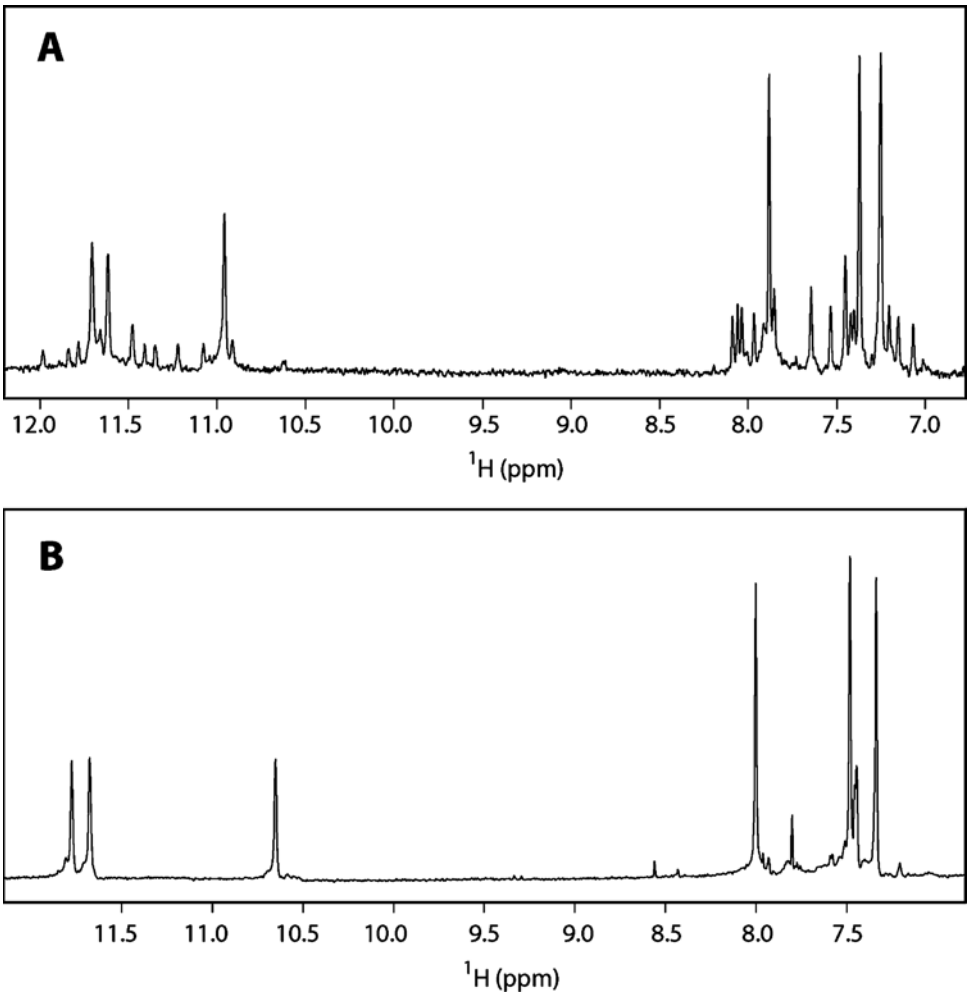

Fig 1 Expanded region of the proton NMR spectra of $[d(TG^{Br}G^{Br}GT)]_4$ (A) and $[d(TG^{Br}GG^{Br}T)]_4$ (B) (500 MHz, T = 25°C).

Thus, these data are consistent with the formation of quadruplex structures characterized by three G-tetrads and possessing a fourfold symmetry with all strands parallel to each other. NOESY experiments (500 MHz, T= 300 K, mixing time = 50, 100, and 200 ms) acquired for

both molecules further corroborated this observation. Moreover, as reported for other parallel quadruplex structures,[2] the observed polarity NOE connectivities (G H8 to ribose protons on the 5' side only) suggested the right handed nature of the helices of the quadruplexes. As for the glycosidic torsion angles of the G residues, useful information could be obtained analyzing the relative intensities of NOEs between G H8 and ribose H2' compared with the NOEs observed between G H8 and H1'. Thus, as for the same sequence containing the unmodified G residue, weak NOE between G H8 and H1' and strong NOEs between G H8 and ribose H2' indicates it is in an *anti* conformation. Unfortunately, the lack of the H8 proton in the 8-Br-dG residues prevented us from determining the conformation of the glycosidic torsion angle for these residues. To obtain further information on the conformational features of the modified quadruplexes, we performed circular dichroism (CD) experiments in 10 mM KH_2PO_4 pH 7.0, 200 mM KCl, 0.1 mM EDTAat single strand concentration of 2×10^{-4} M. CD spectra of the d(TGGGT), d(TG^{Br}G^{Br}GT), d(TG^{Br}GG^{Br}T) were acquired at 10°C (Fig. 2).

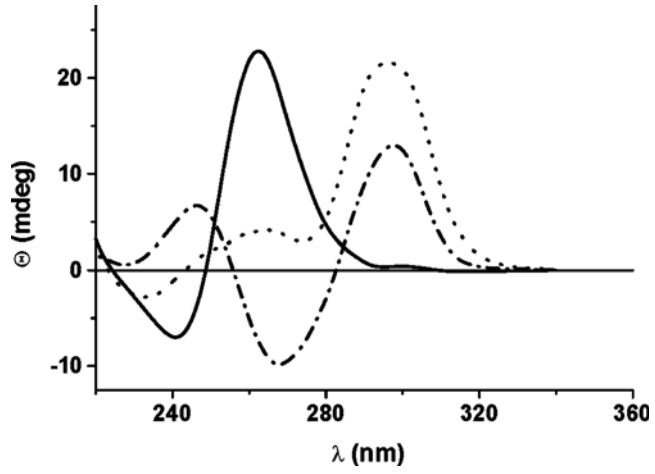

Fig. 2 CD spectra of quadruplexes at 10°C, (—) [d(TGGGT)]$_4$, (· · ·) [d(TG^{Br}G^{Br}GT)]$_4$, (- ·– ·– ·– ·-)[d(TG^{Br}GG^{Br}T)]$_4$.

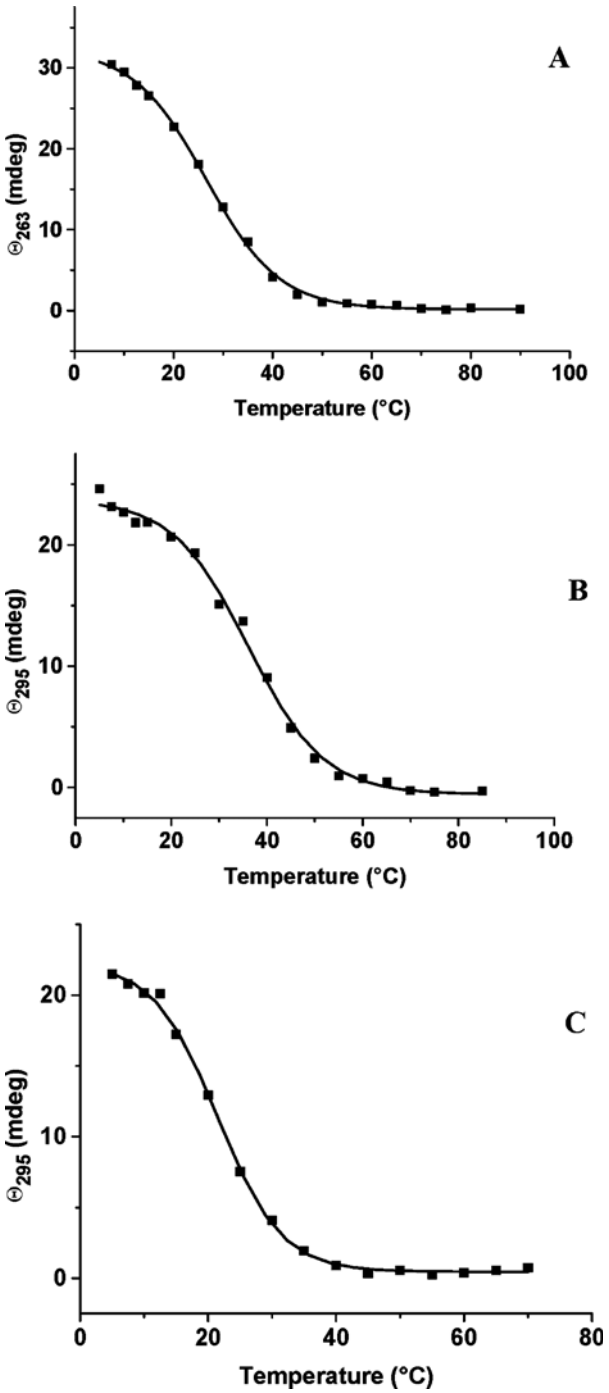

Fig. 3 CD melting curves of the $[d(TGGGT)]_4$ (A), $[d(TG^{Br}G^{Br}GT)]_4$ (B), and $[d(TG^{Br}GG^{Br}T)]_4$ (C).

The spectra of the unmodified sequence show the typical bands for quadruplexes involving four parallel strands, a positive band at 263 nm, and a negative band at 243 nm[3] whereas the introduction of the 8-Br-dG

residues causes significant changes in the CD spectra. Particularly, the CD spectrum of d(TG^{Br}GG^{Br}T) exhibited a negative band at 263 nm and a positive band at 295 nm whereas the spectrum of d(TG^{Br}G^{Br}GT) has two positive bands at about 263 and 295 nm. These data suggest that the 8-Br-dG residues affect the structure of the quadruplex and that the 8-Br-dG residues actually adopt a *syn* glycosidic conformation. In order to evaluate the thermodynamic stability of the modified quadruplexes, CD thermal experiments were performed. Taking into account that the rates of quadruplex formation/dissociation are extremely slow, to avoid a kinetic influence on the collected data, we allowed thermodynamic equilibrium to be reached at each temperature by following a previously reported procedure.[4] All the examined quadruplexes showed sharp transitions and well-shaped sigmoidal curves (Fig. 3).

The thermodynamic parameters obtained from the CD melting analysis are reported in Table 1.

Table 1 Thermodynamic parameters for the dissociation of the quadruplexes

Quadruplex	T_m (°C) ± 1	ΔH°v.H.(KJ mol^{-1})	ΔS°(KJ mol^{-1} K^{-1})	ΔG° 298 (KJ mol^{-1}) ± 1
[d(TGGGT)]$_4$	30	220 ± 11	0.51 ± 0.03	68
[d(TG^{Br}GG^{Br}T)]$_4$	21	200 ± 9	0.47 ± 0.02	60
[d(TG^{Br}G^{Br}GT)]$_4$	37	270 ± 12	0.65 ± 0.04	76

Interestingly, the introduction of two 8-Br-dG residues in different positions of the same sequence d(TGGGT) affects in the opposite way the thermodynamic stability of the canonical [d(TGGGT)]$_4$ quadruplex. Indeed, the modification of two adjacent dG residues near the dT at the 5' position to give d(TG^{Br}G^{Br}GT) results in a more stable quadruplex in comparison with the unmodified one, while modification of two, non-adjacent dG residues to give [d(TG^{Br}GG^{Br}T)]$_4$ results in a decreased thermodynamic stability. Indeed, the melting temperatures were 21°C

and 37°C for $[d(TG^{Br}GG^{Br}T)]_4$ and $[d(TG^{Br}G^{Br}GT)]_4$, respectively, whereas a melting temperature of 30°C was observed for the unmodified quadruplex. In addition, the enthalpy change value for the $[d(TG^{Br}G^{Br}GT)]_4$ melting is 50 kJ mol^{-1} higher than the value obtained for the canonical $[d(TGGGT)]_4$ quadruplex, indicating that the highest thermodynamic stability observed for this quadruplex is enthalpic in origin. These data demonstrate that the substitution of a single atom, such as a bromine atom at the C8 position of guanines, significantly affect the structure and stability of the resulting quadruplex.

Notes and references

1. Piotto, M.; Saudek, V.; Sklenar, V.J. *J. Biomol. NMR* 1992, **2**, 661–665.
2. Wang, Y.; Patel, D.J. *J. Mol. Biol.* 1993, **234**, 1171–1183.
3. Petraccone, L.; Erra, E.; Esposito, V.; Randazzo, A.; Mayol, L.; Nasti, L.; Barone, G.; Giancola, C. *Biochem.* 2004, **43**, 4877–4884.
4. Petraccone, L.; Erra, E.; Esposito, V.; Randazzo, A.; Galeone, A.; Barone, G.; Giancola, C. *Biopolymers* 2005, **77**, 75–85.

5.3 Effect of the introduction of an A-residue into a quadruplex forming oligonucleotide containing a 5'-5' polarity of inversion site

In order to investigate the apparent stabilizing properties of a 5'-5' inversion of polarity site embedded in a G-quadruplex structure, we have undertaken a systematic exploration concerning several sequences containing such backbone modification. Here we report preliminary results about sequence $^{3'}TGA^{5'-5'}GGT^{3'}$. Oligonucleotide $^{3'}TGA^{5'-5'}GGT^{3'}$

was synthesized using standard methods and 3'-phosphoramidites and 5'-phosphoramidites.

NMR samples were prepared at a concentration of approximatively 6 mM, in 0.6 ml (H_2O/D_2O 9:1) buffer solution having 40 mM KH_2PO_4, 280 mM KCl, 0.2 mM EDTA, pH 7.0. The simple appearance of the ^{1}H-NMR spectrum of $^{3'}$TGA$^{5'-5'}$GGT$^{3'}$ (Fig. 1A) indicates that, in the conditions used here, the modified oligomer forms mainly a single well-defined hydrogen-bonded conformation. It shows the presence of three well defined signals in the region 10.5-11.5 ppm, distinctive of imino protons involved in Hoogsteen hydrogen bonds of G-quartets. Moreover, six signals in the aromatic region were clearly observable. A further signal at 9.68 ppm is detectable. We assigned this signal to AH2 consistent with a D_2O exchange study (Fig. 1) and NOE connettivities. Taking into account this result, only three signals can be ascribed to imino protons involved in Hoogsteen hydrogen bonds of G-quartets suggesting the possibility of a four fold symmetry for the quadruplex structure adopted by $^{3'}$TGA$^{5'-5'}$GGT$^{3'}$.

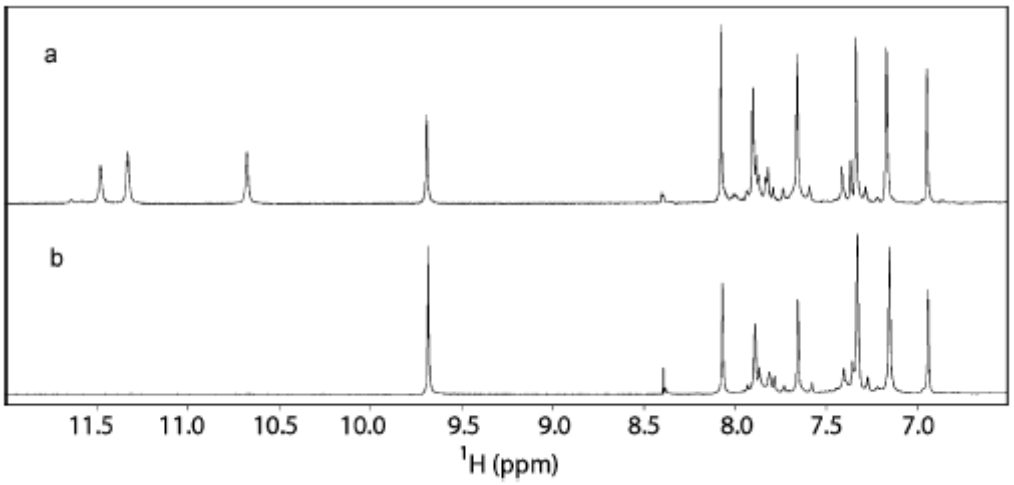

Fig.1 ^{1}H NMR spectra of $^{3'}$TGA$^{5'-5'}$GGT$^{3'}$ in H_2O (A) and in D_2O (B).

The NOESY spectrum (700 MHz, t = 25 °C, mixing time 100 ms) discloses interesting diagnostic information. Particularly, the lack of strong NOEs between any G H8 and H1' of the same residue, strongly suggests that all residues are in the *anti* glycosidic conformation (Fig. 2).

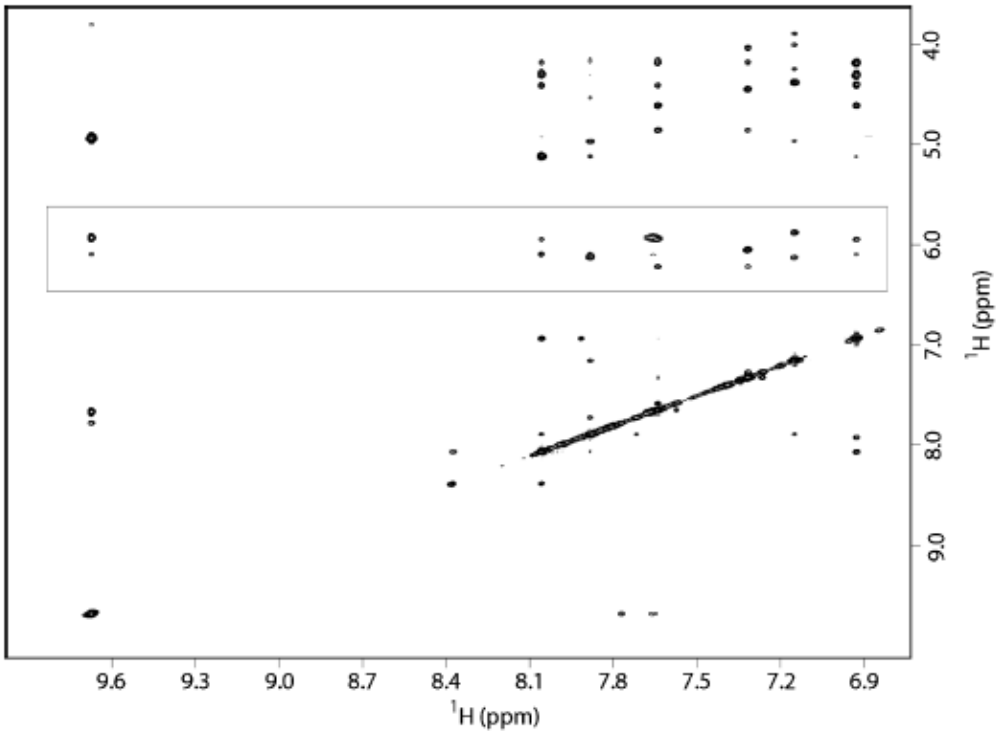

Fig.2 Expanded region of the NOESY spectrum of 3'TGA$^{5'-5'}$GGT$^{3'}$. Section correlating G-H8 and H1' protons is boxed.

Sequence $^{3'}$TGA$^{5'-5'}$GGT$^{3'}$ could, theoretically, arrange in several type of quadruplex structures, among them structures **I** and **II** (Fig. 3) are consistent with NMR data, since according to their symmetry properties, they should show only a single set of signals due to the equivalence of the four strands. Structure **I** could enclose an A-tetrad, a planar arrangement of four A residues similar to a G-tetrad, already evidenced in other quadruplex structures[1-3]. On the other hand, structure **II** could be stabilized by two AGAG-tetrads (Fig. 3, on the right) sited in the centre of the complex. Although some chemical probing studies on

d(G_3T'TAG_3) and d(T'TAG_3) units of different lengths suggest a probable formation of AGAG-tetrads in the quadruplex structures[4], however, other studies question the existence of the AGAG tetrad because of its relatively high energy and the lacking of oxygen atoms required for the formation of the cation site[5,6]. Finally, a theoretical investigation on structures and properties of mixed DNA bases tetrads[7] revealed that a AGAG-tetrad would be characterized by a V-shaped structure suggesting that it cannot be stacked with the G-tetrads in the stem of quadruplexes and hence, that this tetrad arrangement may not be important in the structure stabilization, particularly in presence of cations as Na^+, or K^+.

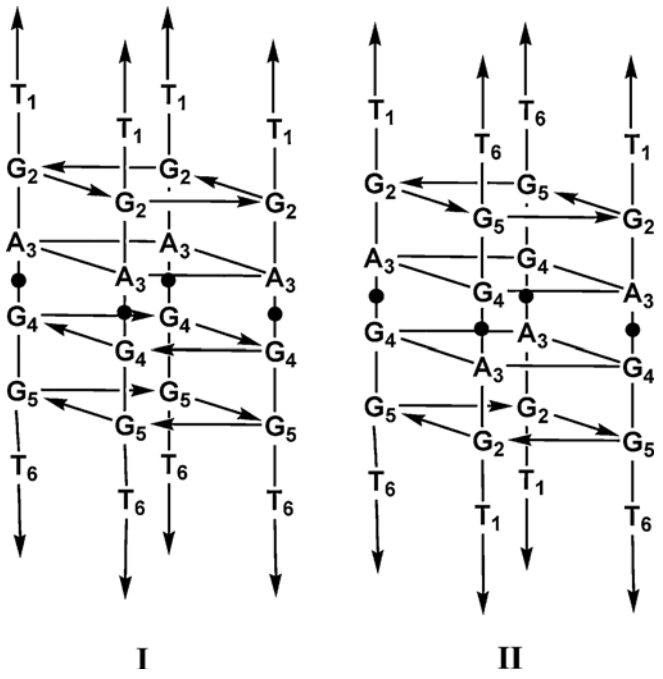

Fig. 3 Possible structures for the quadruplex adopted by 3'-TGA-5'-5'-GGT-3'. Equivalent residues are similarly labeled.

Concerning oligonucleotide $^{3'}$TGA$^{5'-5'}$GGT$^{3'}$ the sequential NOE connectivity pathway along the sequence is broken in the 5'-5' inversion of polarity site. However, in the tract 3'-TGA-5' it gives patters and intensities of NOE consistent with the adenine base stacking within the

quadruplex with all purine glycosidic torsion angles in the anti range. The most probable arrangement of the A residues should involve a putative interstrand hydrogen bonding between the 6-NH2 group and N1. In fact, this brings the adenine H2 of one strand in close proximity to the N7 of an adjacent adenine, potentially accounting for the unusual downfield shift observed for AH2, as reported by other authors[3]. The occurrence of a crosspeak between adenine H2 and H8 further suggests a planar arrangement of the four adenine bases. According to the above results and considerations, structure **I** was supposed to be the most probable one, although further NMR investigations are in progress.

Notes and references

1. Patel, P. K., Koti, A. S. R., and Hosur, R. V. 1999. *Nucleic Acids Res.*, **27** (19), 3836-3843.
2. Gavathiotis, E., and Searle, M. S. 2003. *Org. Biomol. Chem.*, **1**, 1650-1656.
3. Searle, M. S., Williams, H. E. L., Gallagher, C. T., Grant, R. J., and Stevens, M. F. G 2004. *Org. Biomol. Chem.,* **2***,* 810-812.
4. Balagurumoorthy, P., and Brahmachadri, S. K. 1994. *J. Biol. Chem.,* **269***,* 21858-21869.
5. Mohanty, D., and Bansal, M. 1995. Chain Folding and A:T *Biophys. J.,* **69***,* 1046-1067.
6. Sundquist, W. I., and Klug, A. 1989. *Nature (London),* **342***,* 825-829.
7. Gu, J., and Leszczynsky, J. 2000. *J. Phys. Chem. A*, **104**, 1898-1904.

CHAPTER 2

Thrombin Binding Aptamer

1. Thrombin Binding Aptamer: backbone modifications

In 1992, Block et al.[1] screened a pool of about 10^{13} synthetic oligonucleotides for their interaction with thrombin, which is a serine protease with multiple functions in hemostasis whose roles in coronary heart disease and other thrombotic disorders have promoted efforts toward the identification of specific inhibitors. This study led to the identification of a consensus DNA 15-mer, namely $^{5'}$GGTTGGTGTGGTTGG$^{3'}$ (TBA: Thrombin Binding Aptamer), which was found to be a potent inhibitor of thrombin with an EC_{50} of 20 nM in a fibrinogen clotting assay.[2,3] The three-dimensional solution structure of TBA was solved using NMR techniques.[4,5] TBA is characterized by a chair-like structure consisting of two G-tetrads connected by two TT loops and a single TGT loop. As suggested by others,[6] TBA binds at the anion exosite I of thrombin in a conformation little changed from that of the unbound species.

In order to improve the properties of TBA, a number of researches[7-9] have been devoted to its structure-activity relationship to post-SELEX modifications. For example, He et al. [8] synthesized numerous TBA analogues containing modified guanosine carrying several substituents at 8 position or at the exocyclic amino group. In a different paper,[9] the same authors reported the synthesis and the thrombin inhibiting properties of several based TBA oligonucleotides containing neutral formacetal linkages. However, although TBA was the first aptamer to be discovered, no studies dealing with the effects of introduction of sites of polarity inversion have been reported yet, even if such interesting

backbone alteration has already been utilized in another aptamer, namely Macugen[10] (pegaptanib sodium), that has been recently approved by the US FDA for the treatment of neovascular age-related macular degeneration.

Notes and references

1. Block,L.C., Griffin,L.C., Latham,J.A. Vermaas,E.H., Toole,J.J. (1992) *Nature,* **355**, 564-566.
2. Griffin,L.C., Tidmarsh,G.F., Bock,L.C., Toole,J.J., Leung,L.K. (1993) *Bloods*, **81**, 3271-3276.
3. Li,W.X., Kaplan,A.V., Grant,G.F., Toole,J.J., Leung,L.K. (1994) *Blood*, **83**, 677-682.
4. Wang,K.Y., McCurdy,S., Shea,R.G., Swaminathan,S., Bolton,P.H. (1993) *Biochemistry*, **32**,1899-1904.
5. Macaya,R.F., Schultze,P., Smith,F.W. Roe,J.A. Feigon,J. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 3745-3749.
6. Paborsky,L.R., McCurdy,S.N., Griffin,L.C., Toole,J.J., Leung,L.K. (1993) *J. Biol. Chem*., **268**, 20808-20811.
7. Dias,E., Battiste,J.L., Williamson,J.R. (1994) *J. Am. Chem. Soc.*, **116**, 4479-4480.
8. He,G.-X., Krawczyk,S.H., Swaminathan,S., Regan,S.G., Dougherty,J.P., Terhorst,T., Law,V.S., Griffin,L.C., Coutrè,S., Bischofberger,N. (1998) *J. Med. Chem.*, **41**, 2234-2242.
9. He,G.-X., Williams,J.P., Postich,M.J., Swaminathan,S., Regan,S.G., Terhorst,T., Law,V.S., Griffin,L.C., Cheri,M.T., Coutrè,S., Bischofberger,N. (1998) *J. Med. Chem.*, **41**, 4224-4231.
10. Nimjee,S.M., Rusconi,C.P., Sullenger,B.A. (2005) *Annual Review of Medicine*, **56**, 555-583,

1.1 A mini-library of TBA analogues containing 3'-3' and 5'-5' inversion of polarity sites

In this frame, we have undertaken the synthesis and the study of a mini library composed of several TBA analogues containing a 3'-3' or a 5'-5' inversion of polarity site at different positions into the sequence. Particularly, in this paper, we present preliminary results about their structural and biological properties.

The mini-library of TBA analogues consists of two series of oligonucleotides, namely, ODNs 33 containing 3'-3' and ODNs 55 containing 5'-5' inversion of polarity sites. Both series were designed introducing an inversion of polarity site between each residue in the TBA sequence $^{5'}$GGTTGGTGTGGTTGG$^{3'}$. Sequences of the fourteen ODNs obtained are listed in Table 1.

Table 1 Sequences of the fourteen TBA analogues, containing a 3'-3' or a 5'-5' inversion of polarity site at different positions, belonging to the mini library.

5'-5' Series		3'-3' Series	
$^{3'}$G$^{5'-5'}$GTTGGTGTGGTTGG$^{3'}$	**55 I**	$^{5'}$G$^{3'-3'}$GTTGGTGTGGTTGG$^{5'}$	**33 I**
$^{3'}$GG$^{5'-5'}$TTGGTGTGGTTGG$^{3'}$	**55 II**	$^{5'}$GG$^{3'-3'}$TTGGTGTGGTTGG$^{5'}$	**33 II**
$^{3'}$GGT$^{5'-5'}$TGGTGTGGTTGG$^{3'}$	**55 III**	$^{5'}$GGT$^{3'-3'}$TGGTGTGGTTGG$^{5'}$	**33 III**
$^{3'}$GGTT$^{5'-5'}$GGTGTGGTTGG$^{3'}$	**55 IV**	$^{5'}$GGTT$^{3'-3'}$GGTGTGGTTGG$^{5'}$	**33 IV**
$^{3'}$GGTTG$^{5'-5'}$GTGTGGTTGG$^{3'}$	**55 V**	$^{5'}$GGTTG$^{3'-3'}$GTGTGGTTGG$^{5'}$	**33 V**
$^{3'}$GGTTGG$^{5'-5'}$TGTGGTTGG$^{3'}$	**55 VI**	$^{5'}$GGTT GG$^{3'-3'}$TGTGGTTGG$^{5'}$	**33 VI**
$^{3'}$GGTTGGT$^{5'-5'}$GTGGTTGG$^{3'}$	**55 VII**	$^{5'}$GGTT GGT$^{3'-3'}$GTGGTTGG$^{5'}$	**33 VII**

All oligonucleotides were synthesized by solid phase β-cyanoethyl phosphoramidite chemistry and purified by standard methods and then subjected to a preliminary analysis by ^{1}H NMR and CD spectroscopy. Afterwards, the promising sequences were selected for a more detailed

analysis. The simple appearance of ^{1}H NMR spectra of ODNs **55 II**, **55 III** and **55 IV** (Fig. 1) indicates that, in the conditions used, these modified oligomers form mainly a single well-defined hydrogen-bonded conformation. ^{1}H NMR spectra of the other ODNs showed more than one set of signals suggesting the presence of several species in solution. In some cases a severe line broadening prevents us to obtain information about the number of present species.

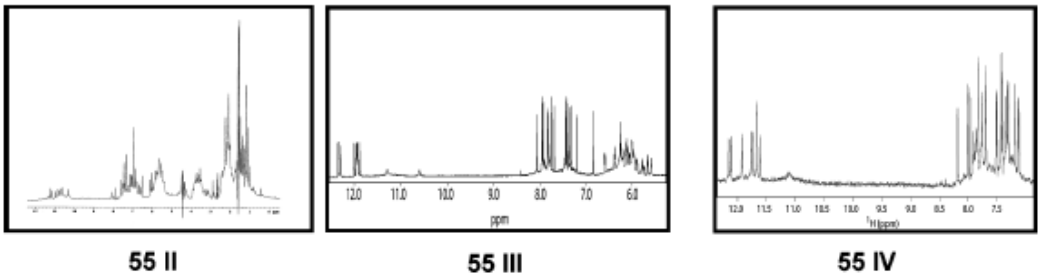

Fig.1 ^{1}H NMR spectra of **55 II**, **55 III** and **55 IV**.

In order to acquire information about the effects of the inversion of polarity site on the structural features of the resulting complexes, CD spectra were acquired for ODNs **55 II**, **55 III** and **55 IV** and their natural counterpart. All CD spectra exhibit two positive bands around 248 and 294 nm and a negative one around 267 nm. In particular, the CD spectra of **55 III** and TBA, performed at 20°C, are almost superimposable suggesting the formation of an anti-parallel quadruplex structure,[1-3] containing residues in *syn*-glycosidic conformations for **55 III**. Melting CD experiments, performed at a scan rate of 10°C/h, provided the melting temperatures of 50°C, 58°C, 61°C and 54°C for complexes **55 II**, **55 III** and **55 IV** and TBA, respectively. These results encouraged us to undertake both a more detailed structural investigation and an estimation of anti-thrombin activity of complexes showing an improved thermal stability compared to TBA. Here we mainly report results on the TBA analogue **55 III** ($^{3'}$G1-G2-T3$^{5'\text{-}5'}$T4-G5-G6-T7-G8-

T9-G10-G11-T12-T13-G14-G15[3'])**.** The ^{1}H-NMR spectrum of **55 III** (700 MHz, T=25°C) shows eight well defined signals in the region 11.5-12.5 ppm, attributable to imino protons involved in Hoogsteen hydrogen bonds of at least two G-quartets, and of two broader and less intense singlets at 10.6 and 11.29 ppm, respectively, probably belonging to two imino protons of two thymines involved in mutual hydrogen bonding. Moreover, fifteen signals in the aromatic region due to the presence of nine guanine H8 and six thymine H6 protons were clearly observable. A combination of the analysis of 2D NOESY (700 MHz, T=25°C), TOCSY spectra (700 MHz, T=25°C), and ^{31}P-NMR spectroscopy (202 MHz, T=25 °C) allowed us to get the almost complete assignment (Tab. 2) of both exchangeable and nonexchangeable protons and phosphorus resonances of **55 III**.

Table 2 Proton chemical shifts (700 MHz) of **55 III** in 10 mM KH_2PO_4, 70 mM KCl, 0.2 mM EDTA (pH 7.0, T=25°C).

Base	H8/H6	H1′	H2′	H2″	H3′	H4′	H5′[a]	H5″[a]	H2/Me	NH″[b]	^{31}P
3′G1	7.94	6.15	2.39	2.56	4.73	4.18	4.21	3.94		11.91	−1.72
G2	6.84	5.68	3.77	2.94	4.83	4.42	4.23	4.27		12.35	−1.51
5′T3	7.83	6.61	2.31	2.45	4.92	4.41	3.69	4.01	1.98		−1.60
5′T4	7.36	5.76	1.80	2.14	4.56	4.45	3.84		1.37	11.35	−1.36
G5	7.42	6.01	3.41	2.92	4.90	4.30	4.39			12.30	−0.89
G6	7.70	5.90	2.58	2.71	5.10	4.40	4.21	3.95		12.03	−1.09
T7	7.85	6.40	2.44	2.58	4.84	4.41	4.22	3.76	1.97		−1.56
G8	7.43	5.68	2.02	2.33	4.69	4.13	4.07	3.81			−1.76
T9	7.22	5.80	1.90	2.31	4.59	4.48	3.67	3.51	1.54		−1.98
G10	7.45	6.02	3.64	2.90	4.90	4.37	4.15	4.34		11.91	−1.84
G11	8.07	5.98	2.32	2.90	5.09	4.40	4.25			12.37	−0.53
T12	7.77	6.26	2.27	2.59	4.90	4.35	4.24	4.34	1.96		−2.20
T13	7.34	6.20	2.19	2.76	4.93	4.30	4.24	4.33	1.35	10.65	−1.01
G14	7.41	6.09	3.63	3.00	4.96	4.58	4.30	4.38		11.95	−1.80
3′G15	7.96	6.12	2.39	2.59	4.82	4.53	4.20	4.30		11.97	

[a] No stereospecific assignment has been done for H5' and H5"

[b] Imino proton chemical shifts have been assigned at T=5°C

Interestingly, the intensities of NOESY (700 MHz, T=25 °C, mixing time 100 ms) crosspeaks between the H8 proton bases and sugar H1' resonances indicate that four (G2, G5, G10, G14) out of nine Gs of **55 III** adopt a *syn* glycosidic conformation, (Fig. 2) where the H8 resonances of *syn* G residues are upfield shifted with respect to those of the *anti* ones.[4-6] Then, four *anti*-Gs (G1, G6, G11 and G15) have classical H8/H2'-H2" sequential connectivities to 5' neighboring *syn*-Gs (G2, G5, G10 and G14, respectively), indicating that the subunits G1-G2, G5-G6, G10-G11 and G14-G15 are involved in the formation of a four-stranded helical structure (underlined residues adopt a *syn* glycosidic conformation). The alternation of *syn* and *anti* G residues within each strand suggests that, as TBA, **55 III** folds into a monomolecular foldback quadruplex, characterized by two G tetrads. Further, a number of unusual NOE connectivities was observed between *syn*-Gs and Ts, indicating that $^{5'}$TG$^{3'}$ and $^{5'}$GT$^{3'}$ tracts do not adopt a helical winding, thus suggesting that, most probably, the TT and TGT tracts form loops.

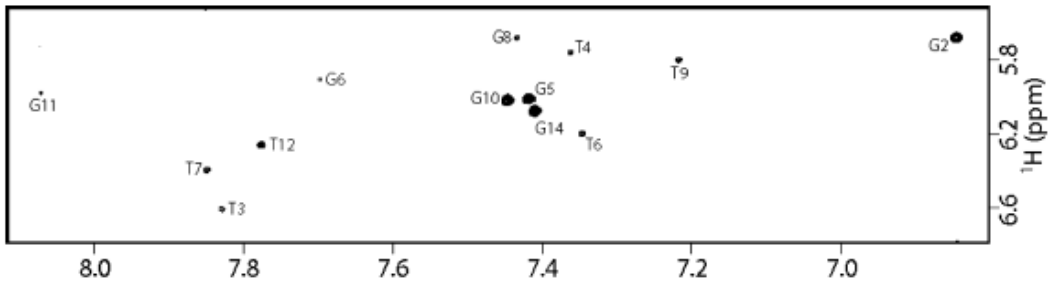

Fig.2 Expanded region of the NOESY spectrum of **55 III** (700 MHz, T= 25°C, mt=200 ms) correlating G-H8 and H1' protons.

The acquisition of a ^{1}H-^{15}N HSQC experiment allowed us to assign exchangeable protons to their respective G or T residues. To be noted that signals at 10.65 and 11.35 ppm belonging to T residues are characterized by strong NOE effects between each other, thus suggesting that two thymines are in close proximity and interacting by a

mutual hydrogen bond, likewise as observed in the unmodified TBA. The whole of data suggests that the structure of **55 III** is a foldback quadruplex characterized by two G tetrads formed by the residues G1, G6, G10, G15 and G2, G5, G11, G14, respectively, where the underlined residues possess a *syn* glycosidic conformation. Interestingly the two tetrads assume a *anti-anti-anti-syn* and *syn-syn-syn-anti* arrangement of the bases. Hence, **55 III** folds into a chair-like quadruplex structure possessing three strands parallel to each other and only one strand oriented in the opposite direction (Fig. 3).

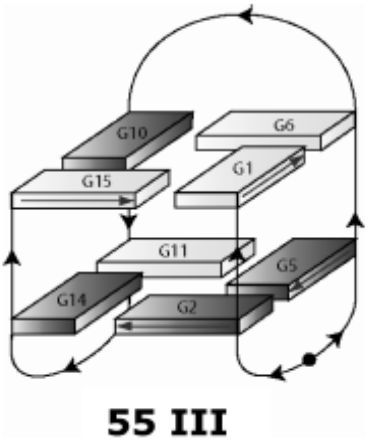

Fig.3 Schematic illustration of the structure adopted by **55 III**.

In order to study the effects of the structural changes on the biological activity of **1**, if any, we performed a prothrombin time (PT) assay. PT assay on **55 III** was performed on human plasma in strict comparison with TBA. In order to eliminate the variation caused by the measurements performed on different days, the samples of **55 III** and TBA were prepared at the same time and folded together by heating the samples for 10 minutes at 80°C and slowly cooling them down at room temperature. The assays have been conducted after one week from the preparation procedure. As shown in Fig. 4, overall, TBA showed a statistically significant more marked effect of **55 III** both as onset of the

activity and endurance. Indeed, incubation studies demonstrated that TBA effect is already present after 10 seconds of incubation and lasts up to 15 minutes of incubation. On the other hand, **55 III** showed a slower onset of the effect, indeed PT time started to be significantly prolonged following 5 minutes of incubation. Also the magnitude of the activity of **55 III** resulted reduced compared to TBA. The biological data reported here imply that the modifications of the structure adopted by **55 III** affect sensibly the biological activity, and therefore its interaction with the thrombin.

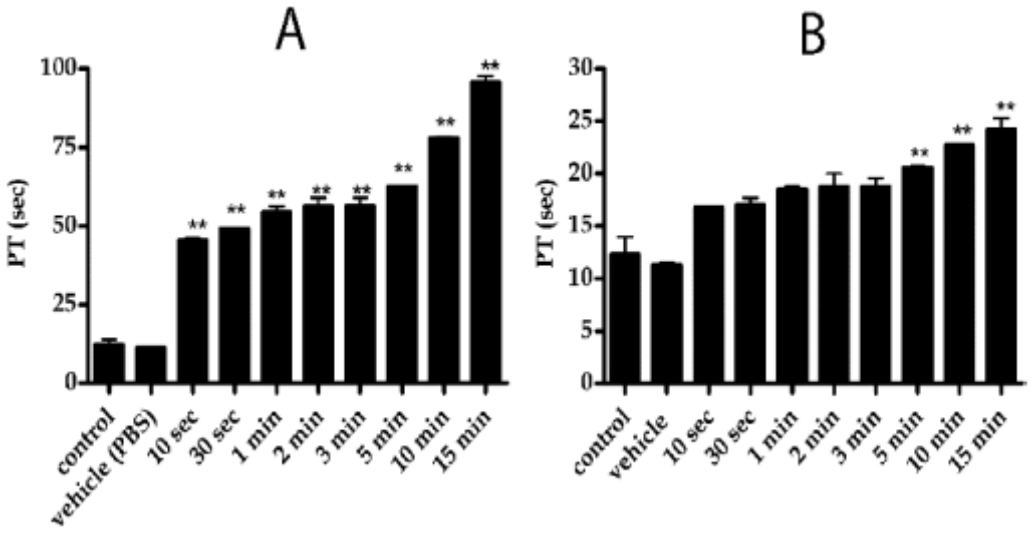

Fig.4 Prothrombin time assessed by using human plasma. On the X axis the minutes of incubation are reported, while PT time in seconds is reported on Y axis. Panel A refers to TBA, while panel B to **55 III**.

The ODNs of the mini-library show quite different structural behaviour. There is evidence for a main stable conformation only for the 5'-5' series (ODNs **55 II**, **55 III** and **55 IV**). In fact, most of ^{1}H NMR spectra of ODNs belonging to the 3'-3' series showed more than one set of signals suggesting the presence of several species in solution. Furthermore, an inversion of polarity site 5'-5' positioned in the G-tract prevents the formation of a stable quadruplex structure. However, it could be of

interest investigate the different conformations present in solution, since, in some cases (for example ODN **55 VI**) a significant anticoagulant activity resulted, in spite of its rather ambiguous ^{1}H NMR spectrum (data not shown).

Notes and references

1. Lu, M., Guo, Q. and Kallenbach, N. R. 1993. *Biochemistry*, **32**, 598–601.
2. Lu, M., Guo, Q. and Kallenbach, N. R. 1993. *Biochemistry*, **32**, 3596–3603.
3. Smirnov, I. and Shafer, R. H. 2000. *Biochemistry*, **39**, 1462–1468.
4. Wang, K. Y., McCurdy, S., Shea, R. G., Swaminathan, S. and Bolton, P. H. 1993. *Biochemistry*, **32**,1899-1904.
5. Smith, F. W. and Feigon, J. 1993. *Biochemistry*, **32**, 8682-8692.
6. Wang, Y. and Patel, D. J. 1993. *Structure*, **1**, 263-282.

1.2 A New Modified Thrombin Binding Aptamer Containing a 5'-5' Inversion of Polarity Site

In this frame, we have undertaken a study whose aim is to use a biological driven approach in order to improve the knowledge of the interactions between thrombin and TBA that are critical for the biological activity. Particularly, we have designed and synthesized a TBA based oligodeoxynucleotide containing a 5'-5' site of polarity inversion, namely $^{3'}$GGT$^{5'}$-$^{5'}$TGGTGTGGTTGG$^{3'}$ (**1**). Here we report a detailed structural study, along with thermal stability analysis and a structure-activity relationship of this new modified TBA (Fig. 1).

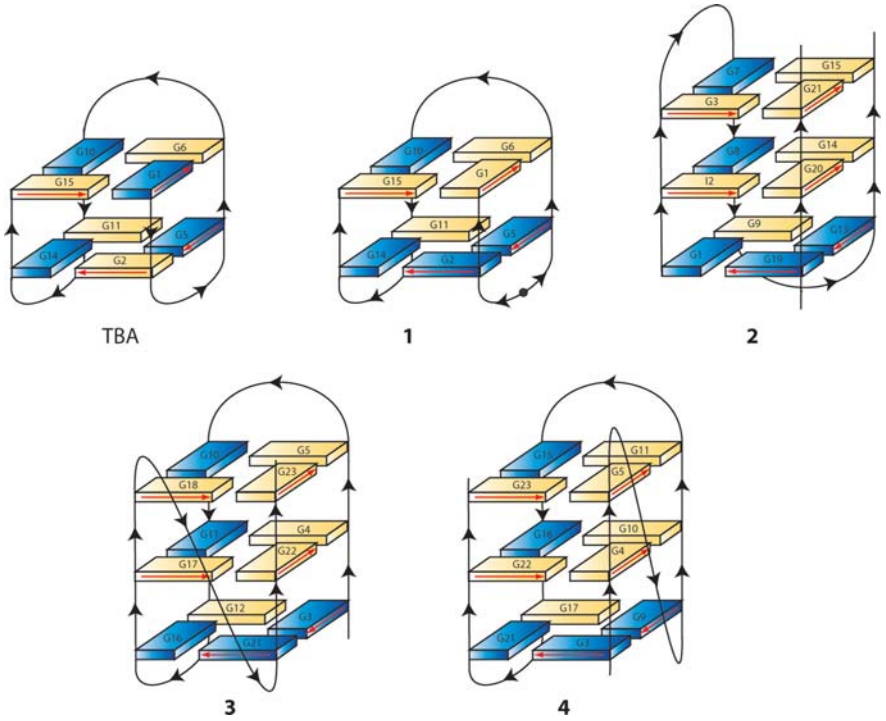

Fig. 1 Schematic illustration of the structures adopted by TBA, modified TBA (1), the three human telomeric repeats sequence into a dimeric quadruplex (11) (2), the four-repeat Tetrahymena $d(T_2G_4)_4$ sequence (12) (3), and the four-repeat human telomeric d[TAGGG(TTAGGG)$_3$] sequence in K^+ solution (13,14). Black arrowheads indicate 5'-3' polarity of the strands. Black circle in 1 represents the 5'–5' inversion of polarity site. Anti and syn guanines are depicted as yellow and blue solids, respectively. Red arrows indicate the direction of the proton donors and acceptors in Hoogsteen hydrogen bonds.

Materials and Methods

The oligonucleotide $^{3'}$GGT$^{5'}$-$^{5'}$TGGTGTGGTTGG$^{3'}$ (1) and the natural counterpart $^{5'}$GGTTGGTGTGGTTGG$^{3'}$ (TBA) were synthesized on a Millipore Cyclon Plus DNA synthesizer, using solid phase β-cyanoethyl phosphoramidite chemistry at 15 μmol scale. The synthesis of 1 was performed by standard methods for the $^{5'}$TGGTGTGGTTGG$^{3'}$ tract, and using 5'-phosphoramidites for the $^{3'}$GGT$^{5'}$ tract. The oligomers were detached from the support and deprotected by treatment with conc. aqueous ammonia at 55°C for 12 h. The combined filtrates and washings were concentrated under reduced pressure, redissolved in H_2O and

analysed and purified by HPLC on a Nucleogel SAX column (Macherey-Nagel, 1000-8/46); using buffer A: 20 mM KH_2PO_4 aq. solution, pH 7.0, containing 20% (v/v) CH_3CN; buffer B: 1 M KCl, 20 mM KH_2PO_4 aqueous solution, pH 7.0, containing 20% (v/v) CH_3CN; a linear gradient from 0 to 100 % B in 30 min and flow rate 1 mL/min were used. The isolated oligomers have the following retention times: 1 = 19.0 min; TBA = 20.5 min. The oligomers were collected and successively desalted by Sep-Pak cartridges (C18). The isolated oligomers resulted to be more than 99% pure (NMR).

Nuclear Magnetic Resonances

NMR samples were prepared at a concentration of approximatively 2 mM, in 0.6 ml (H_2O/D_2O 9:1) buffer solution having 10 mM KH_2PO_4, 70 mM KCl, 0.2 mM EDTA, pH 7.0. For D_2O experiments, the H_2O was replaced with D_2O by drying down the sample, lyophilization and redissolution in D_2O alone. NMR spectra were recorded with Varian ^{Unity}INOVA 500 MHz and 700 MHz spectrometers. ^1H chemical shifts were referenced relative to external sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS), whereas ^{31}P chemical shifts were referenced to external phosphoric acid (H_3PO_4 85% v/v). 1D proton spectra of samples in H_2O were recorded using pulsed-field gradient WATERGATE[1] for H_2O suppression. A proton-detected ^1H-^{31}P heteronuclear COSY was recorded in D_2O in the hypercomplex mode with 2048 t_2 points and 96 t_1 increments, and a spectral width of 500 Hz along the ^{31}P dimension. Phase sensitive NOESY spectra[2] were recorded with mixing times of 100 and 200 ms (T=25°C and 5°C). Pulsed-field gradient WATERGATE was used for NOESY and ^1H-^{15}N HSQC[3] spectra in H_2O. TOCSY spectra[4] with mixing times of 100 ms were recorded with D_2O solutions.

All experiments were recorded using STATES-TPPI[5] procedure for quadrature detection. In all 2D experiments the time domain data consisted of 2048 complex points in t_2 and 400-512 fids in t1 dimension. The relaxation delay was kept at 3 s for NOESY experiments used in the structure determination. A relaxation delay of 1.2 s was used for all other experiments. The NMR data were processed on a SGI Octane workstation using FELIX 98 software (Accelrys, San Diego, CA).

Structural calculations

Cross peak volume integrations were performed with the program FELIX 98, using the NOESY experiment collected at mixing time of 100 ms. The NOE volumes were then converted to distance restraints after they were calibrated using known fixed distances of H2'/H2" of G1, G2, G5, G8, T9, G10, G11, T12, T13 and G14. Then a NOE restraint file was generated with three distance classifications as follows: strong NOEs (1.8 Å $\leq r_{ij} \leq$ 3.0 Å, where 1.8 Å is the van der Waals radius and r_{ij} is the interproton distance between protons i,j), medium NOEs (2.5 Å $\leq r_{ij} \leq$ 4.0 Å) and weak NOEs (3.5 Å $\leq r_{ij} \leq$ 5.5 Å). A total of 198 NOE derived distance restraints were used.

Hydrogen bonds constraints were used: upper and lower distance limits of 2.0 Å and 1.7 Å for hydrogen-acceptor distance, and 3.0 Å and 2.7 Å for donor-acceptor distance, respectively. These constraints for H-bonds did not lead to an increase in residual constraints violation. 54 backbone torsion angles were used in the calculations too. Moreover, according to NMR data, glycosidic torsion angles were kept in a range of -160°/-70° for the G anti, whereas a range of 10°/100° was used for the G syn.

The calculations have been performed using a distance dependant macroscopic dielectric constant of 4*r and an infinite cut-off for

nonbonded interactions to partially compensate for the lack of the solvent[6] have been used. Thus the three-dimensional structures which satisfy NOE and dihedral angle constraints were constructed by simulated annealing calculations. An initial structure of the oligonucleotide was built using a completely random array of atoms. Using the steepest descent followed by quasi-Newton-Raphson method (VA09A) the conformational energy was minimized. Restrained simulations were carried out for 500 ps using the CVFF force field as implemented in Discover software (Accelrys, San Diego, USA). The simulation started at 1000 K, and then the temperature was decreased stepwise until 273 K. The final step was again to energy-minimize to refine the obtained structures, using successively the steepest descent and the quasi-Newton-Raphson (VA09A) algorithms. Both dynamic and mechanic calculations were carried out by using 1 (kcal/mol)/Å^2 flatwell distance restraints. 20 structures were generated. RMSD values of 0.75 ± 0.27 Å and 0.90 ± 0.31 Å for the backbone and heavy atoms, respectively, were calculated for the best 7 structures.

Illustrations of structures were generated using the INSIGHT II program, version 2005 (Accelrys, San Diego, USA). All the calculations were performed on a PC running Linux WS 4.0.

The final set of coordinates has been deposited in the Protein Data Bank (accession code: 2IDN).

Circular Dichroism

CD samples of **1** and TBA were prepared at a concentration of 1 x 10^{-4} M, by using the buffer solution used for NMR experiments: 10 mM KH_2PO_4, 70 mM KCl, 0.2 mM EDTA, pH 7.0. CD spectra and CD melting curves were registered on a Jasco 715 circular dichroism

spectrophotometer in a 0.1 cm pathlength cuvette. For the CD spectra the wavelength was varied from 220 to 320 nm at 100 nm min^{-1}. The spectra were recorded with a response of 16 s at 2.0 nm bandwidth and normalized by subtraction of the background scan with buffer. The temperature was kept constant at 20°C with a thermoelectrically controlled cell holder (Jasco PTC-348). CD melting curves were registered as a function of temperature from 20 to 90°C at 294 nm with a scan rate of 10°C h^{-1} for both quadruplexes. The CD melting curves of TBA and modified aptamer showed sigmoidal profiles and were modeled by a two-state transition, using a theoretical equation for an intramolecular association, according to the van't Hoff analysis.[7] The T_m and ΔH° values provide the best fit of the experimental melting data. The reported errors for thermodynamic parameters are the standard deviations of the mean from the multiple determinations. The ΔS° values were calculated by equation $\Delta S° = \Delta H° / T_m$ and free energy change values by the equation $\Delta G°(T) = \Delta H° - T\Delta S°$.

Prothrombin Time (PT) assay

Human plasma samples were collected by venipuncture, in presence of 0.1 volumes 3.8% sodium citrate and fractionated by centrifugation at 2000g for 5 minutes. PT times were measured by using Koagulab MJ Coagulation System with a specific kit RecombiPlas Tin HemosIL (Inst. Labs, Lexinton, USA). Briefly, this method relies on the high sensitivity thromboplastin reagent based on recombinant human tissue factor. The addition of recombiplastin to the plasma in presence of calcium ions initiates the activation of extrinsic pathway. This results ultimately in the conversion of fibrinogen to fibrin, with a formation of solid gel. The procedure was performed according to the manufacturer's instructions.

TBA, **1** or vehicle (PBS) were added at different time points in volume of 2 μl at final concentration of 20 μM. Data are expressed as mean ± s.e.m. and are representative of at least three different measurements.

Results and Discussion

The NMR sample of 1 was heated for 10 minutes at 80°C and slowly cooled down to room temperature, then its ^{1}H-NMR spectrum was recorded by using pulsed-field gradient WATERGATE[1] for H_2O suppression (Fig. 2).

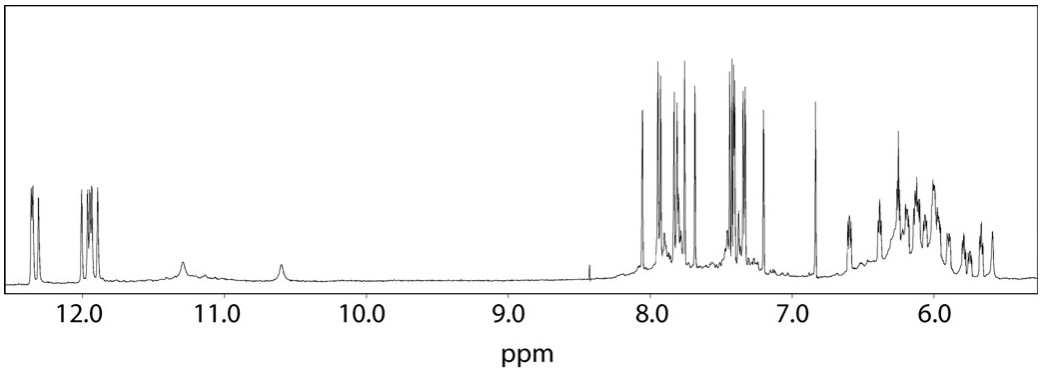

Fig. 2 Expanded region of the proton NMR spectrum of 1 (700 MHz, T=25_C).

With the exclusion of some weak resonances due to very minor conformations also present in solution (whose relative intensities turned out to be insensitive to temperature changes), the simple appearance of one-dimensional spectra of 1 indicates that, in the conditions used here, the modified oligomer forms mainly a single well-defined hydrogen-bonded conformation. In fact, the ^{1}H-NMR spectrum of 1 (700 MHz, T = 25°C) shows the presence of eight well defined signals in the region 11.5-12.5 ppm, attributable to imino protons involved in Hoogsteen hydrogen bonds of at least two G-quartets, and of two broader and less intense singlets at 10.6 and 11.29 ppm, respectively, probably belonging to two imino protons of two thymines involved in mutual hydrogen

bonding (see below). Moreover, fifteen signals in the aromatic region due to the presence of nine guanine H8 and six thymine H6 protons were clearly observable.

A combination of the analysis of 2D NOESY (700 MHz, T = 25°C), TOCSY spectra (700 MHz, T = 25°C), and ^{31}P-NMR spectroscopy (202 MHz, T = 25 °C) allowed us to get the almost complete assignment of both exchangeable and nonexchangeable protons and phosphorus resonances of 1. Particularly, the one-dimensional proton decoupled phosphorus spectrum displays fourteen signals. After assigning the ^{1}H resonances within each deoxyribose by a 2D TOCSY experiment, the 2D proton-detected heteronuclear ^{1}H-^{31}P COSY allowed us to assign the entire backbone correlating each phosphorus resonance to the respective 5'-coupled H3' proton and 3'-coupled H5' and H5" protons of adjacent ribose. Further, NOEs between base protons and H1', H2' and H2" protons allowed us to assign all aromatic protons to the pertinent base.

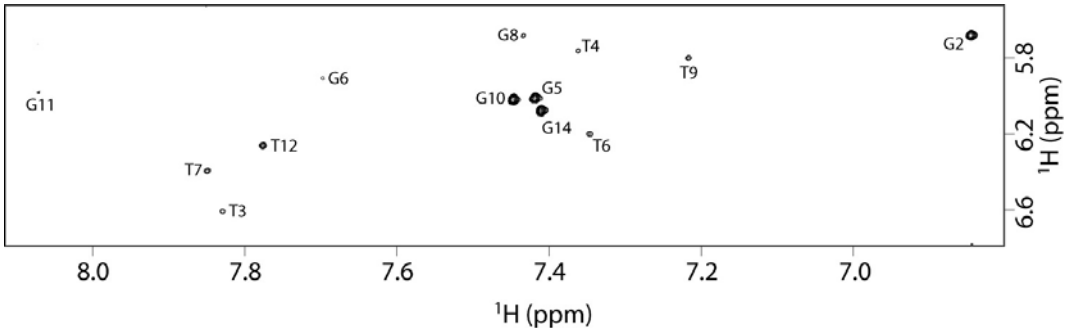

Fig. 3 Expanded region of the NOESY spectrum of 1 (700 MHz, T = 25°C, mt = 200 ms) correlating G–H8 and H1' protons.

It is interesting to note that the intensities NOESY (700 MHz, T=25 °C, mixing time 100 ms) crosspeaks between the H8 proton bases and sugar H1' resonances indicate that four (G2, G5, G10, G14) out of nine Gs of

1 adopt a *syn* glycosidic conformation (Fig. 3), where the H8 resonances of *syn* G residues are upfield shifted with respect to those of the *anti* ones.[8-9]

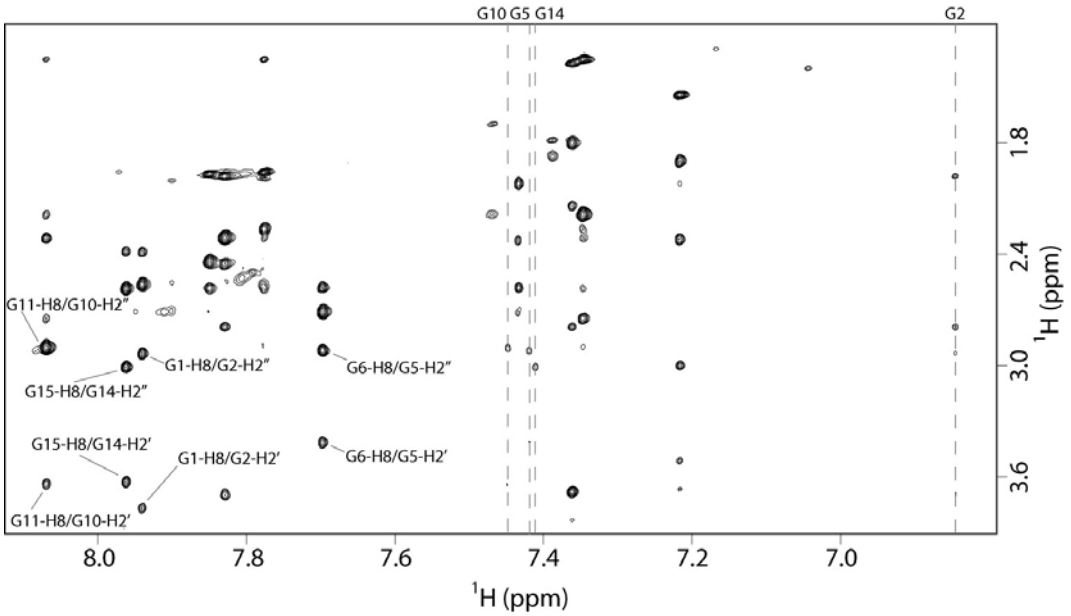

Fig. 4 Expanded region of NOESY spectrum of 1 (700 MHz, T = 25_C, mt = 100 ms) correlating G–H8 and H2'-H2" protons. Classical G–H8/H2'–H2" NOEs (nuclear Overhauser effect) correlations for the four-stranded part of the complex are reported. Pattern of NOEs for the Gs in syn glycosidic conformation are highlighted by vertical dashed lines.

Then, four *anti*-Gs (G1, G6, G11 and G15) have classical H8/H2'-H2" sequential connectivities to 5' neighboring *syn*-Gs (G2, G5, G10 and G14, respectively) (Fig. 4), indicating that the subunits G1-G2, G5-G6, G10-G11 and G14-G15 are involved in the formation of a four-stranded helical structure (underlined residues adopt a *syn* glycosidic conformation).

Moreover, the entire pattern of NOEs observed for all cited Gs indicates that the backbone conformations of these tracts resemble those of the unmodified TBA possessing a right-handed helix structure (Fig. 4).

The alternation of *syn* and *anti* G residues within each strand suggests that, as TBA, **1** folds into a monomolecular foldback quadruplex, characterized by two G tetrads. Further, a number of unusual NOE connectivities were observed between *syn*-Gs and Ts, indicating that $^{5'}$TG$^{3'}$ and $^{5'}$GT$^{3'}$ tracts do not adopt a helical winding, thus suggesting that, most probably, the TT and TGT tracts form loops.

In order to assign exchangeable protons to their respective G or T residues, a ^{1}H-^{15}N HSQC experiment (700 MHz, T= 5°C) has been acquired. Eight exchangeable protons (the eight at lower field) correlate with ^{15}N in the region between 143 and 147 ppm indicating that they are attributable to imino protons of G residues,[10] whereas the signals at 10.65 and 11.35 ppm correlate with nitrogens at 155.7 and 155.6 ppm, respectively, thus indicating that the latter belong to T residues.[10] Interestingly these two signals are characterized by strong NOE effects between each other, thus suggesting that two thymines are in close proximity and, verisimilarly, interacting by a mutual hydrogen bond, exactly as observed in the unmodified TBA. Moreover, NOEs between H1' proton of G8, methyl protons of T9 and three G imino protons (δ_H 12.03, 11.97 and 11.91 ppm), and NOEs between other four imino protons (δ_H 12.37, 12.35, 12.30 and 11.95 ppm) and H6, methyl, H3' and H1' protons of T4 and H3', H2', H4' and H1' of T13, led us to assign the exchangeable protons of two tetrads

The whole of data indicates that the structure of **1** is a foldback quadruplex characterized by two G tetrads formed by the residues G1, G6, G10, G15 and G2, G5, G11, G14, respectively, where the underlined residues possess a *syn* glycosidic conformation. Interestingly the two tetrads assume a *anti-anti-anti-syn* and *syn-syn-syn-anti* arrangement of the bases. It is noteworthy that even the T imino protons (δ_H 10.65

and 11.35 ppm, T=5°C) are characterized by several NOE connectivities with a set of G imino protons belonging to only one tetrad (δ_H 11.95, 12.80, 12.35 and 12.37 ppm). Moreover, the pattern of NOEs attributable to the loop T12-T13 and T3-T4 let us to identify the T4 and T13 as the residues that are H-bonded to each other. Interestingly, T13 and T4 are mutually hydrogen bonded in TBA as well. Then, the facts that TGT loop possesses some NOEs with G imino protons of one tetrad and that the two TT loops are characterized by NOEs with G imino protons of the other tetrad indicate that no loop assumes a dog-eared conformation. All this means that the **1** folds into a chair-like quadruplex structure possessing three strands parallel to each other and only one strand oriented in reverse manner (Fig. 1).

In order to get the three-dimensional structure of **1** at atomic level, an estimation of proton-proton distances has been retrieved from cross-peak intensities in 2D NOESY experiments acquired at 700 MHz, both at T=25°C and T=5°C (mixing time 100 ms for the experiment acquired in D_2O and 200 ms for the experiments acquired in H_2O). Pseudo-atoms were introduced where needed. 198 distances were used for the calculations and, as suggested by the presence of eight G and two T imino protons in the 1D ^{1}H-NMR spectrum, 32 supplementary distance restraints (HN1-O6, N1-O6, HN2-N7, N2-N7) for 16 hydrogen bonds corresponding to the two G-quartets, and 4 distance restraints (HN3-O4, N3-O4) for 2 hydrogen bonds between T13 and T4 were also incorporated during the computations (Table 1). The backbone torsion angles β and ε were estimated from the scalar coupling measured in the 2D proton-detected heteronuclear ^{1}H-^{31}P COSY by using the semi-empirical Karplus equation.[11] Thus, since for all sugars in **1** all $J_{P,H5'}$ and $J_{P,H5''}$ were < 8 Hz, β backbone torsion angles were restricted to 180 ±

30°.[11] For residues 3, 4, 5, 6, 7, 8, 9, and 11, $J_{P,H3'}$ was larger than 8 Hz, and the corresponding ε angles were restricted to -120± 30°.[10] Sugar ring conformations were obtained from the analysis of the PE-COSY spectrum. From the $\Sigma J_{H4'}$ estimation (< 10Hz), no dihedral restraints for γ were applied to the four G nucleotides in the *syn* glycosidic conformation, whereas, γ backbone torsion angles for residues G6 and G11 (possessing an *anti* conformation) were restricted to 60± 30°.[10] All measurable $J_{H1',H2'}$ were reasonably large, indicating a predominant S-type nature of sugar ring conformations: δ and ν_2 angle constraints have been added consistently. Further, in agreement with NMR data, glycosidic torsion angles for four out of eight guanines involved in the formation of the two G-tetrads were fixed in the *anti*-domain (-160°/-70°), whereas the χ angle was kept in a range of 10°/100° (*syn*-conformation) for the remaining four G residues.

Therefore, three-dimensional structures which satisfy NOEs were constructed by simulated annealing (SA) calculations. An initial structure of the oligonucleotide was constructed possessing a random conformation and minimized, in order to eliminate any possible source of initial bias in the folding pathway. Restrained simulations were carried out for 500 ps using the CVFF force field. The restrained SA calculations started at 1000 K, and, thereafter, the temperature was decreased stepwise down to 273 K. The aim step was to energy-minimize and refine the structures. A total of 20 structures was generated. Average RMSD values of 0.75 ± 0.27 Å and 0.90 ± 0.31 Å for the backbone and all heavy atoms, respectively, were obtained from the superimposition of the best 7 structures (Fig. 5). These data, along with the lack of significant violations of the experimental restraints (see Table 1), suggest

that the obtained structures are representative of the structure actually adopted in solution by **1**.

Table 1 Experimental constraints and structure statistics of the best 20 structures

Experimental constraints	
Total NOEs	198
NOEs from non-exchangeable protons	185
NOEs from exchangeable protons	13
Hydrogen bonds constraints	36
Dihedral angle constraints	62
CVFF energy (kcal mol^{-1}) of the minimized structures	
Total	370.158 ± 4.529
nonbond	40.783 ± 5.145
restraint	13.213 ± 1.417
NOEs violations	
Number > 0.2 Å	3.40 ± 1.63
Maximum (Å)	0.560 ± 0.263
Sum (Å)	3.821 ± 0.693
Average violation (Å)	0.016
r.m.s. deviations from the mean structure (Å)	
All backbone heavy atoms	0.75 ± 0.27
All heavy atoms	0.90 ± 0.31

As expected, **1** shows a right-handed helical backbone geometry and three edge-wise connecting TT, TGT, and TT loops. As the unmodified TBA, **1** involves two stacked G-tetrads, but differs in guanine *syn*/*anti* distribution around the GGGG tetrads, being characterized by the unusual *anti-anti-anti-syn* and *syn-syn-syn-anti* arrangement of the two tetrads respectively, and in the relative strand orientations, with three strands parallel to each other and with only one oriented in opposite manner. Interestingly, a similar strand orientation and glycosidic angle arrangement of the bases of the tetrads have already been observed for the assembly of three human telomeric repeats into a dimeric G-quadruplex[12] (**2**), the four-repeat *Tetrahymena* $d(T_2G_4)_4$ sequence[13] (**3**), and the four-repeat human telomeric d(TAGGG(TTAGGG)$_3$] sequence in K^+ solution[14-15] (**4**) (Fig. 1).

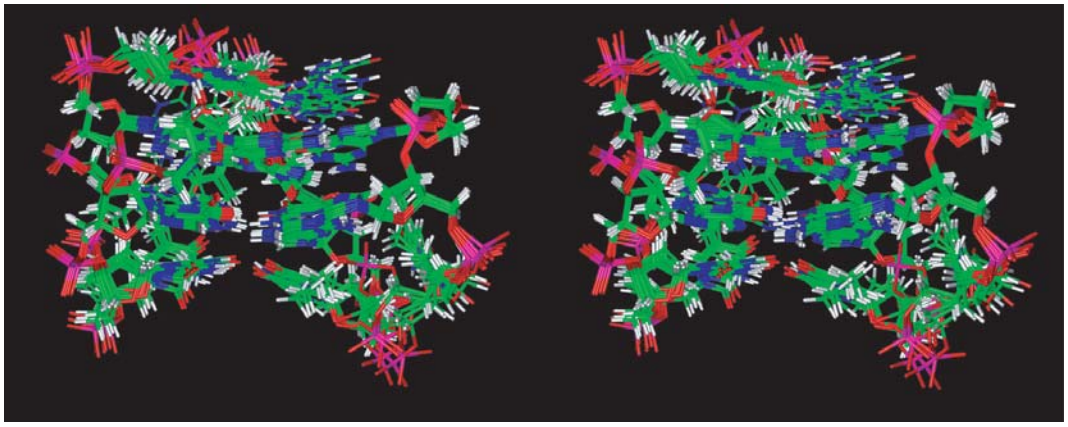

Fig. 5 Stereoview representation of the superimposed seven best structures of 1. Backbones and bases are depicted in colored "stick" (carbons, green; nitrogens, blue; oxygens, red; hydrogens, white).

Even though sequences **2**-**4** differ to each other in both molecularity and folding topology, when their structures are represented in such a way that the four strands possess the same orientation (Fig. 1), a certain resemblance among the three models is clearly apparent. Particularly, all the corresponding guanines are characterized by the same glycosidic angle arrangements. It is noteworthy that the only two tetrads in the modified TBA (**1**) show the same glycosidic angle arrangements as those in the lower two planes in **2**-**4**. A further structural element joining the four complexes is the presence in all of them of one narrow, one wide and two medium width grooves.

Furthermore, a comparison of the most representative structures of **1** (the one with lowest energy after minimization), and the already reported NMR structure of the unmodified TBA (PDB code: 148D) has been accomplished by analyzing the helix parameters by CURVES[16-17] (Table 2 and 3).

Table 2 Rise, tilt, roll and twist of TBA and of the best-minimized structure of 1

TBA	Rise	Tilt	Roll	Twist	1
G1/G2	3.44	−10.60	−179.75	50.93	
	3.25	7.84	160.55	−45.79	G1/G2
G5/G6	2.67	6.22	−161.77	48.90	
	3.64	−1.25	−178.12	−49.88	G5/G6
G10/G11	2.80	10.45	−178.83	48.50	
	3.92	−4.04	151.25	−52.77	G10/G11
G14/G15	4.67	−17.19	−177.35	55.65	
	3.24	−12.25	−156.98	−49.22	G14/G15

Table 3 Shear, stretch, stagger, buckle, propeller, opening of TBA and of the best-minimized structure of 1

TBA	Shear	Stretch	Stagger	Buckle	Propeller	Opening	1
G1/G6	−6.15	1.72	−0.88	11.32	176.45	−89.48	
	−6.67	2.57	−0.38	3.51	174.13	91.17	G1/G6
G1/G10	4.82	2.32	−2.69	3.47	−176.25	−0.95	
	−6.59	2.09	−0.71	38.43	164.49	6.02	G1/G10
G1/G15	−6.02	1.87	−0.59	−5.12	−162.35	86.32	
	−6.95	2.01	−0.37	28.42	−172.46	−86.37	G1/G15
G2/G5	6.42	2.77	−0.11	−5.50	157.97	−91.51	
	6.23	2.08	0.02	−5.58	−168.29	87.08	G2/G5
G2/G11	−6.04	1.75	−2.05	24.52	−177.68	−3.39	
	7.00	2.51	−0.04	26.55	−147.31	−0.96	G2/G11
G2/G14	5.41	2.07	−1.82	−5.12	−162.35	86.32	
	6.59	1.88	−0.37	8.34	−176.04	−89.79	G2/G14

These data clearly suggest that the overall structure adopted by **1** is similar to that of TBA. The base stacking of the two quadruplexes is quite similar, being the five-membered rings of the top guanines overlapped to the five-membered rings of the underneath guanine bases (Fig. 6). Furthermore, as suggested by *Tilt* and *Stagger* values, the stacking between G pairs is slightly better in **1** than in TBA. The groove widths for **1** are just a little bit different from that of unmodified TBA. Particularly, TBA is characterized by two wide grooves (ca. 19 Å) between the strands G5-G6 and G10-G11 and between the strands G1-G2 and G14-G15 respectively, and other two narrow grooves that are calculated to be ca. 13 Å. On the other hand, **1** possesses the groove between the strands G1-G2 and G5-G6 just a little bit wider (ca. 16 Å)

and the groove between the strands G1-G2 and G14-G15 just a little bit narrower (ca. 15 Å) (Fig.6).

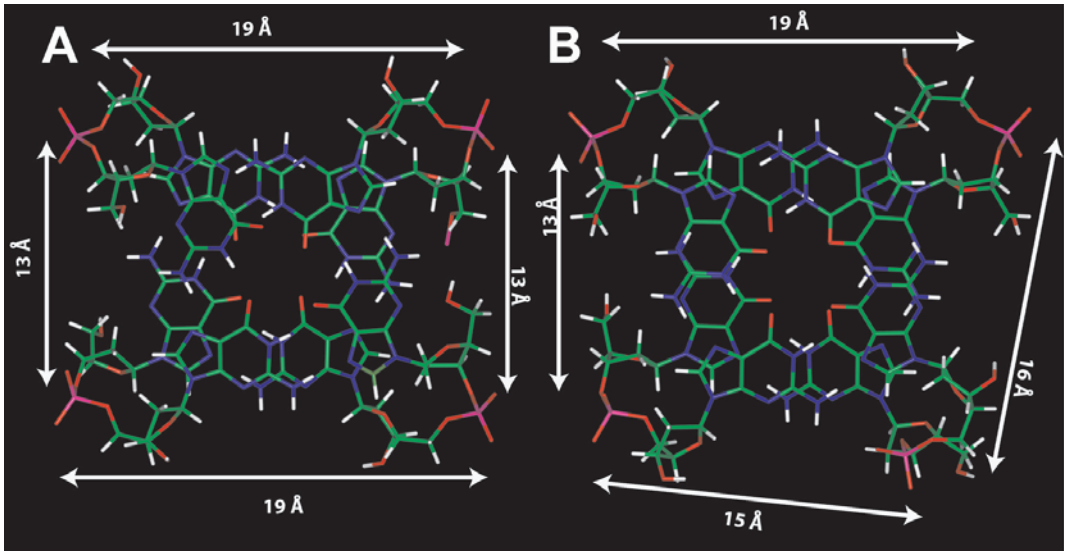

Fig. 6 Top view of the two G-tetrads of TBA (A) and 1 (B). Backbones and bases are depicted in colored 'stick' (carbons, green; nitrogens, blue; oxygens, red; hydrogens, white). White arrows indicate groove widths.

As for the sugar puckering, the two molecules possess the same sugar ring conformations, being in S-type conformation. It is interesting to note that the main structural difference rely on the orientation of T3. In fact, as indicated by the diagnostic NOE connectivities between T3-Me and G2-H8, and between T3-H6 and G2-H8, the base of T3 folds back into the groove between the strands G1-G2 and G5-G6 in **1**, and faces closely the G2-H8, thus justifying the very low chemical shift value of T3-H6 (δ_H 6.84) (Fig. 7). Furthermore, another slight structural difference is to be ascribed to the orientation of T7. In fact, T7 folds back into the groove formed by the strands G5-G6 and G10-G11. Actually, this latter structural change seems due to a different folding of the backbone of this part of the molecule, where the sugar of T7 is closer to the bases G8 and G6 than the unmodified TBA (Fig. 7).

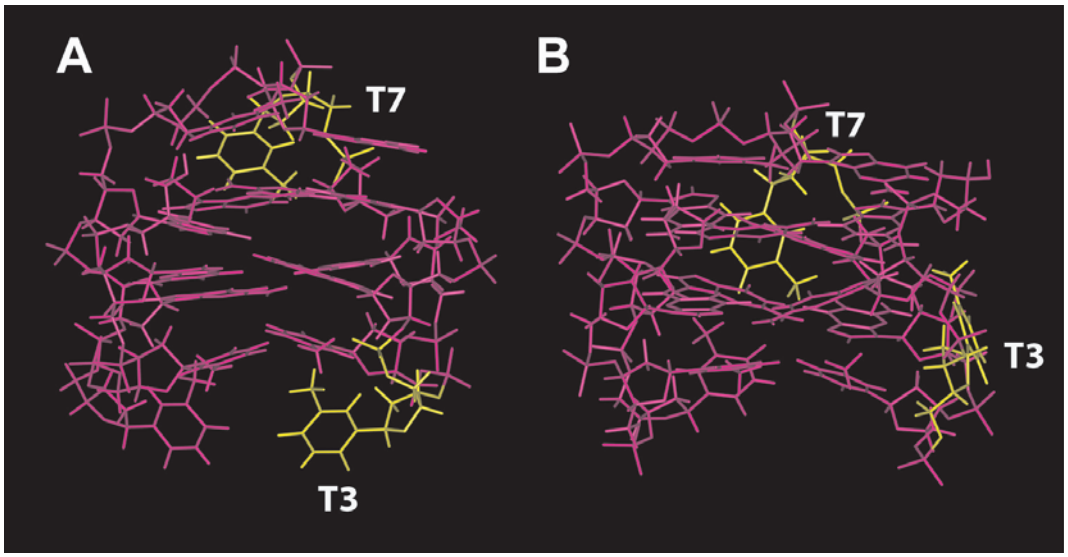

Fig. 7 Side view of TBA (A) and 1 (B). All residues are depicted in magenta, except T3 and T7 (yellow). The different orientations of T3 and T7 are clearly observable.

In order to determine the effects of the inversion of polarity site on the CD profile and the thermal stability of the resulting quadruplex structure, CD spectra and CD melting and annealing experiments were acquired for **1** and its natural counterpart (Fig. 8).

In particular, the CD spectra of **1** and TBA, performed at 20°C, are almost superimposable, both exhibiting two positive bands at 248 and 294 nm and a negative one at 267 nm. As expected, the TBA spectrum is typical of anti-parallel quadruplex structures.[18-20] The spectrum of **1** is also typical of antiparallel quadruplex structures because **1** possesses two G-tetrads with an alternating *anti* and *syn* glycosidic conformation along each strand. However, this spectrum differs from CD spectra previously reported for the same "3+1" strand arrangements, where the presence of an additional peak at 260 nm was attributed to the presence of third G-tetrad that implies a non-alternating guanine glycosidic conformation between the middle and bottom G-tetrads.

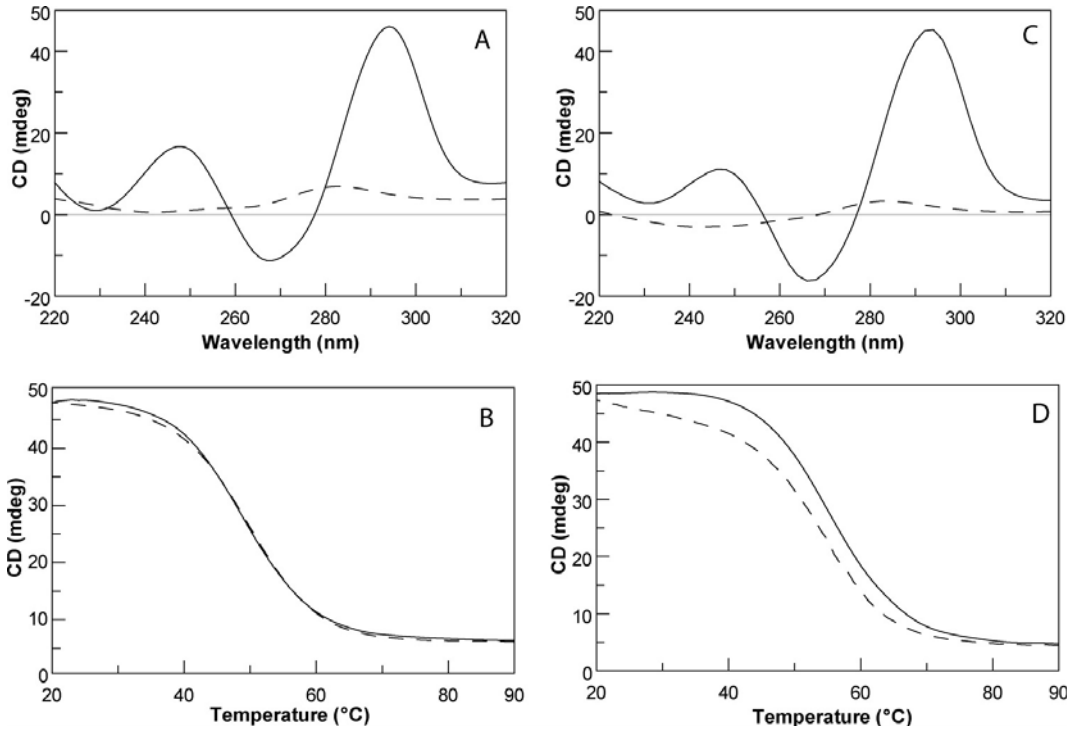

Fig. 8 CD spectra of TBA (A) and 1 (C) at 20°C (continuous lines) and 90°C (dashed lines). Melting (continuous lines) and annealing (dashed lines) CD experiments of TBA (B) and 1 (D).

Table 4 Thermodynamic parameters for the unfolding of 1 and TBA

	T_m (°C) (±1)	$\Delta H°$ (kJ mol^{-1}) (±20)	$\Delta S°$ (kJ mol^{-1} K^{-1}) (±0.05)	$\Delta G°$(298 K) (kJ mol^{-1}) (±1)
1	57	180	0.54	17
TBA	52	160	0.49	13

As for melting and annealing CD experiments, taking into account that the rates of quadruplex formation/dissociation are very slow, we collected the data at 10°C/h. No significant hysteresis emerged for each oligonucleotide comparing annealing and melting curves thereby indicating that, at the scan rate used, both systems were at equilibrium. The melting curves were analysed using van't Hoff analysis and thermodynamic parameters are shown in Table 4. The values for TBA are in good agreement with those recently obtained by UV technique.[21-22] The transition for the modified aptamer occurs 5 degrees above the

transition of TBA and the enthalpy change values are consistent with the opening of two G-tetrads.[23-24] The entropy values suggest that the modified aptamer possesses a more rigid structure respect to TBA, according to a better stacking between G pairs found in the NMR study. The whole set of thermodynamic parameters shows that the quadruplex formed by **1** is more stable than its unmodified counterpart.

In order to study the effects of the structural changes on the biological activity of **1**, if any, we performed a prothrombin time (PT) assay. PT assay on **1** was performed on human plasma in strict comparison with TBA. In order to eliminate the variation caused by the measurements performed on different days, the samples of **1** and TBA were prepared at the same time and folded together by heating the samples for 10 minutes at 80°C and slowly cooling them down at room temperature.

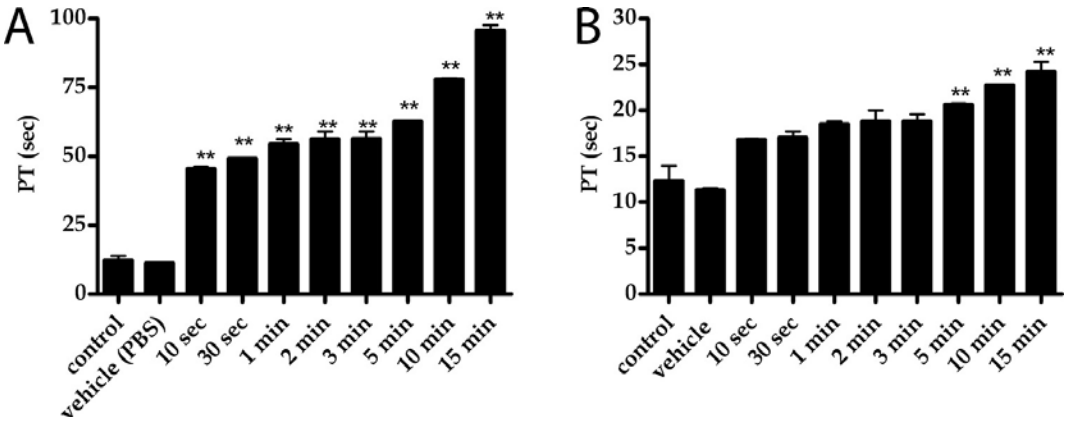

Fig. 9 Prothrombin time assessed by using human plasma. On the x-axis the incubation time is reported, while PT time in seconds is reported on y-axis. (A) refers to TBA, while (B) to 1.

The assays have been conducted after one week from the preparation procedure. As shown in figure 9, overall, TBA showed a statistically significant more marked effect of **1** both as onset of the activity and endurance. Indeed, incubation studies demonstrated that TBA effect is already present after 10 seconds of incubation and lasts up to 15 minutes

of incubation. On the other hand, **1** showed a slower onset of the effect, indeed PT time started to be significantly prolonged following 5 minutes of incubation. Also the magnitude of the activity of **1** resulted reduced compared to TBA. The biological data reported here imply that the modifications of the structure adopted by **1** affect sensibly the biological activity, and therefore its interaction with the thrombin. Looking to the overall folding of the modified aptamer, it was not expected such reduction in the biological activity. However, it is interesting to note that the TBA inhibits the thrombin activity when interacting with its fibrinogen exosite.[25] In particular, it has been proposed the TBA could associate with thrombin in two different ways: through the T7-G8-T9 loop (as suggested by X-ray data) or through the T3 and T12 residues (as suggested by NMR models).[25] As far as the X-ray data are concerned, among all the observed interactions, T7 is buried in a pocket of hydrophobic residues of the fibrinogen exosite of thrombin made by Ile24, His71, Ile79 and Tyr117, whereas, according to NMR data, the T3 residue interacts with the thrombin residues Tyr76 and Ile82. Therefore, the reduced biological activity of **1** could be ascribed to the different orientation adopted by T3, and, at a less extent, by T7 in **1** with respect of the unmodified TBA (Fig. 7). The longer onset displayed by **1** could be tentatively explained considering that **1** needs to properly orientate T3 and T7 upon its binding to thrombin. Nevertheless, the anticoagulant activity of **1** could not be simply ascribed to these changes, but could even due to a combination of all structural changes observed such as: i) relative strand orientation; ii) glycosidic arrangements of the bases iii) groove widths.

The present data give a further insight into the understanding of the variables involved in the formation and the stability of quadruplex

structures. In fact, **1** clearly prefers to adopt a "3+1" strand orientation instead of a classical "2+2" typical of this kind of structure (see TBA, Fig. 1), in spite of both arrangements are equally possible. Thus, these findings further enforce the idea that the "3+1" core topology has not to be considered an anomaly, but it is rather a robust G-quadruplex scaffold. The structure reported here represents also the first example of a "3+1" core topology quadruplex structure containing only two G-tetrads. Furthermore, our findings may contribute to the comprehension of the folding of telomeric DNA, and could shed more light on the understanding of the thrombin/TBA interaction phenomenon. Finally, considering that the "3+1" arrangement increases the thermal stability of the quadruplex and that **1** is not devoid of biological activity, the scaffold described in the present paper can be seen as a unique platform for structure-base anticoagulant drug design. Further investigations concerning ODNs containing either 3'-3' or 5'-5' inversion of polarity sites and their thermodynamic study are in progress in our laboratories.

Notes and references

1. Piotto,M., Saudek,V. and Sklenar,V.J. (1992). *J. Biomol. NMR*, **2**, 661-665.
2. Jeener, J., Meier, B., Bachmann, H.P. and Ernst, R. R. (1979) *J. Chem. Phys.,* **71**, 4546-4553.
3. Sklenar, V., Piotto, M., Leppik, R. and Saudek, V. (1993) *J. Magn. Reson.*, **102**, 241-245.
4. Braunschweiler, L. and Ernst, R. R. (1983), *J. Magn. Reson.*, **53**, 521-528.
5. Marion,D., Ikura,M., Tschudin,R. and Bax,A. (1989) *J. Magn. Reson.*, 85, 393–399.

6. Weiner,S.J., Kollman,P.A., Case,D.A., Singh,U.C., Ghio,C., Alagona,G., Profeta,S., Weiner,P.J. (1984) *J. Am. Chem. Soc.*, **106**, 765-784.
7. Marky, L.A., Breslauer, K.J. (1987) *Biopolymers,* **26** 1601-1620.
8. Smith,F.W., Feigon,J. (1993) *Biochemistry*, **32**, 8682-8692.
9. Wang,Y., Patel,D.J., (1993) *Structure*, **1**, 263-282.
10. Fernandez,C., Szyperski,T., Ono,A., Iwai,H., Tate,S., Kainosho,M., Wuthrich,K. (1998) *J. Biomol. NMR*, **12(1)**, 25-37.
11. Kim,S, Lin,L., Reid,B.R. (1992), *Biochemistry*, **31**, 3564-3574.
12. Zhang,N., Phan,A.T., Patel,D.J. (2005) (*J. Am. Chem. Soc.*, **127**, 17277-17285.
13. Wang,Y. and Patel,D.J. (1994) *Structure*, **2**, 1141-1155.
14. Luu,K.N., Phan,A.T., Kuryavyi,V., Lacroix,L., Patel,D.J. (2006) *J. Am. Chem. Soc.*, **128**, in press
15. Ambrus,A., Chen,D., Dai,J., Bialis,T., Jones,R.A., Yang,D. (2006) *Nucleic Acids Res*, **34**, 2723-2735.
16. Lavery,R. and Sklenar,H. (1988) *J. Biomol. Struct. Dyn.*, **6**, 63-91.
17. Lavery,R. and Sklenar,H. (1989). *J. Biomol. Struct. Dyn.*, **6**,655-667.
18. Lu,M., Guo,Q. and Kallenbach,N.R. (1993) *Biochemistry*, **32**, 598–601.
19. Lu,M., Guo,Q. and Kallenbach,N.R. (1993) *Biochemistry*, **32**, 3596–3603.
20. Smirnov,I. and Shafer,R.H. (2000) *Biochemistry*, **39**, 1462–1468.
21. Miyoshi, D., Karimata, H. and Sugimoto, N. (2006) *J. Am. Chem. Soc.*, **128**, 7957-7963.
22. Olsen,C.M., Gmeiner,W.H., and Marky, L.A., (2006) *J. Phys. Chem B*, **110**, 6962-6969.
23. Shafer, R.H., (1998) *Prog. Nucleic Acids Res Mol Biol*, **59**, 55-94.

24. Petraccone, L., Erra, E., Nasti, L., Galeone, A., Randazzo, A., Mayol, L., Barone, G., Giancola, C., (2002) *J Biol Macromol*, **31**, 131-137.
25. Padmanachan,K., Tulinsky,A. (1996) *Acta Cryst.*, **D52**, 272-282.

2. Thrombin Binding Aptamer: base modifications

The prospect of preparing therapeutically active analogues of natural nucleic acids has stimulated much interest to improve their characteristics, such as hybridization affinity, stability towards cellular nucleases and the ability to penetrate the cell membrane.[1]

In the last two decades a variety of modified oligonucleotides have been synthesized and they are used now in biophysical and biochemical studies.[2] In 1998 the first oligonucleotides containing one or more 2'-O, 4'-C methylene-linked bicyclic ribonucleosides (LNA, Locked Nucleic Acid) were prepared by the groups of Wengel[3,4] and Imanishi.[5] They also described the unprecedented hybridization affinity of LNAs towards complementary nucleic acids. In LNA, the furanose conformation is chemically locked in a C3'-endo (N-type) conformation by the introduction of a 2'-O, 4'-C methylene linkage. LNAs have shown unprecedented thermal affinities when hybridized with either DNA (ΔT_m=1-8°C per modification), RNA (ΔT_m=2-10°C per modification) or LNA (ΔT_m > 5°C per modification).[6-9]

The unprecedented hybridization properties suggest that LNAs could be used as powerful agents for fine tuning drugs with a very specific target potential, thus providing a new class of therapeutics.[7] Actually, LNAs represent very versatile tools for the control of gene expression, the treatment of various human diseases and diagnostic assays. In fact, chimeric 2'-O-methyl/LNA oligoribonucleotides were found to inhibit RNA-protein interactions important for HIV replication by sterically blocking the trans-activation responsive region TAR.[10] Furthermore, incorporation of LNA monomers into the binding arms of the '10-23'

DNAzyme, yielding an LNAzyme, markedly increases cleavage properties towards the target RNA.[11]

2.1 A novel thrombin binding aptamer containing a G-LNA residue

Although TBA was the first aptamer to be discovered and considering the advantageous characteristics of LNAs,[4] no studies dealing with the effects of substitution of a regular nucleotide with an LNA residue in the TBA sequence have been reported yet.

In light of the above, in this work we report here the study of four new TBA based oligonucleotides containing LNA residues, namely $^{5'}$ggttggtgtggttgg$^{3'}$ (1), $^{5'}$ggT̲TggTGTggT̲Tgg$^{3'}$ (2), $^{5'}$gGT̲TGGTGTGGT̲TGG$^{3'}$ (3) and $^{5'}$GGT̲TGGTGTGGT̲TGg$^{3'}$ (4), where upper case and lower case letters represent DNA and LNA residues, respectively. Moreover, the chemical-physical properties and the three-dimensional characterization, based on NMR and CD spectroscopy, associated with molecular mechanics and dynamics calculations, of 4 are also reported.

Results and Discussion

The synthesis of the oligoribonucleotides $^{5'}$ggttggtgtggttgg$^{3'}$ (1), $^{5'}$ggT̲TggTGTggT̲Tgg$^{3'}$ (2), $^{5'}$gGT̲TGGTGTGGT̲TGG$^{3'}$ (3) and $^{5'}$GGT̲TGGTGTGGT̲TGg$^{3'}$ (4), were performed by standard methods and the incorporation of modified residues was performed using g-LNA or t-LNA phosphoramidites. The NMR samples of 1, 2, 3 and 4 were prepared at a concentration of 2.0 mM (0.6 ml, 90% H_2O / 10% D_2O)

in 10 mM KH_2PO_4 buffer containing 70 mM KCl, 0.2 mM EDTA (pH 7.0). The samples were heated for 10 minutes at 80°C and slowly cooled down to room temperature, then their ^{1}H-NMR spectra were recorded by using pulsed-field gradient WATERGATE[27] for H_2O suppression. The lack of immino signals in the 1D ^{1}H-NMR spectrum of the oligomer containing all LNA residues (1) indicates that, in the conditions used here, it is unstructured, most probably due to the less flexibility of the molecule now containing LNA residues. On the other hand, 2 and 3 give NMR spectra characteristic of mixtures of different structures. As fas as 1, 2 and 3 are concerned, neither varying buffer solutions (using alternatively potassium and sodium buffers and changing the concentrations of both KCl and NaCl), nor varying the temperature could improve the quality of their spectra.

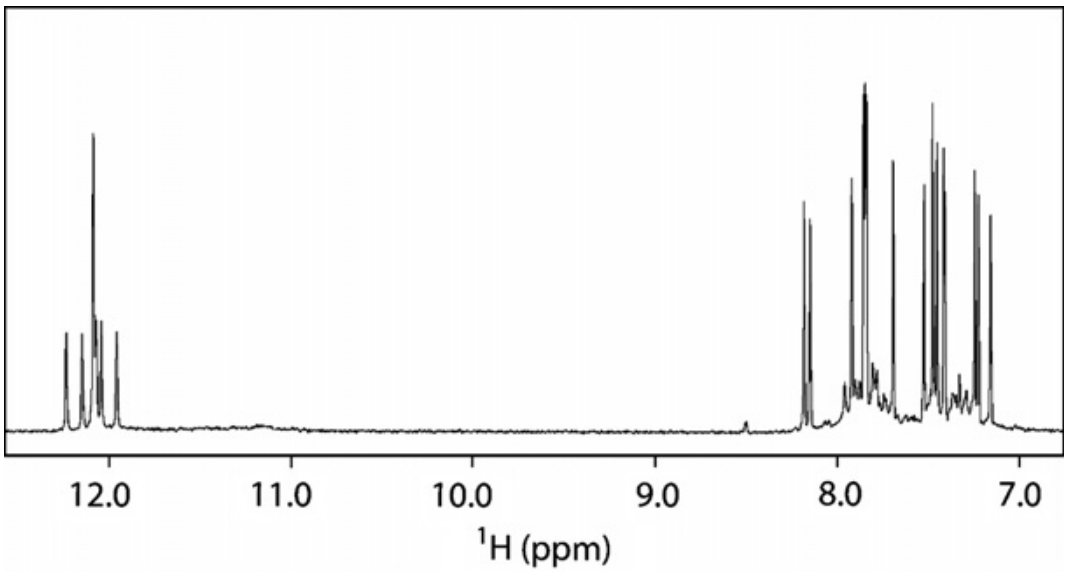

Fig. 1 Expanded region of the proton NMR spectrum of 4 (700 MHz, T = 25°C).

The oligonucleotides $^{5'}$GGTTGGTGTGGTTGg$^{3'}$ (4), instead, turned out to be the only analyzed oligonucletide capable to form a single well-defined hydrogen-bonded structure in solution. In fact, with the exclusion of some weak resonances due to very minor conformations

also present in solution (whose relative intensities turned out to be insensitive to temperature changes), only fifteen signals attributable to nine guanine H8 and six thymine H6 protons were clearly observable in the aromatic region (Fig. 1).

A combination of the analysis of 2D NOESY (700 MHz, T = 25°C), TOCSY spectra (700 MHz, T = 25°C), and ^{31}P-NMR spectroscopy (202 MHz, T = 25 °C) allowed us to get the almost complete assignment (Table 1) of both exchangeable and nonexchangeable protons, and phosphorus resonances of 4. Particularly, the one-dimensional proton decoupled phosphorus spectrum displays fourteen signals. After assigning the ^{1}H resonances within each deoxyribose by a 2D TOCSY experiment, the 2D proton-detected heteronuclear ^{1}H-^{31}P COSY allowed us to assign the entire backbone correlating each phosphorus resonance to the respective 5'-coupled H3' proton and 3'-coupled H5' and H5" protons of adjacent ribose. Further, NOEs between base protons and H1', H2' and H2" protons allowed us to assign all aromatic protons to the pertinent base.

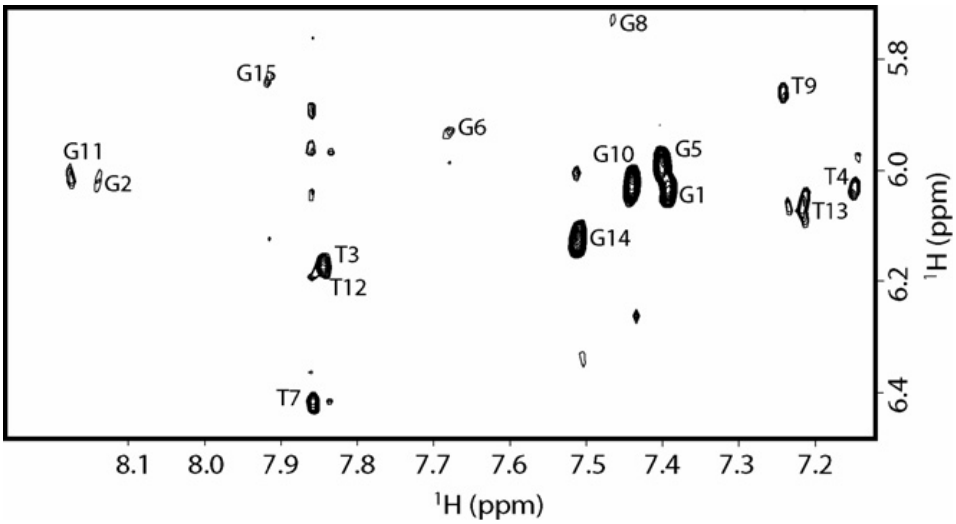

Fig. 2 Expanded region of the NOESY spectrum of 4 (700 MHz, T=25 °C, mt = 200 ms) correlating G–H8/H10 protons.

It is interesting to note that the intensities NOESY (700 MHz, T=25 °C, mixing time 100 ms) crosspeaks between the H8 proton bases and sugar H1' resonances indicate that four (G1, G5, G10, G14) out of nine Gs of 4 adopt a *syn* glycosidic conformation (Fig. 2), where the H8 resonances of *syn* G residues are upfield shifted with respect to those of the *anti* ones.[20,26-27]

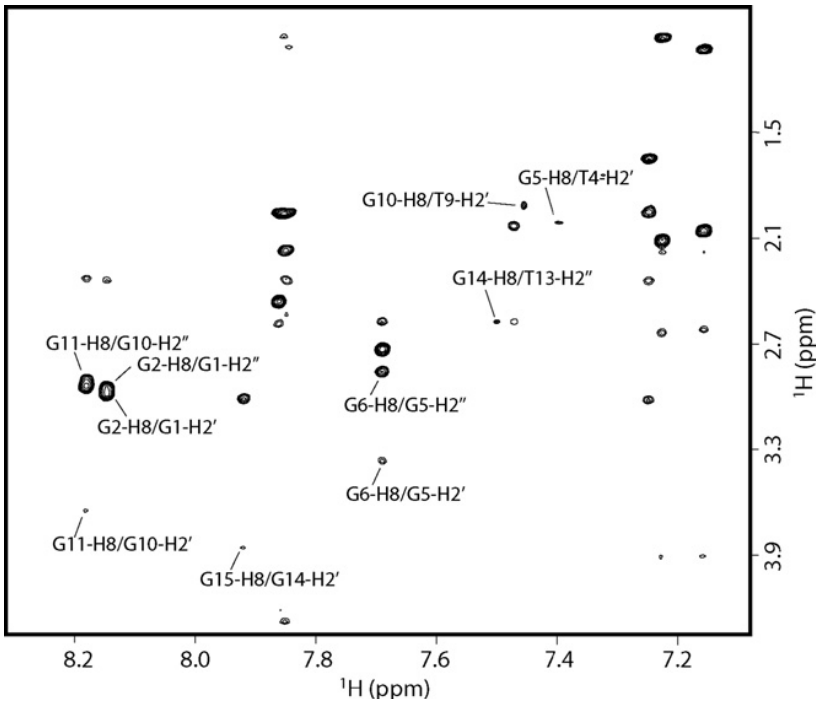

Fig 3 Expanded region of NOESY spectrum of 4 (700 MHz, T = 25°C, mt = 100 ms) correlating G–H8/H2'–H2" protons. Classical G–H8/H2'–H2" NOEs correlations for the four-stranded part of the complex are reported

Then, four *anti*-Gs (G2, G6, G11 and G15) have classical H8/H2'-H2" sequential connectivities to 5' neighboring *syn*-Gs (G1, G5, G10 and G14, respectively) (Fig. 3), indicating that the subunits G1-G2, G5-G6, G10-G11 and G14-G15 are involved in the formation of a four-stranded helical structure (underlined residues adopt a *syn* glycosidic conformation). Moreover, the entire pattern of NOEs observed for all cited Gs indicates that the backbone conformations of these tracts

resemble those of the unmodified TBA possessing a right-handed helix structure (Fig. 3).

Table 1 Proton (700 MHz) and phosphorus (202 MHz) chemical shifts of 4 in 10 mM KH_2PO_4, 70 mM KCl, 0.2 EDTA (pH 7.0, T = 25 °C)

Base (5′–3′)	H8/H6	H1′	H2′	H2″	H3′	H4′	H5′/H5″[a]	H6′/H6″[a]	H2/Me	^{31}P
G1	7.41	6.04	2.94	2.94	4.98	4.37	3.98/4.05			
G2	8.14	6.01	2.98	2.35	5.13	4.39	4.40			−2.09
T3	7.84	6.18	2.17	2.54	4.89	4.28	4.29/4.27		1.96	−0.18
T4	7.15	6.04	2.06	2.61	4.88	4.23	4.38/4.30		1.03	−2.24
G5	7.40	5.99	3.37	2.86	4.85	4.42	4.27			−1.03
G6	7.68	5.94	2.74	2.57	5.10	4.41	4.23/4.41			−2.01
T7	7.85	6.42	2.47	2.59	4.84	4.42	4.23/3.91		1.96	−1.10
G8	7.47	5.72	2.03	2.34	4.74	4.07	4.03			−1.57
T9	7.24	5.86	1.95	2.36	4.61	3.01	3.66/3.55		1.65	−1.81
G10	7.45	6.03	3.66	2.94	4.89	4.26	4.10/4.12			−2.13
G11	8.17	6.00	2.95	2.35	5.13	4.39	4.20/4.18			−1.10
T12	7.84	6.18	2.17	2.54	4.89	4.28	4.29/4.27		1.96	−0.24
T13	7.22	6.07	2.12	2.64	4.90	4.27	4.16/3.90		0.96	−0.95
G14	7.51	6.13	3.86	3.01	4.97	4.37	4.10/4.34			−1.46
g15	4.91	5.81	4.86		4.49		4.36	4.11/3.92		−1.90

[a] No stereospecific assignment has been done.

The alternation of *syn* and *anti* G residues within each strand suggests that, as TBA, 4 folds into a monomolecular foldback quadruplex, characterized by two G tetrads. Further, a number of unusual NOE connectivities were observed between *syn*-Gs and Ts, indicating that $^{5'}$TG$^{3'}$ and $^{5'}$GT$^{3'}$ tracts do not adopt a helical winding, thus suggesting that, most probably, the TT and TGT tracts form loops. Thus, 4 is characterized by two G tetrads formed by the residues G1, G6, G10, G15 and G2, G5, G11, G14, respectively, where the underlined residues possess a *syn* glycosidic conformation. Interestingly the two tetrads assume a *syn-anti-syn-anti* and *anti-syn-anti-syn* arrangement of the bases. Then, the facts that TGT loop possesses some NOEs with G imino protons of one tetrad and that the two TT loops are characterized by NOEs with G imino protons of the other tetrad indicate that no loop assumes a dog-eared conformation. All this means that the 4 folds into a chair-like quadruplex structure characterized by the same folding

observed for TBA, being two strands parallel to each other and two strands oriented in opposite manner.

In order to get the three-dimensional structure of 4 at atomic level, an estimation of proton-proton distances has been done analyzing the cross-peak intensities in 2D NOESY experiments acquired at 700 MHz, both at T=25°C and T=5°C (mixing time 100 ms for the experiment acquired in D_2O and 200 ms for the experiments acquired in H_2O). Pseudo-atoms were introduced where needed. 138 distances were used for the calculations and, as suggested by the presence of eight G imino protons in the 1D ^{1}H-NMR spectrum, 32 supplementary distance restraints (HN1-O6, N1-O6, HN2-N7, N2-N7) for 16 hydrogen bonds corresponding to the two G-quartets, were incorporated during the computations (Table 2).

Table 2 Experimental constraints and structure statistics of the best 20 structures of 4

Experimental constraints	
Total NOEs	138
NOEs from nonexchangeable protons	117
NOEs from exchangeable protons	21
Hydrogen bonds constraints	36
Dihedral angle constraints	8
CVFF energy (kcal mol^{-1}) of the minimized structures	
Total	−4282.250 ± 2.396
Nonbond	143.685 ± 3.112
Restraint	18.979 ± 1.085
RMS deviations from the mean structure (Å)	
All backbone heavy atoms	0.80 ± 0.24
All heavy atoms	0.88 ± 0.29

Further, in agreement with NMR data, glycosidic torsion angles for four out of eight guanines involved in the formation of the two G-tetrads were fixed in the *anti*-domain (-160°/-70°), whereas the χ angle was kept in a range of 10°/100° (*syn*-conformation) for the remaining four G residues.

Therefore, three-dimensional structures which satisfy NOEs were constructed by simulated annealing (SA) calculations. An initial structure

of the oligonucleotide was constructed possessing a random conformation and minimized, in order to eliminate any possible source of initial bias in the folding pathway. Restrained simulations were carried out for 500 ps using the CVFF force field as implemented in Discover software (Accelrys, San Diego, USA). The restrained SA calculations started at 1000 K, and, thereafter, the temperature was decreased stepwise down to 273 K. The aim step was to energy-minimize and refine the structures obtained by using the steepest descent followed by the quasi-Newton-Raphson (VA09A) algorithms. A total of 20 structures was generated. Average RMSD values of 0.80 ± 0.24 Å and 0.88 ± 0.29 Å for the backbone and all heavy atoms, respectively, were obtained from the superimposition of all the 20 structures (Fig. 4).

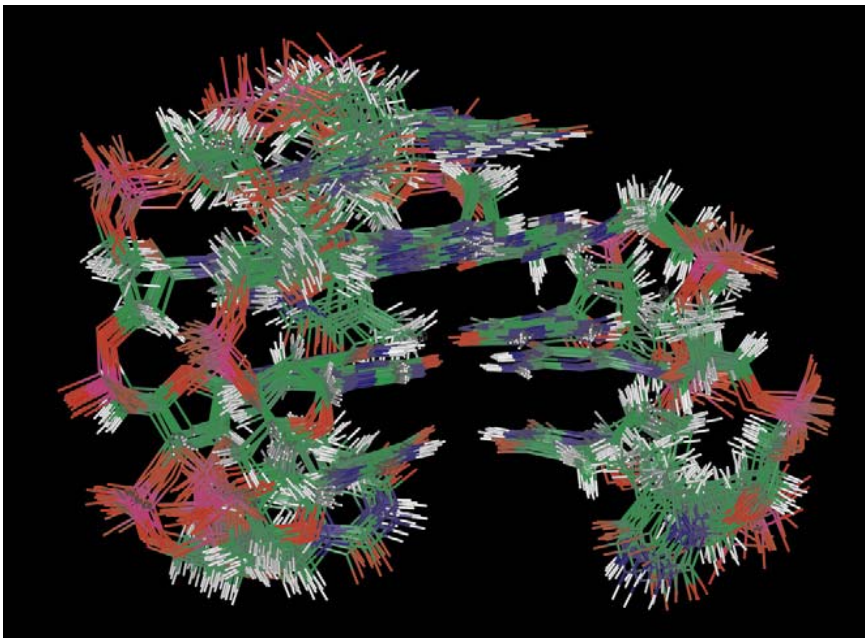

Fig 4 Side view representation of the superimposed 20 structures of 4. Backbones and bases are depicted in colored 'stick' (carbons, green; nitrogens, blue; oxygens, red; hydrogens, white).

These data, along with the lack of significant violations of the experimental restraints, suggest that the obtained structures are representative of the structure actually adopted in solution by 4.

Table 3 Rise, Tilt, Roll, and Twist of TBA and of the best-minimized structure of 4

TBA	Rise	Tilt	Roll	Twist	4
G1/G2	3.44	−10.60	−179.75	50.93	
	−3.27	−13.21	−171.63	53.01	G1/G2
G5/G6	2.67	6.22	−161.77	48.90	
	−2.78	5.99	−177.02	53.01	G5/G6
G10/G11	2.80	10.45	−178.83	48.50	
	−2.69	19.14	162.67	53.27	G10/G11
G14/G15	4.67	−17.19	−177.35	55.65	
	−3.79	−4.11	−174.67	51.53	G14/G15

Table 4. Shear, Stretch, Stagger, Buckle, Propeller, Opening of TBA and of the best-minimized structure of 4

TBA	Shear	Stretch	Stagger	4
G1/G6	−6.15	1.72	−0.88	
	−6.39	0.59	−0.47	G1/G6
G1/G10	4.82	2.32	−2.69	
	6.59	2.66	−1.05	G1/G10
G1/G15	−6.02	1.87	−0.59	
	−4.47	2.83	−0.18	G1/G15
G2/G5	6.42	2.77	−0.11	
	6.50	3.80	0.03	G2/G5
G2/G11	−6.04	1.75	−2.05	
	−5.88	4.29	−0.47	G2/G11
G2/G14	5.41	2.07	−1.82	
	5.95	5.06	−0.69	G2/G14
TBA	**Buckle**	**Propeller**	**Opening**	**4**
G1/G6	11.32	176.45	−89.48	
	9.96	−173.55	−90.27	G1/G6
G1/G10	3.47	−176.25	−0.95	
	−13.52	179.48	−0.57	G1/G10
G1/G15	−5.12	−162.35	86.32	
	4.42	−156.38	88.74	G1/G15
G2/G5	−5.50	157.97	−91.51	
	−9.24	175.09	−91.16	G2/G5
G2/G11	24.52	−177.68	−3.39	
	18.83	−171.56	−1.20	G2/G11
G2/G14	−5.12	−162.35	86.32	
	−4.68	−170.08	86.37	G2/G14

As expected, 4 shows a right-handed helical backbone geometry and three edge-wise connecting TT, TGT, and TT loops. As the unmodified TBA, 4 involves two stacked G-tetrads, with the same guanine *syn*/*anti* distribution around the GGGG tetrads, being characterized by the usual

syn-anti-syn-anti and *anti-syn-anti-syn* arrangement of the two tetrads respectively, and in the relative strand orientations, with two strands parallel to each other and two strands oriented in opposite manner.

A direct comparison of the most representative structure of 4 (the one with lowest energy after minimization), and the already reported NMR structure of the unmodified TBA (PDB code: 148D) has been accomplished by analyzing the helix parameters by CURVES[28-29] (Table 3 and 4).

These data clearly suggest that the overall structure adopted by 4 is similar to that of TBA. The base stacking of the two quadruplexes is quite similar, being the five-membered rings of the top guanines overlapped to the five-membered rings of the underneath guanine bases.

As for the sugar puckering, the two molecules possess almost the same sugar ring conformations except for the G1, G10, G11 and G15 that in 4 adopt a C4'-*exo,* C3'-*endo,* C3'-*endo* and C3'-*endo* conformations, respectively.

It is interesting to note that one structural difference relies on the orientation of T7, that folds back into the groove formed by the strands G5-G6 and G10-G11(Fig. 5).

In order to determine the effects of the introduction of a modified residue (g-LNA) on the CD profile and the thermal stability of the resulting quadruplex structure, CD spectra and CD melting and annealing experiments were acquired for 4 (Fig. 6A) and its natural counterpart. In particular, the CD spectra of 4 and TBA, performed at 20°C, are almost superimposable, both exhibiting two positive bands at 248 and 294 nm and a negative one at 267 nm, typical of anti-parallel quadruplex structures,[30-32] containing residues in *syn*-glycosidic conformations.

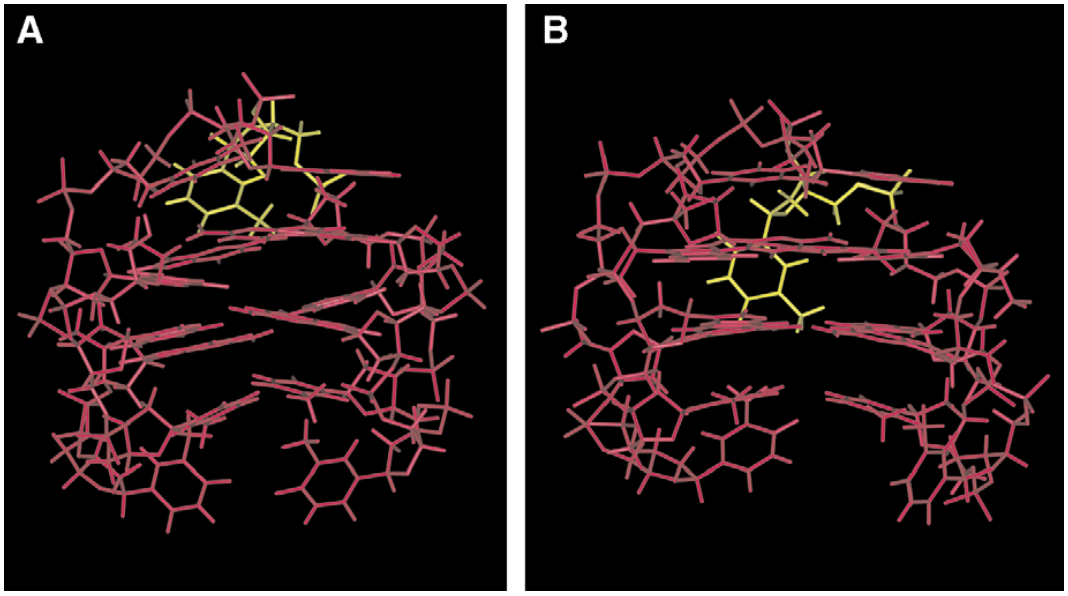

Fig 5 Side view of TBA (A) and 4 (B). Backbones and bases are depicted in colored 'stick' (magenta). The different orientation of T7 (in yellow) is highlighted

As for melting and annealing CD experiments, taking into account that the rates of quadruplex formation/dissociation are very slow, we collected the data at 10°C/h. No significant hysteresis emerged for both 4 (Fig. 6B) and TBA comparing annealing and melting curves thereby indicating that, at the scan rate used, both systems were at equilibrium. Therefore, the melting temperatures of 46°C and 52°C could be measured for 4 and TBA, respectively.

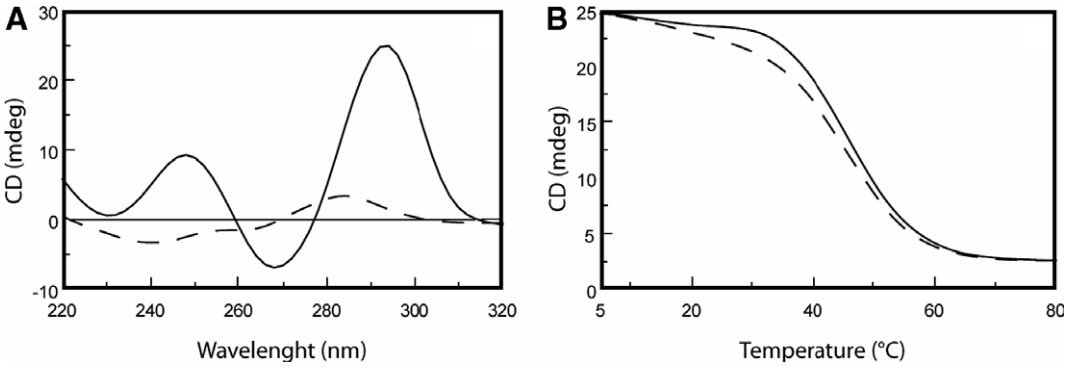

Fig. 6 (A) CD spectra of 4 at 20°C (continuous lines) and 90°C (dashed lines). (B) Melting (continuous lines) and annealing (dashed lines) CD experiments of 4.

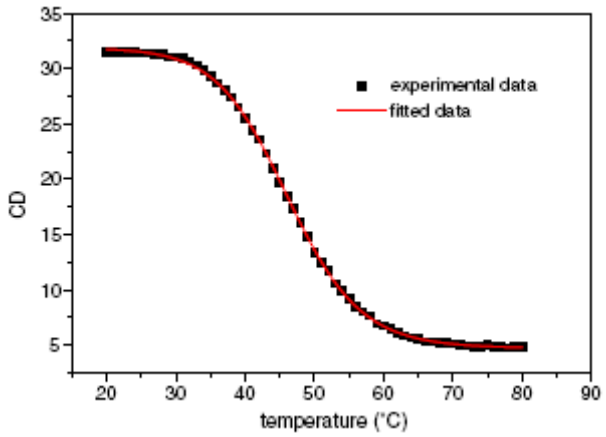

Fig. 7 Fitting of CD melting data of 4 according to the van't Hoff analysis

The melting curves were analyzed using van't Hoff analysis and thermodynamic parameters are shown in Table 5. The data, compared with those of unmodified TBA, indicate that the introduction of a LNA residue destabilizes the TBA structure. ΔH° values for the two aptamers are similar and correspond to a value of ΔH° of dissociation of about 80 kJmol^{-1} for G- tetrad, in accord with previously reported values.[33,34] The enthalpic data are in line with NMR data that show for the two quadruplexes a very similar geometry of the G-tetrads. Since ΔH° values are comparable and the T_m of 4 is lower, the lower thermodynamic stability is entropically controlled. In fact, the presence of a conformationally locked nucleotide increases the local organization of the phosphodiester backbone lowering the entropic content in the quadruplex with a consequence of a major gain in entropy in the dissociation process.

Table 5 Thermodynamic parameters for the unfolding of 4 and TBA

	T_m °C (±1)	$\Delta H°$ kJ mol^{-1} (±20)	$\Delta S°$ kJ mol^{-1} K^{-1} (±0.05)	$\Delta G°$(298 K) kJ mol^{-1} (±1)
4	46	163	0.51	10
TBA	52	160	0.49	13

The biological activity of 1, 2, 3, and 4 have been tested with a prothrombin time (PT) assay. PT assays on all four samples were performed on human plasma in strict comparison with TBA. In order to eliminate the variation caused by the measurements performed on different days, the samples of 1, 2, 3, 4 and TBA were prepared at the same time and folded together by heating the samples for 10 minutes at 80°C and slowly cooling them down at room temperature. The assays have been conducted after one week from the preparation procedure. Sample 1, 2 and 3 did not show any significant activity. On the other hand, 4 displayed an anticoagulant activity. However, as shown in figure 8, the activity of TBA on PT measurement turned to be more marked when compared with 4 in the same time frame. It is interesting to note that only the magnitude of the activity of 4 was reduced when compared to TBA since the onset as well as the endurance of the activity were almost comparable. The biological data reported here imply that the modifications of the structure adopted by 4 affect the biological activity, and therefore its interaction with the thrombin. Looking to the overall folding of the modified aptamer, it was not expected such reduction in the biological activity. However, it is interesting to note that TBA inhibits the thrombin activity when interacting with its fibrinogen exosite.[35] In particular, it has been proposed TBA could associate with thrombin in two different ways: through the T7-G8-T9 loop (as suggested by X-ray data) or through the T3 and T12 residues (as suggested by NMR models).[35] It is already highlighted the role of the conformation of the loops in another modified TBA.[36] However, the very little difference in the orientation of T7 does not seem to fully justify such a reduction of biological activity. Therefore, the reduced activity of 4 might suggest that the mode of action of TBA actually

requires a more wide recognition process that involve even locally a single residue.

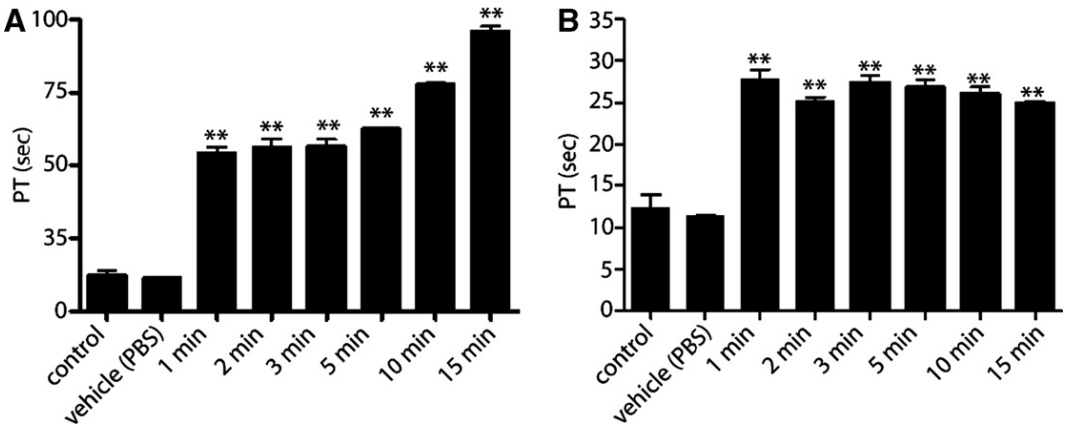

Fig. 8 Prothrombin time assessed by using human plasma. On the x-axis the minutes of incubation are reported, while PT in seconds is reported on y-axis. (A) Refers to TBA, while (B) to 4.

The structures reported here are, to the best of our knowledge, the first example of a monomolecular quadruplex structure containing an LNA residue. These results, particularly when viewed in the context of other recent findings about locked nucleic acids, serve to underscore further the polymorphic nature of LNAs. Furthermore, the many advantageous characteristics of LNA usage, demonstrate that LNAs might amplify the range of applicability of synthetic oligonucleotides as aptamers. Moreover, the data reported here might give a further insight into the understanding of the variables involved in the mode of action of TBA and thus it is possible to take advantage of these findings to design new thrombin binding aptamers.

Materials and methods

The oligonucleotides $^{5'}$ggttggtgtggttgg$^{3'}$ (1), $^{5'}$ggT'TggTGTggT'Tgg$^{3'}$ (2), $^{5'}$gGT'TGGTGTGGT'TGG$^{3'}$ (3) and $^{5'}$GGT'TGGTGTGGT'TGg$^{3'}$ (4), and the natural counterpart $^{5'}$GGT'TGGTGTGGT'TGG$^{3'}$ (TBA) were

synthesized on a Millipore Cyclon Plus DNA synthesizer, using solid phase β-cyanoethyl phosphoramidite chemistry at 15 μmol scale. The syntheses of the oligonucleotides were performed by standard methods and using g- and t-LNA phosphoramidites for the modified residue. The oligomers were detached from the support and deprotected by treatment with conc. aqueous ammonia at 55°C for 12 h. The combined filtrates and washings were concentrated under reduced pressure, redissolved in H_2O and analysed and purified by HPLC on a Nucleogel SAX column (Macherey-Nagel, 1000-8/46); using buffer A: 20 mM KH_2PO_4 aq. solution, pH 7.0, containing 20% (v/v) CH_3CN; buffer B: 1 M KCl, 20 mM KH_2PO_4 aqueous solution, pH 7.0, containing 20% (v/v) CH_3CN; a linear gradient from 0 to 100 % B in 30 min and flow rate 1 mL/min were used. The oligomers were collected and successively desalted by Sep-Pak cartridges (C18). The isolated oligomers resulted to be more than 99% pure (NMR).

Nuclear Magnetic Resonances

NMR samples were prepared at a concentration of approximatively 2 mM, in 0.6 ml (H_2O/D_2O 9:1) buffer solution having 10 mM KH_2PO_4, 70 mM KCl, 0.2 mM EDTA, pH 7.0. For D_2O experiments, the H_2O was replaced with D_2O by drying down the sample, lyophilization and redissolution in D_2O alone. NMR spectra were recorded with Varian ^{Unity}INOVA 500 MHz and 700 MHz spectrometers. ^1H chemical shifts were referenced relative to external sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS), whereas ^{31}P chemical shifts were referenced to external phosphoric acid (H_3PO_4 85% v/v). 1D proton spectra of samples in H_2O were recorded using pulsed-field gradient WATERGATE[37] for H_2O suppression. A proton-detected ^1H-^{31}P heteronuclear COSY was

recorded in D_2O in the hypercomplex mode with 2048 t_2 points and 96 t_1 increments, and a spectral width of 500 Hz along the ^{31}P dimension. Phase sensitive NOESY spectra[38] were recorded with mixing times of 100 and 200 ms (T=25°C and 5°C). Pulsed-field gradient WATERGATE was used for NOESY spectra in H_2O. TOCSY spectra[39] with mixing times of 100 ms were recorded with D_2O solutions.

All experiments were recorded using STATES-TPPI[40] procedure for quadrature detection. In all 2D experiments the time domain data consisted of 2048 complex points in t_2 and 400-512 fids in t_1 dimension. The relaxation delay was kept at 3s for NOESY experiments used in the structure determination. A relaxation delay of 1.2 s was used for all other experiments. The NMR data were processed on a SGI Octane workstation using FELIX 98 software (Accelrys, San Diego, CA).

Structural calculations

Structure calculations of 4 has been performed as follow: cross peak volume integrations were performed with the program FELIX 98, using the NOESY experiment collected at mixing time of 100 ms. The NOE volumes were then converted to distance restraints after they were calibrated using known fixed distances of H2'/H2" of G5, G8, G10, T12 and G14. Then a NOE restraint file was generated with three distance classifications as follows: strong NOEs (1.8 Å $\leq r_{ij} \leq$ 3.0 Å, where r_{ij} is the interproton distance between protons *i,j*), medium NOEs (2.5 Å $\leq r_{ij} \leq$ 4.0 Å) and weak NOEs (3.5 Å $\leq r_{ij} \leq$ 5.5 Å). A total of 138 NOE derived distance restraints were used.

Hydrogen bonds constraints were used: upper and lower distance limits of 2.0 Å and 1.7 Å for hydrogen-acceptor distance, and 3.0 Å and 2.7 Å for donor-acceptor distance, respectively. These constraints for H-bonds

did not lead to an increase in residual constraints violation. No backbone torsion angles were used. According to NMR data, glycosidic torsion angles were kept in a range of -160°/-70° for the G *anti*, whereas a range of 10°/100° was used for the G *syn*.

The calculations have been performed using a distance dependant macroscopic dielectric constant of 4*r and an infinite cut-off for nonbonded interactions to partially compensate for the lack of the solvent.[41] Thus the three-dimensional structures which satisfy NOE and dihedral angle constraints were constructed by simulated annealing calculations. An initial structure of the oligonucleotide was built using a completely random array of atoms. Using the steepest descent followed by quasi-Newton-Raphson method (VA09A) the conformational energy was minimized. Restrained simulations were carried out for 500 ps using the CVFF force field as implemented in Discover software (Accelrys, San Diego, USA). The simulation started at 1000 K, and then the temperature was decreased stepwise until 273 K. The final step was again to energy-minimize to refine the structures obtained, using successively the steepest descent and the quasi-Newton-Raphson (VA09A) algorithms. Both dynamic and mechanic calculations were carried out by using 1 (kcal/mol)/Å^2 flatwell distance restraints. 20 structures were generated. RMSD values of 0.80 ± 0.24 Å and 0.88 ± 0.29 Å for the backbone and heavy atoms, respectively, were calculated for all the 20 structures.

Illustrations of structures were generated using the INSIGHT II program, version 2005 (Accelrys, San Diego, USA). All the calculations were performed on a PC running Linux WS 4.0.

Circular Dichroism

CD samples of 4 and TBA were prepared at a concentration of 1 x 10^{-4} M, by using the buffer solution used for NMR experiments: 10 mM KH_2PO_4, 70 mM KCl, 0.2 mM EDTA, pH 7.0. CD spectra and CD melting curves were registered on a Jasco 715 circular dichroism spectrophotometer in a 0.1 cm pathlength cuvette. For the CD spectra, the wavelength was varied from 220 to 320 nm at 100 nm min^{-1}, and the spectra recorded with a response of 16 s, at 2.0 nm bandwidth and normalized by subtraction of the background scan with buffer. The temperature was kept constant at 20°C with a thermoelectrically controlled cell holder (Jasco PTC-348). CD melting curves were registered as a function of temperature from 20 to 90°C at 294 nm with a scan rate of 10°C h^{-1} for both quadruplexes. The CD melting curves of TBA and modified aptamer showed sigmoidal profiles and were modeled by a two-state transition, using a theoretical equation for an intramolecular association, according to the van't Hoff analysis.[42] The T_m and ΔH° values provide the best fit of the experimental melting data. The reported errors for thermodynamic parameters are the standard deviations of the mean from the multiple determinations. The ΔS° values were calculated by equation $\Delta S° = \Delta H° / T_m$ and free energy change values by the equation $\Delta G°(T) = \Delta H° - T\Delta S°$.

Prothrombin Time (PT) assay

Human plasma samples were collected by venipuncture, in presence of 0.1 volumes 3.8% sodium citrate and fractionated by centrifugation at 2000g for 5 minutes. PT times were measured by using Koagulab MJ Coagulation System with a specific kit RecombiPlas Tin HemosIL (Inst. Labs, Lexinton, USA). Briefly, this method relies on the high sensitivity

thromboplastin reagent based on recombinant human tissue factor. The addition of recombiplastin to the plasma in presence of calcium ions initiates the activation of extrinsic pathway. This results ultimately in the conversion of fibrinogen to fibrin, with a formation of solid gel. The procedure was performed according to the manufacturer's instructions. TBA, 1, 2, 3, and 4 or vehicle (PBS) were added at different time points in volume of 2 μl at final concentration of 20 μM. Data are expressed as mean ± s.e.m. and are representative of at least three different measurements.

Notes and references

1 Hanss, B.; Leal-Pinto, E.; Bruggeman, L. A.; Copeland; T. D., Klotman, P. E. *Proc. Natl. Acad. Sci.*, 1998, **95**, 1921.

2 Verma, S.; Eckstein, F. *Annu. Rev. Biochem.*, 1998, **67**, 99.

3 Koshkin, A. A.; Rajwanshi, V. K., Wengel, J. *Tetrahedron Lett.*, 1998, **39**, 4381.

4 Wengel, J. *Acc. Chem. Res.*, 1998, **32**, 301-310.

5 Obika, S., Nanbu, D.; Hari, Y.; Andoh, J.; Morio, K.; Doi, T.; Imanishi, T.; *Tetrahedron Lett.*, 1998, **39**, 5401.

6 Singh, S. K.; Nielsen, P.; Koshkin A. A.; Wengel, J. *Chem. Commun.*, 1998, **4**, 455.

7 Nielsen, K. E.; Singh, S. K.; Wengel, J.; Jacobsen, J. P. *Bioconjugate Chemistry*, 2000, **11**, 228.

8 Bondensgaard, K.; Petersen, M.; Singh, S. K.; Rajwanshi, V. K.; Kumar, R.; Wengel J.; Jacobsen, J. P. *Chem. Eur. J.*, 2000, **6,** 2687.

9 Braasch, D. A.; Corey, D. R. *Chem. Biol.*, 2001, **8,** 1.

10 Arzumanov, A.; Walsh, A. P. , Rajwanshi, V. K.; Kumar, R.; Wengel J.; Gait, M. J. *Biochemistry*, 2001, **40,** 14645.

11 Vester, B.; Lundberg, L. B.; Sorensen, M. D.; Babu, B. R.; Douthwaite S.; Wengel, J. *J. Am. Chem. Soc.*, 2002, **124**, 13682.

12 Randazzo, A.; Esposito, V.; Ohlenschlaeger, O.; Ramachandran R.; Mayol, L. *Nucleic Acids Res.*, 2004, **32***(10)*, 3083.

13 Williamson, J. R.; Raghuraman, M. K.; Cech, T. R. *Cell,* 1989, **59**, 871.

14 Simonsson, T.; Pecinka, P.; Kubista, M. *Nucleic Acid Res.*, 1998, **26**, 1167.

15 Nimjee, S. M.; Rusconi, C. P.; Sullenger, B. A. *Annual Review of Medicine,* 2005, **56**, 555.

16 Block, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J. *Nature,* 1992, **355**, 564.

17 Gold, L.; Polisky, B.; Uhlenbeck O.; Yarsus, M. *Annu. Rev. Biochem.*, 1995, **64**, 763.

18 Griffin, L. C.; Tidmarsh, G. F.; Bock, L. C.; Toole, J. J.; Leung, L. K. *Bloods*, 1993, **81**, 3271.

19 Li, W. X.; Kaplan, A. V.; Grant, G. W.; Toole, J. J.; Leung, L. L., *Blood*, 1994, **83**, 677.

20 Wang, K. Y., McCurdy, S.; Shea, R. G.; Swaminathan, S.; Bolton, P. H. *Biochemistry*, 1993, **32**, 1899.

21 Macaya, R. F.; Schultze, P.; Smith, F. W.; Roe, J. A.; Feigon, J. *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 3745-3749.

22 Paborsky, L. R.; McCurdy, S. N.; Griffin, L. C.; Toole J. J.; Leung, L. K. *J. Biol. Chem.*, 1993, **268**, 20808.

23 Dias, E.; Battiste, J. L.; Williamson, J. R. *J. Am. Chem. Soc.*, 1994, **116**, 4479.

24 He, G. -X.; Krawczyk, S. H.; Swaminathan, S.; Regan, S. G.; Dougherty, J. P., Terhorst, T.; Law, V. S.; Griffin, L. C.; Coutrè, S.; Bischofberger, N. *J. Med. Chem.*, 1998, **41**, 2234.

25 He, G. -X.; Williams, J. P.; Postich, M. J.; Swaminathan, S.; Regan, S. G.; Terhorst, T.; Law, V. S.; Griffin, L. C.; Cheri, M. T.; Coutrè, S.; Bischofberger, N. *J. Med. Chem.*, 1998, **41**, 4224.

26 Smith, F.W.; Feigon, J. *Biochemistry*, 1993, **32**, 8682.

27 Wang, Y.; Patel, D. J. *Structure*, 1993, **1**, 263.

28 Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.*, 1988, **6**, 63.

29 Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.*, 1989, **6**,655.

30 Lu, M.; Guo, Q.; Kallenbach, N. R. *Biochemistry*, 1993**, 32**, 598.

31 Lu, M.; Guo, Q.; Kallenbach, N. R. *Biochemistry*, 1993, **32**, 3596.

32 Smirnov, I.; Shafer, R. H. *Biochemistry*, 2000, **39**, 1462.

33 Shafer, R. H. *Prog. Nucleic Acids Res. Mol. Biol.*, 1998, **59**, 55.

34 Petraccone, L.; Erra, E.; Nasti, L.; Galeone, A.; Randazzo, A.; Mayol, L.; Barone, G.; Giancola, C. *J. Biomol. Macromol.*, 2002, **31**, 131.

35 Padmanachan, K.; Tulinsky, A. *Acta Cryst.*, 1996, **D52**, 272.

36 Martino, L.; Virno, A.; Randazzo, A.; Virgilio, A.; Esposito, V.; Giancola, C.; Bucci, M.; Cirino, G.; Mayol, L. *Nucleic Acid Res.*, 2006, **34***(22)*, 6653.

37 Piotto, M.; Saudek, V.; Sklenar, V. J. *J. Biomol. NMR*, 1992, **2**, 661.

38 Jeener, J., Meier, B.; Bachmann, H. P.; Ernst, R. R. *J. Chem. Phys.,* 1979, **71**, 4546.

39 Braunschweiler, L.; Ernst, R. R. *J. Magn. Reson.*, 1983, **53**, 521.

40 Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.*, 1989, *85*, 393.

41 Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta S.; Weiner, P. J. *J. Am. Chem. Soc.*, 1984, **106**, 765.

42 Marky, L. A.; Breslauer, K. J. *Biopolymers,* 1987, **26**, 1601.

CHAPTER 3

Targeting G-quadruplexes

1. Structural and Thermodynamic Studies of the Interaction of Distamycin A with the Parallel Quadruplex Structure $[d(TGGGGT)]_4$

The ends of the chromosomes in all eukaryotic species have specialized non-coding DNA sequences that, together with associated proteins, are known as telomeres. Telomere protects the ends of the chromosome from damage and recombination and its shortening has been implicated in cellular senescence. Telomeric DNA consists of tandem repeats of simple short sequences, rich in guanine residues. In the presence of metal ions such as K^+ or Na^+, telomeric DNA can form structures of potential biological significance, the G-quadruplexes.[1] G-quadruplex structures comprise stacks of G-tetrads, which are the planar association of four guanines in a cyclic Hoogsteen hydrogen-bonding arrangement.[2] Telomerase, the enzyme which elongates the G-rich strand of telomeric DNA, is active in about 85% of tumors, leading the cancer cells to infinite lifetime. The inhibition of telomerase has become an attractive strategy for the anticancer therapy and, because telomerase requires a single-stranded telomeric primer, the formation of G-quadruplex complexes by telomeric DNA inhibits the telomerase activity. Furthermore, small molecules that stabilize G-quadruplex structures have been found to be effective telomerase inhibitors and, then, the use of drugs to target G-quadruplexes is emerging as a promising way to interfere with telomere replication in the tumors cells and to act as anticancer agents.[3]

The truncated telomeric sequence from *Oxytricha*, d(TGGGGT), has been previously reported in literature to form a tetramer, with a parallel-stranded structure. This G-quadruplex has been determined by NMR and X-ray techniques, the strands associate generating a right-handed helix, containing four equivalent grooves and all bases in the *anti* glycosidic conformation.[4,5]

Distamycin A (Dist-A) is a small molecule with antibiotic properties which binds with high affinity to B-form DNA.[6,7] In particular, it has been demonstrated that two Dist-A molecules, side by side in an antiparallel orientation, are able to bind the expanded minor groove of an A and T rich regions of duplex DNA in a 2:1 drug/DNA stoichiometry.

Moreover, Dist-A has been shown to interact with four-stranded parallel DNA quadruplex containing oligonucleotides of different sequence.[8,9] Two opposite models have been proposed for the distamycin-quadruplex complex: the first, proposed by some of us, suggests that distamycin molecules bind as dimers in two opposite grooves of quadruplex $[d(TGGGGT)]_4$;[8] the second one suggests that two distamycin molecules stack on the terminal G-tetrad planes of the quadruplexes $[d(TAGGGTTA)]_4$, $[d(TAGGGGTT)]_4$ and $[d(TAGGGGGT)]_4$.[9] More recently, some distamycin analogues have been found interacting with a DNA quadruplex of sequence $dGGG(TTAGGG)_3$;[10] a mixed groove/G-quartet stacking binding mode was suggested for the interaction on the basis of molecular modeling. All these results suggest that, most probably, there is no a general code that explain the drug-quadruplex binding and a sequence-dependent interaction could also be important. Since it is commonly believed that a deep understanding of factors which determine the complex formation

and stabilization is important for the drug design, we tried to shed light on the binding mode of Dist-A to the quadruplex $[d(TGGGGT)]_4$, analyzing the NMR titration profile of a modified quadruplex, namely $[d(TGG^{Me}GGT)]_4$ (where dG^{Me} is 8-methyl-2'-deoxyguanosine) and applying the recently developed STD method.[11] Furthermore, we determined the three-dimensional structure of the Dist-A/$[d(TGGGGT)]_4$ complex by an in-depth NMR study followed by dynamic and mechanic calculations. We also report the energetics of the interaction by isothermal titration calorimetry (ITC).

All results unequivocally assess that distamycin molecules interact with $[d(TGGGGT)]_4$ in a 4:1 binding mode, with two distamycin dimers binding simultaneously two opposite grooves of the quadruplex.

Materials and methods

Oligonucleotide synthesis

The oligonucleotides $d(TGG^{Me}GGT)$ and d(TGGGGT) were synthesized on a Millipore Cyclone Plus DNA synthesizer using solid phase β-cyanoethyl phosphoramidite chemistry at 15 μmol scale. The synthesis of the suitably protected dG^{Me} phosphoramidite monomer was performed following the synthetic strategy proposed by Kohda et al.[12] with minor modifications. The oligomers were detached from the support and deprotected by treatment with concentrated aqueous ammonia at 55°C for 12 h. The combined filtrates and washings were concentrated under reduced pressure, redissolved in H_2O, analyzed and purified by high-performance liquid chromatography (HPLC) on a Nucleogel SAX column (Macherey–Nagel, 1000-8/46); using buffer A: 20 mM KH_2PO_4/K_2HPO_4 aqueous solution (pH 7.0), containing 20% (v/v) CH_3CN; buffer B: 1 M KCl, 20 mM KH_2PO_4/K_2HPO_4 aqueous

solution (pH 7.0), containing 20% (v/v) CH_3CN; a linear gradient from 0 to 100% B for 30 min and flow rate 1 ml/min were used. The fractions of the oligomer were collected and successively desalted by Sep-pak cartridges (C-18). The isolated oligomers proved to be >98% pure NMR.

Nuclear magnetic resonance experiments

NMR samples were prepared at a concentration of 2 mM, in 0.6 ml (H_2O/D_2O 9:1) buffer solution having 10 mM KH_2PO_4, 70 mM KCl, 0.2 mM EDTA, pH 7.0. For D_2O experiments, the H_2O was replaced with D_2O by drying down the sample, followed by lyophilization and redissolution in D_2O alone. NMR spectra were recorded with Varian ^{Unity}INOVA 700 MHz spectrometer. ^1H chemical shifts were referenced relative to external sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS). 1D proton spectra of samples in H_2O were recorded using pulsed-field gradient DPFGSE[13,14] for H_2O suppression. Phase-sensitive NOESY spectra[15] were recorded with mixing times of 100 and 200 ms (T = 25°C). Pulsed-field gradient DPFGSE sequence was used for NOESY experiments in H_2O. Pulsed-field gradient WATERGATE[16] sequence was used for $^1H–^{15}N$ HSQC[17] spectrum in H_2O. TOCSY spectra[18] with mixing times of 100 ms were recorded with D_2O solutions.

All experiments were recorded using STATES-TPPI[19] procedure for quadrature detection. In all 2D experiments, the time domain data consisted of 2048 complex points in t2 and 400–512 fids in t1 dimension. The relaxation delay was kept at 3 s for NOESY experiments used in the structure determination. A relaxation delay of 1.2 s was used for all other experiments. The NMR data were processed on a SGI Octane workstation using FELIX 98 software (Accelrys, San Diego, CA).

Structure calculations

Cross-peak volume integrations were performed with the program FELIX 98, using the NOESY experiment collected at mixing time of 100 ms. The NOE volumes were then converted to distance restraints after they were calibrated using known fixed distances of H2'/H2" of GA_5, TA_6, GB_3, GB_4 and GB_5. Then a NOE restraint file was generated with three distance classifications as follows: strong NOEs (1.0 Å < r_{ij} < 3.5 Å), medium NOEs (3.0 Å < r_{ij} < 4.5 Å) and weak NOEs (4.0 Å < r_{ij} < 6.0 Å). A total of 440 NOE derived distance restraints were used.

Hydrogen bonds constraints were used (1.7 Å < r_{ij} < 2.3 Å). These constraints for H-bonds did not lead to an increase in residual constraints violation. A total of 96 backbone torsion angles were used in the calculations too. Particularly, the backbone torsion angles □, β, □, □ and ε were restrained in the range -150°/-30°, -230°/-110°, 20°/100°, 95°/175° and -230°/-110°, respectively, and glycosidic torsion angles χ were fixed in the *anti*-domain (-155°/-75°). Moreover, 64 planar constraints were used for G bases. The calculations have been performed using a distance dependent macroscopic dielectric constant of 4*r and an infinite cut-off for non bonded interactions to partially compensate for the lack of the solvent have been used.[20] Thus the 3D structures which satisfy NOE and dihedral angle constraints were constructed by simulated annealing calculations.[21] An initial structure of the oligonucleotide was built using a completely random array of atoms. Using the steepest descent followed by quasi-Newton–Raphson method (VA09A), the conformational energy was minimized. Restrained simulations were carried out for 300 ps using the CVFF force field as implemented in Discover software (Accelrys, San Diego, CA). The

simulation started at 500 K, and then the temperature was decreased stepwise until 100 K. The final step was again to energy-minimize to refine the obtained structures, using successively the steepest descent and the quasi-Newton–Raphson (VA09A) algorithms. Both dynamic and mechanic calculations were carried out by using 20 kcal mol^{-1} $Å^{-2}$ flatwell distance restraints: 20 structures were generated. RMSD (root mean square deviation) value of 0.28 ± 0.069 Å for heavy atoms was calculated for the best ten structures. Illustrations of structures were generated using the INSIGHT II program, version 2005 (Accelrys, San Diego, CA), Chimera,[22] Python,[23] APBS[24] and MSMS.[25] All the calculations were performed on a PC running Linux WS 4.0.

The final set of coordinates has been deposited in the Protein Data Bank (accession code:2JT7).

Isothermal Titration Calorimetry

The binding energetics was obtained with a CSC 4200 Calorimeter from Calorimetry Science Corporation (CSC, Utah) at 25°C. Calorimetric titrations were carried out by injecting 15 μL aliquots of a 300-900 μM ligand solution into a 30-50 μM quadruplex solution at 400 seconds intervals with stirring at 297 rpm at 25 °C for a total of 16 injections. Titration curves were corrected for heat of dilution determinated by injecting the ligand into the buffer.

The solutions of Dist-A (Sigma-Aldrich, USA) and quadruplex DNA samples were prepared in the same buffer. Control experiments were done by injecting distamycin in buffer to obtain the heat effects for dilution of the drug. The calorimetric enthalpy for each injection was calculated after correction for the heat of Dist-A dilution. To obtain the thermodynamic properties of interaction a non linear regression analysis

was performed using a simple two sets of independent binding-site model.[26]

STD experiments

STD-NMR experiments were performed on a Bruker Avance DRX 600-MHz spectrometer equipped with cryoprobe, at 27 °C. NMR samples were prepared by dissolving the ligands and the $[d(TGGGGT)]_4$ oligomer in 200 μl (3 mm NMR tube) of a solution H_2O:D_2O (90:10) (20 μL, 99.996%, Sigma Aldrich) containing phosphate buffered saline (10 mM) at pH 7.1. A high ligand–quadruplex molar excess (50:1) was used for the best STD effects. In particular, the concentration of ligands was set at 4.5 mM. The concentration of the tetra-stranded DNA olygomer was 90 μM.

Typically, 32 scans were recorded for the reference STD spectrum, whereas 64 scans were recorded for each DF-STD spectrum (saturation time=4 sec). The water ^1H signal was suppressed using a 3-9-19[17] pulse sequence. The STD effects of the individual protons were calculated for each compound relative to a reference spectrum with off-resonance saturation at δ= -16 ppm.

The calculated binding mode indexes (BMIs) are:

$$BMI = \frac{\sum_i \left(\frac{SN_{aromatic}/SN_{rif}}{SN_{aliphatic}/SN_{rif}} \right)}{n_i} \qquad \text{(eq. 1)}$$

$$BMI' = \frac{\sum_i \left(\frac{SN_{imino}/SN_{rif}}{SN_{aliphatic}/SN_{rif}} \right)}{n_i} \qquad \text{(eq. 2)}$$

$$BMI'' = \frac{\sum_i \left(\frac{SN_{imino}/SN_{rif}}{SN_{aromatic}/SN_{rif}} \right)}{n_i} \qquad \text{(eq. 3)}$$

where $SN_{aliphatic}$, $SN_{aromatic}$ and SN_{imino} are the differences between the

intensities (expressed as S/N ratio) of the signals of the drug when irradiated in the on-resonance STD spectrum (*i.e.* in the aliphatic, aromatic and imino regions, respectively) and that of the signals in the off-resonance NMR spectrum. SN_{rif} is instead the intensity of the same signal in the off-resonance spectrum (see Experimental Section), and n_i is the number of the signals.

Results

Synthesis and NMR titration of the quadruplex $[d(TGG^{Me}GGT)]_4$

In order to clarify the binding mode of Dist-A with the quadruplex $[d(TGGGGT)]_4$, we have designed and synthesized a new oligonucleotide, namely $d(TGG^{Me}GGT)$, where dG^{Me} is 8-methyl-2'-deoxyguanosine, potentially capable to form quadruplex structure and possessing a bulky group (methyl) into the grooves. NMR sample of $d(TGG^{Me}GGT)$ was prepared at a concentration of 2 mM, in 0.6 ml (H_2O/D_2O 9:1) buffer solution having 10 mM KH_2PO_4, 70 mM KCl, 0.2 mM EDTA, pH 7.0. The sample was annealed for 5-10 minutes at 80°C and slowly cooled down to room temperature, then ^{1}H-NMR spectra of it were recorded by using DPFGSE pulse sequence for H_2O suppression.[13,14] The ^{1}H-NMR spectrum (700 MHz, T = 25 °C) shows the presence of four well defined singlets in the region 11-12 ppm, attributable to imino protons involved in Hoogsteen hydrogen bonds of G quartets (Fig. 1, molar ratio = 0).

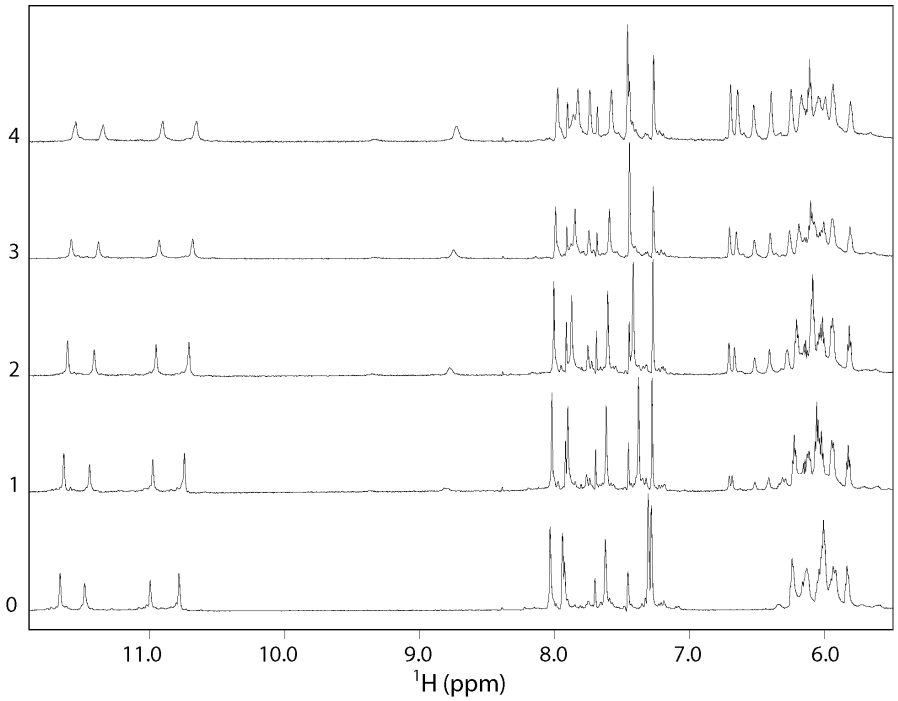

Fig. 1. NMR titration of [d(TGG^{Me}GGT)]$_4$ with Dist-A. The drug:quadruplex molar ratios are shown along the side of the spectra.

Moreover, the presence of five signals, belonging to three guanine H8 and two thymine H6 protons in the aromatic region, and of a methyl resonance at δ_H 2.24 ppm (attributable to dG^{Me}) indicates that a single well defined quadruplex species is present in solution, consisting of four G-tetrads and possessing a fourfold symmetry with all strands parallel to each other. The entire pattern of NOEs observed in the NOESY experiment (700 MHz, T=25 °C, mixing time 100 ms) (see Supporting Information) indicates that d(TGG^{Me}GGT) forms a tetramolecular parallel quadruplex possessing a right-handed B-form helix structure and that the backbone conformations resemble closely that of the [d(TGGGGT)]$_4$. Noteworthy, the direct comparison of the intensities of the NOESY crosspeaks between H8 (Me-8 in the case of dG^{Me}) proton

bases and sugar H1', and between H8/Me-8 proton bases and sugar H2'/H2" resonances indicates that, exactly as in $[d(TGGGGT)]_4$,[4,5] all Gs residues (including dG^{Me}) adopt an *anti* glycosidic conformation.

Thus, the quadruplex $[d(TGG^{Me}GGT)]_4$ has been titrated with Dist-A and the results are reported in Fig. 1. The initial addition of 1 mol equiv. of Dist-A to the quadruplex caused the down-field shift of the signal at δ_H 7.30 belonging to the H6 proton of T6. The progressive increase of Dist-A concentration up to 4 mol equiv. caused only a drift of DNA signals and drug resonances to gradually grow in intensity (between 6 and 7 ppm). A further addition of drug did not lead to significant changes.

DF-STD method

The binding mode of Dist-A with the quadruplex $[d(TGGGGT)]_4$ has also been deepened by means of the DF-STD[11] method. This method consists in an acquisition of a set of experiments at different saturation frequencies, and it can provide useful information on binding mode of DNA interacting molecules, discriminating among base-pair intercalators, minor groove binders, and external backbone binders.

In its original version, where duplex-DNA is investigated, the DF-STD method is based on the acquisition of two parallel sets of STD experiments performed under the same experimental conditions, irradiating on aromatic protons of purine/pyrimidine bases and on deoxyribose/backbone resonances[11] (both spectral regions where no ligand resonances were present).[27,28] In the present work we have used a third saturation frequency that is in the range of the imino proton frequencies, since such protons, pointing inside the quadruplex core, may be very diagnostic for probing this structural portion of the DNA

quadruplex.

For our investigation, we have firstly employed a known DNA ligand, the ethidium bromide, a base-pairs intercalator which also interacts with the phosphate backbone thank to its two charged amino group,[29,30] and then we used the Dist-A.

The aim to use this intercalator is to correctly interpret the experimental outcomes at different saturation frequency in presence of a DNA quadruplex as biological target, considering the established binding mode to DNA. For Dist-A, the saturation frequencies are 11.0 ppm, 9.0 ppm and 5.1 ppm, whereas for ethidium bromide are 11.0 ppm, 9.5 ppm and 2.4 ppm. The results of the irradiations for Dist-A and ethidium bromide are reported in the Supporting Information. We have calculated three binding mode indexes (BMIs)[11] for both molecules and the results are reported in Table 1.

Table 1: BMI values for ligands **1** and **2** calculated by using equation 1-3

ligand	BMI	BMI'	BMI''
ethidium bromide	1.3	0.8	0.7
distamycin A	0.8	0.2	0.4

NMR study

As already reported,[8] the NMR titration profile of $[d(TGGGGT)]_4$ with Dist-A is peculiar. An increase of Dist-A concentration up to 2 mol equiv. caused drug resonances to gradually grow in intensity and a progressive drift of DNA signals. Nearly at 2:1 ligand:quadruplex stoichiometry, a further addition of drug caused a complication of the

spectrum due to the appearance of a separate set of proton resonances. The intensities of these new resonances rose by increasing the amount of drug with the concomitant falling off of the original signals which completely disappeared at a ratio of 4:1 drug-quadruplex. The final NMR spectrum showed that the binding of the ligand to the quadruplex caused the loss of the original fourfold symmetry of the free quadruplex. Particularly, 8 imino proton, 4 methyl and 12 aromatic proton resonances were discernible in the 1D proton spectra of the complex, counting for two pairs of not magnetically equivalent strands (named A and B).

An almost complete assignment (see Supporting Information) of the nonexchangeable/exchangeable protons of the complex has been accomplished by means of a combination of the analysis of 2D NOESY (700 MHz, 25°C), TOCSY spectra (700 MHz, T = 25°C) and ^{1}H-^{15}N-HSQC experiments (700 MHz for ^{1}H, 70 MHz for ^{15}N, T = 25°C). In particular, as for DNA, ^{1}H resonances within each deoxyribose were identified by 2D TOCSY experiment, while the analysis of NOEs among base protons and H1', H2' and H2" protons allowed us to assign all base protons. It is interesting to note that the direct comparison of the intensities of the NOESY crosspeaks (700 MHz, T=25 °C, mixing time 100 ms; see Supporting Information) between the H8 proton bases and sugar H1' resonances, and among H8 proton bases and sugar H2'/H2" resonances indicates that all Gs residues of the complexed DNA adopt an *anti* glycosidic conformation. Then, all bases have classical H8/H2'-H2" sequential connectivities to 5' neighboring (see Supporting Information), indicating that the four strands are involved in the formation of a helical structure. Moreover, the entire pattern of NOEs observed indicates that the backbone conformations resemble closely

that of the uncomplexed $[d(TGGGGT)]_4$ possessing a right-handed B-form helix structure.

Guanine imino protons and Dist-A amide and amidinium protons were firstly identified by means of an ^{1}H-^{15}N-HSQC experiment. Thus, eight exchangeable protons correlate with ^{15}N in the region between 139.2-144.6 ppm, indicating that they are attributable to imino protons of G residues.[31] Furthermore, other four exchangeable protons correlate with ^{15}N in the region between 110.6-130.8 ppm, indicating that they are attributable to amide protons of Dist-A and two exchangeable protons correlate with ^{15}N in the region between 80.4-80.7 ppm indicating that they are attributable to amidinium protons of Dist-A (see Supporting Information).

Each exchangeable proton signal was then assigned to the pertinent hydrogen by an in-depth analysis of the NOESY spectra. In particular, as far as DNA is concerned, NOEs among H1 protons at δ_H 10.62 and 10.68, and two methyl protons at δ_H 1.52 and 1.94 (belonging to T1_A and T1_B residues, respectively) led us to identify the imino protons belonging to the tetrad that is in proximity of the 5' edge of the quadruplex. At the same way, NOEs among H1 protons at δ_H 10.80 and 10.62 ppm and the methyl groups at δ_H 1.74 and 1.75 ppm of the thymines T6_A and T6_B, respectively, led us to identify the imino protons belonging to the tetrad that is in proximity of the 3' edge. The other four H1 resonances of the remaining two central tetrads were identified by analyzing the NOE connections with adjacent tetrads.

The resonances of Dist-A were assigned, firstly, identifying the resonances belonging to the hydrogen on the convex side of the molecule (Fig. 2).

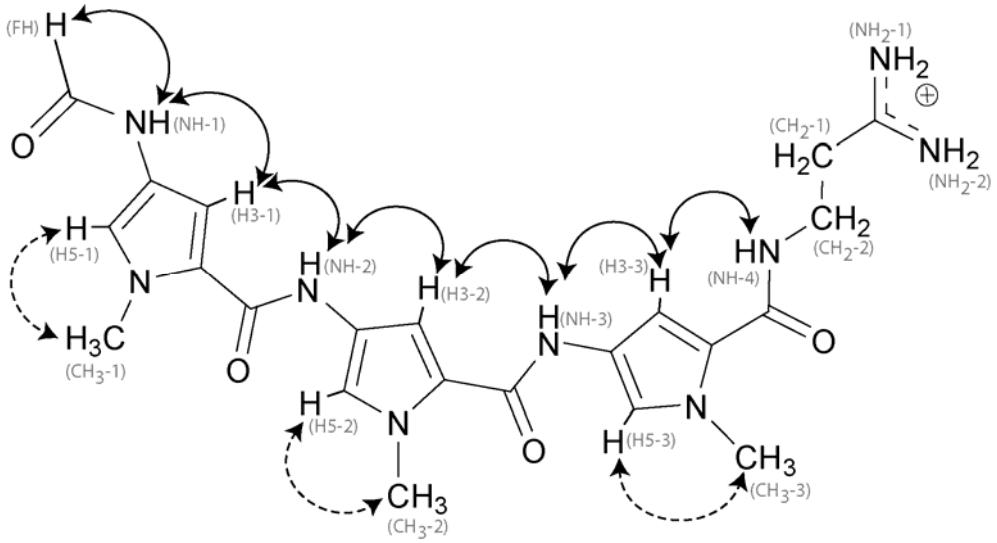

Fig. 2 NOE connectivities observed within Dist-A molecule. Regular and dashed lines represent connectivities on the concave and convex side of the molecule, respectively. Nomenclature used for Dist-A is reported in brackets (in gray).

Particularly, the H5 pyrrole protons (δ_H 6.66, 6.72 and 6.51 ppm for H5-1, H5-2 and H5-3, respectively) have been assigned taking into account that they both exhibit scalar and dipolar coupling with methyls that are linked to the same pyrrole ring (δ_H 3.39, 3.55 and 3.40 ppm for CH_3-1, CH_3-2, CH_3-3, respectively). As for the protons on the concave side of Dist-A, they have been assigned following the sequential NOE starting from the formyl proton at δ_H 7.70 ppm to the adjacent NH-1 (δ_H 9.80 ppm) and then sequentially to all other H3/NH hydrogens (Fig. 2).

The NOESY spectrum of the complex contains many intermolecular drug:drug and drug:DNA NOEs, in addition to intramolecular ones. Particularly, as for drug:drug contacts, 9 head to tail NOEs for each Dist-A molecule (counting for 18 NOE contacts for each dimer) were clearly discernable in the NOESY spectra (Table 2). In addition, 22 NOEs were observed between Dist-A and $[d(TGGGGT)]_4$, predominantly involving aromatic protons of strand A and H1' protons

of strand B (Fig. 3 and Table 3).

Table 2. Intermolecular head to tail drug-drug NOE contacts.

FH	CH_2
H3-1	H3-3
H5-2	CH_3-3, H5-3
H5-3	CH_3-2, H3-1, CH_3-1
NH-3	NH-2
NH-4	CH_3-1

Table 3. Intermolecular drug-DNA NOE contacts.

Drug	**DNA Proton**
FH	G5_B:H1, T1_A:H6
NH-1	G2_B:H1, T1_A:CH_3, T6_B:H1', T1_A:H6
NH-2	G5_B:H1', T1_A:H5", G2_A:H8, T1_A:H6, G3_A:H8
NH-3	G3_A:H8
NH-4	G2_B:H1', G3_B:H1'
H3-1	G5_B:H1', T1_A:H2", G2_A:H8
H3-3	G4_A:H8, G3_B:H1
H5-1	T1_A:H2'
CH_3-2	G2_A:H2', G5_B:H5"

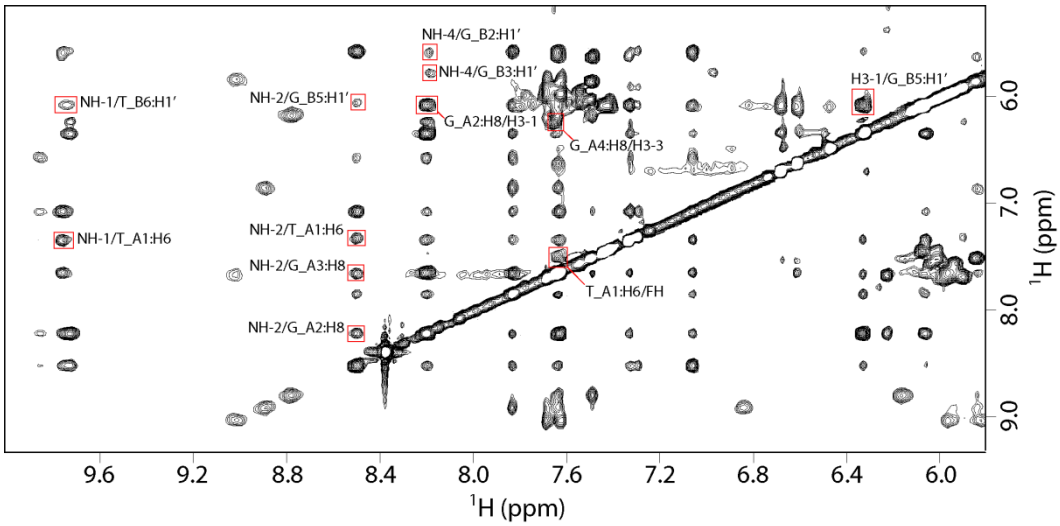

Fig. 3. Expanded region of the NOESY (700 MHz, 25°C, mt=100 ms) spectrum of the 4:1 complex Dist-A/[d(TGGGGT)]$_4$, displaying some of the intermolecular drug-DNA NOE contacts (boxed in red).

Structure calculations

In order to obtain the three-dimensional structure of the 4:1 complex at atomic level, an estimation of proton–proton distances has been retrieved from cross-peak intensities in 2D NOESY experiments (700 MHz, T=25°C). Pseudo-atoms were introduced where needed. A total of 440 distances were used for the calculations and, as suggested by the presence of eight guanine imino protons in the 1D ^{1}H-NMR spectrum, 32 supplementary distance restraints (HN1–O6, HN2–N7) for 32 hydrogen bonds corresponding to the four G-quartets were also incorporated during the computations (Table 4).

In agreement with NMR data, the backbone torsion angles □, β, □, □ and ε were restrained in the range -150°/-30°, -230°/-110°, 20°/100°, 95°/175° and -230°/-110°, respectively.[32] Further, glycosidic torsion angles χ were fixed in the *anti*-domain (-155°/-75°). Therefore, 3D structures which satisfy NOEs were constructed by simulated annealing (SA) calculations. An initial structure of the 4:1 complex was constructed possessing a random conformation and minimized. Restrained simulations were carried out for 300 ps using the CVFF (Consistent Valence Force Field).

Table 4: Experimental constraints and structure statistics of the best 10 structures.

Experimental constraints	
Total NOEs	440
NOEs from non-exchangeable protons	352
NOEs from exchangeable protons	88
Hydrogen bonds constraints	32
Dihedral angle constraints	96
Planarity constraints for G bases	64
CVFF energy (kcal mol^{-1}) of the minimized structures	
Total	1187.44 ± 3.78
nonbond	-148.27 ± 3.48
restraint	226.06 ± 2.14
NOEs violations	
Number > 0.2 Å	0.4 ± 0.1
Maximum (Å)	0.196 ± 0.028
Sum (Å)	3.026 ± 0.232
Average violation (Å)	0.007
r.m.s. deviations from the mean structure (Å)	
All heavy atoms	0.28 ± 0.069

The restrained SA (Simulated Annealing) calculations started at 500 K, and thereafter, the temperature was decreased stepwise down to 100 K. The aim stem was to energy-minimize and refine the structures. A total of 20 structures was generated. Average RMSD value of 0.28 ± 0.069 Å for the heavy atoms was obtained from the superimposition of the best 10 structures (Fig. 4).

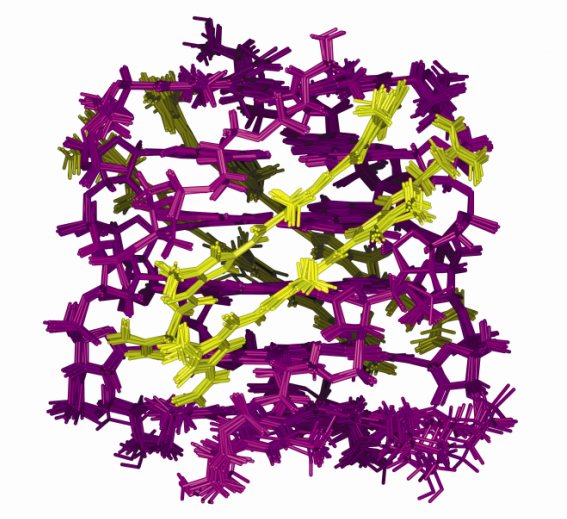

Fig. 4 Side view of the superimposition of the 10 best structures of the 4:1 complex Dist-A/[d(TGGGGT)]$_4$. Dist-A is reported in yellow and DNA is colored in magenta.

Isothermal Titration Calorimetry measurements

ITC was used to provide direct information about the interaction enthalpy (ΔH°), binding constant (K_b), and stoichiometry (n).[33] In addition, to fully characterize the thermodynamics of the binding reactions we determined the change in free energy (ΔG°) and entropy (T·ΔS°).

Figure 5 shows the ITC results for the titration of [d(TGGGGT)]$_4$ with Dist-A. The interaction heats were corrected for the heats of dilution associated with the addition of Dist-A into buffer solution. Fitting of the

corrected data allows ΔH°, K_b, and n to be determined (Table 5).

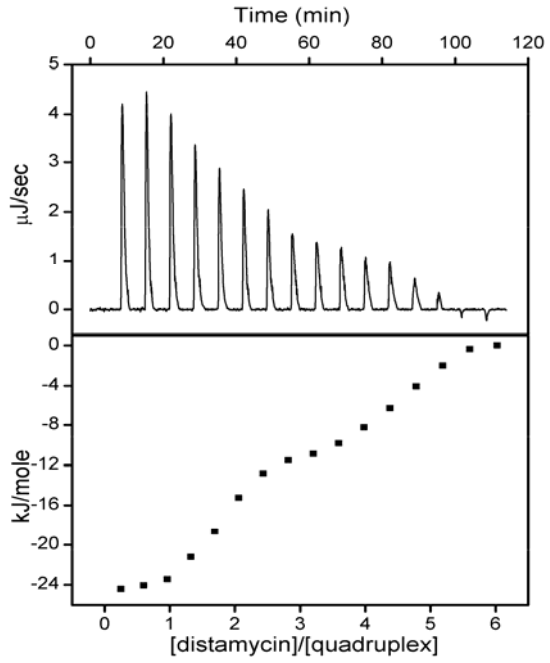

Figure 5. ITC data for the binding of Dist-A to $[d(TGGGGT)]_4$. Top panel shows the calorimetric response of 16 x 15 µl injections of 940 µM Dist-A to 30 µM $[d(TGGGGT)]_4$ at pH 7.3 and T=25 °C. Bottom panel shows the integrated injections heats for the above data.

ITC measurements show that the affinity of Dist-A toward $[d(TGGGGT)]_4$ follows the same behavior observed by NMR; there are two distinct binding events, with 2:1 and 4:1 drug:quadruplex stoichiometry, respectively. Examination of the thermodynamic data reveals that Dist-A/$[d(TGGGGT)]_4$ interactions are exothermic. K_b values of 4 x 10^5 M^{-1} and 4 x 10^6 M^{-1} have been found for 2:1 and 4:1 complexes, respectively.

Table 5 Thermodynamic parameters for the interaction of Dist-A with $[d(TGGGGT)]_4$ determined by ITC at 25°C.

n	K_b /10^6 x M^{-1}	$\Delta H°$/kJ mol^{-1}	$T\Delta S°$/kJ mol^{-1}	$\Delta G°$/kJ mol^{-1}
1.8 ± 0.2	0.4 ± 0.3	-8 ± 1	24 ± 2	-32 ± 2
4.2 ± 0.2	4.0 ± 3.0	-10 ± 1	27 ± 2	-37 ± 2

Discussion

In order to unambiguously assert that Dist-A interacts with the grooves of the quadruplex $[d(TGGGGT)]_4$, we have designed and synthesized a new modified oligonucleotide, namely $d(TGG^{Me}GGT)$. Particularly, we have taken advantage of the finding of a recent study of some of us[34] concerning the aptitude of 8-methyl-2'-deoxyguanosine (dG^{Me}) containing oligonucleotides to form quadruplex structures. According to the NMR data, the oligonucleotide $d(TGG^{Me}GGT)$ is actually able to fold into a quadruplex, and its structure is almost superimposable to that of $[d(TGGGGT)]_4$. Thus, each 8-methyl group of the four dG^{Me} of the quadruplex $[d(TGG^{Me}GGT)]_4$ faces right in to the very central region of the grooves, pointing outward the quadruplex. As a result, if Dist-A interacts with the groove of the quadruplex, the presence of these bulky groups in the very central region of the four grooves should prevent (or at least should limit) the insertion of Dist-A molecules, and consequently, the formation of a stable complex.

In the case of $[d(TGGGGT)]_4$, Dist-A displayed an high affinity towards the quadruplex as suggested by the appearance of a new set of signals during the NMR titration. Furthermore, Dist-A caused the loss of the

original four-fold symmetry of the free quadruplex. On the other hand, during the entire titration process of quadruplex $[d(TGG^{Me}GGT)]_4$, the four strands were magnetically equivalent and only a general change of the chemical shift resonances could be observed (Fig. 1). This behavior could be explained assuming that really the presence of the methyl groups in the central region of the grooves does affect the binding of Dist-A to the quadruplex, and hence that Dist-A interacts with the grooves of the quadruplex.

The interaction of Dist-A with the grooves of the quadruplex $[d(TGGGGT)]_4$ was further inferred by the analysis of the DF-STD data by using the binding mode indexes (BMIs)[11] (see Experimental Section). BMI is a numerical parameter that takes into account the relative STD intensities at different saturation frequencies, giving information on the binding contact of the ligand on the target molecule surface.

As for ethidium bromide, the expected BMI value would be ≈ 1, as found in the original DF-STD paper,[11] because this small ligand is a base-pairs intercalator[28,29] that establishes secondary electrostatic interactions by extending its two positively charged amino groups toward the phosphate backbone. In analogy to that observed for duplex DNA, the calculated BMI index for ethidium bromide suggests an intercalating binding mode (Table 1). The BMI' and BMI" values are consistent with the first one (BMI), confirming the well defined intercalated position of ethidium bromide.

The above described experimental data show that the DF-STD approach may be used for describing the intercalating mode of binding of ethidium bromide inside the quadruplex, suggesting that similar information could be derived for characterizing the general binding mode of the Dist-

A/quadruplex complex.

The experimental outcomes for Dist-A reveal that the ligand establishes simultaneous interactions with external backbone and aromatic protons of the bases (that face into the groove) (BMI = 0.8), because the saturation diffusion is similar either irradiating on aliphatic protons and on aromatic resonances. Moreover, the BMI value is in agreement with the same one that describes the well known interaction of Dist-A with a DNA duplex[11] as minor groove binder. On the other hand, the very low values of both BMI' and BMI" (Table 1) show that saturation spreads less than the STD effects obtained irradiating deoxyribose/backbone or purine/pyrimidine bases protons, indicating that Dist-A is closer to the backbone and to the base protons with respect to the imino ones, since these point inside the quadruplex core.

Finally, the whole set of NMR data, along with the structure calculations, confirmed definitively the binding mode of Dist-A with the quadruplex $[d(TGGGGT)]_4$, at least in the experimental condition used here. In particular, the presence of eighteen head to tail drug:drug contacts unambiguously indicates that the drug molecules bind to the quadruplex, two by two, with each term of the dimeric pairs with an antiparallel orientation and in close contact to its partner (as observed with duplex DNA)[6,7] (Fig. 6).

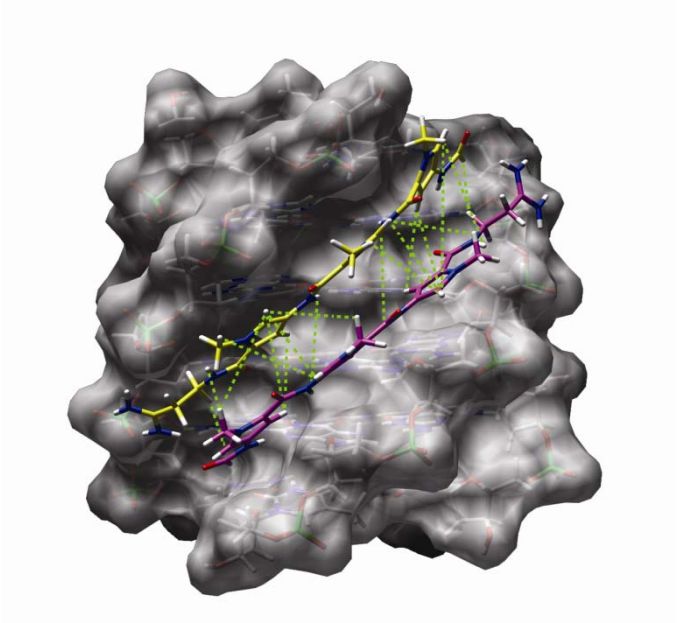

Fig. 6 Side view representation of the 4:1 complex Dist-A/[d(TGGGGT)]$_4$. Intermolecular head to tail drug-drug NOE contacts are indicated with green dashed lines.

Moreover, it is noteworthy that, throughout the whole NMR titration, a single set of signals was present for Dist-A protons, which only grew in intensity and did not show any significant change in chemical shift values by increasing drug concentration. This observation suggests that (i), even at low ligand:DNA stoichiometries (e.g., 0.5:1), simultaneous binding of two Dist-A side by side, in a highly cooperative mode, is dominant; (ii), both in the 2:1 and in the 4:1 complexes, the bound pair of Dist-A, reorientates itself in a fast process on the NMR time scale, similar to that observed for the binding of Dist-A to duplex DNA structures.[6] NOESY spectra acquired at different stage of the titration (data not shown), display the presence of the same head to tail drug:drug contacts observed for the 4:1 complex, and this further supports the above observations.

Furthermore, Dist-A binds two opposite grooves of the quadruplex as suggested also by the presence of NOE contacts only with the H8/H6 protons of strand A and H1' of strand B (Fig. 7).

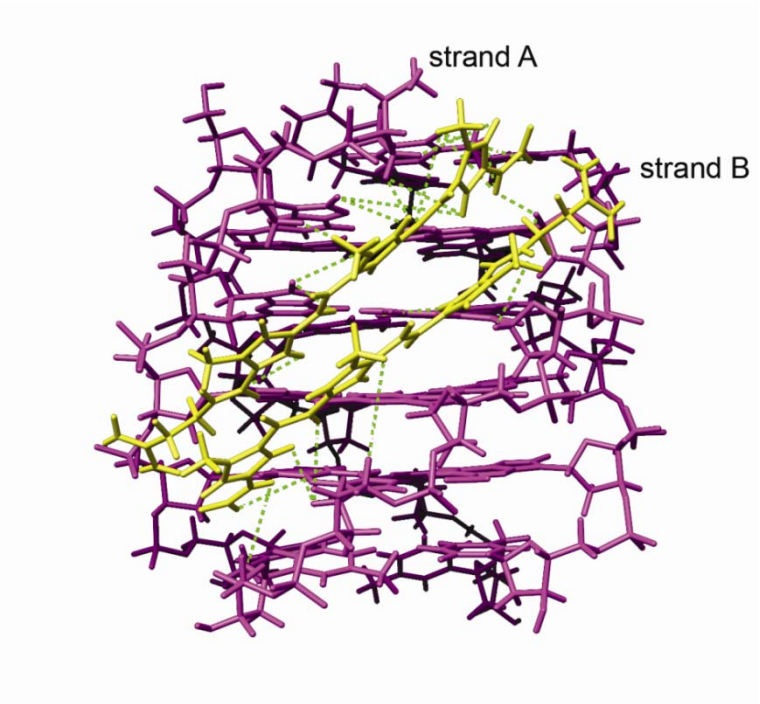

Fig 7 Side view representation of the of the 4:1 complex Dist-A/[d(TGGGGT)]$_4$. Dist-A and [d(TGGGGT)]$_4$ are reported in yellow and magenta, respectively. Drug-DNA NOE contacts are indicated with green dashed lines.

This explains the dyad symmetry of the final 4:1 complex, that is, therefore, consistent with a structure comprising two Dist-A dimers simultaneously spanning, in fast reorientation, two opposite grooves of the quadruplex.

In order to evaluate the changes of the DNA conformation upon binding of Dist-A, we have analyzed by CURVES[35,36] the helix parameters of the most representative structure of the complex (we used lowest energy structure after minimization), and the already reported NMR structure of [d(TTGGGGT)]$_4$ (PDB code: 139D) (see

Supplementary Information). The DNA quadruplex remains in the B-DNA family, although local variations are observed. The deoxyribose rings are predominantly in a S-type conformation, except for residues G2_Bs which possess a N-type conformation.

Furthermore, as suggested by lower *Twist* values and higher *Rise* values, the helical winding of the complexed DNA is smaller than that of $[d(TTGGGGT)]_4$. Notably, each Dist-A dimer expands its binding groove (similarly to that observed with duplex DNA),[6,7] with concomitant reduction of the size of the adjacent ones. In particular, the bounded grooves are 17.8 Å (~2 Å wider than that of the uncomplexed quadruplex), while the adjacent ones are 15.0 Å (~1 Å narrower) (Fig. 8). This, basically, prevents a further interaction with other Dist-A molecules.

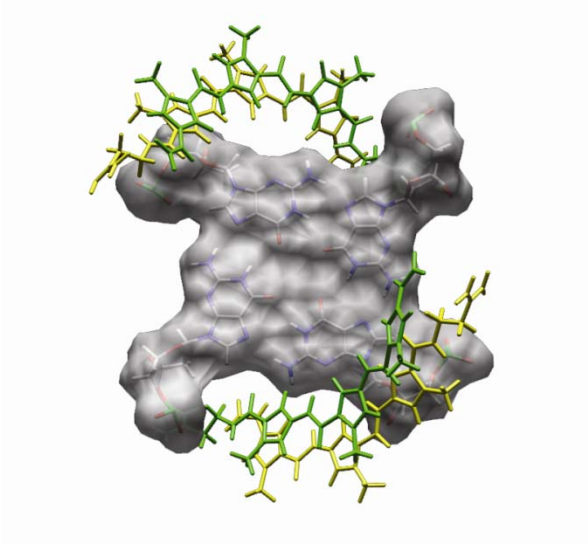

Fig. 8 Top view representation of the G3-tetrad showing the different size of the occupied and unoccupied grooves. Dist-A molecules are reported in yellow and green.

The Dist-A dimers span almost the entire length of the grooves. However, the Dist-A dimers are slightly shifted towards the 5' end of the

quadruplex. The peptide bonds and the *N*-methylpyrrole rings of Dist-A turned out to be planar, while the Dist-A molecules are twisted due to the flexibility of the links between these two planar units. The peptide-pyrrole hinges have similar twist angles of pyrrole-peptide hinges, ranging between 2°-26°.

The two staggered antiparallel Dist-A molecules overlap by nearly 90%, with each *N*-methylpyrrole ring of one molecule facing a peptide bond of the other one, thus making synchronized twisting to fit the curvature of the DNA quadruplex grooves easier. The preference for such orientation could lie in the electronic π-π interactions and in the energetically favorable dipole-dipole interactions, since the dipole moments of the pyrrole ring and the carbonyl group are exactly in opposite directions

The CPK model of the refined structures clearly shows the van der Waals complementarity between Dist-A dimers and the quadruplex grooves. In fact, the crescent shape of the ligand maximizes favorable interactions with the quadruplex by following the curvature of the grooves. Moreover, the negative value for van der Waals energy obtained for the calculated structures (see Table 4) indicates the absence of any unfavorable contacts in the proposed structures.

Electrostatic potential surface has been calculated and visualized with the programs Python,[23] APBS[24] and MSMS.[25] Figure 9 clearly indicates that the positively charged amidinium moiety interacts with the G5_A and G4_A phosphate groups of the quadruplex.

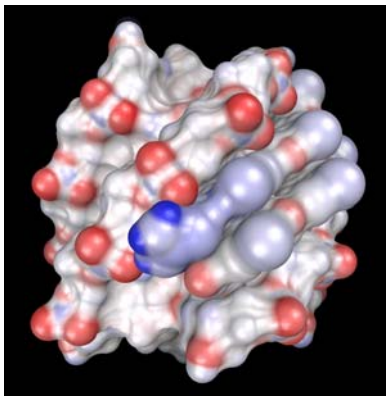

Fig. 9 Electrostatic potential surface of the representative structure. The surface is colored according to electrostatic potential, positive potential is shown in blue and negative potential is in red. Electrostatic interactions between amidinium group of distamycin and G5_A and G4_A phosphate groups of the quadruplex are clearly descernable.

Furthermore, the complex is also stabilized by four drug:DNA H-bonds: NH_2-2/ G5_A:O2P, NH_2-1/ G5_A:N2 and HF/ G2_B:N3. It is interesting to note that, when Dist-A binds the duplex DNA,[7] four peptide NH groups of the drug make four hydrogen bonds with N3 (purine) or O2 (pyrimidine) base atoms of DNA. In our case no H-bond involving NH groups has been observed, and this can be explained taking into account that the quadruplex and duplex grooves are chemically different.

Finally, ITC measurements show that the affinity of Dist-A toward $[d(TGGGGT)]_4$ follows the same behavior observed by NMR; there are two distinct binding events, with 2:1 and 4:1 drug:quadruplex stoichiometry, respectively. The affinity between Dist-A and $[d(TGGGGT)]_4$ is enhanced when the ratio of Dist-A to the quadruplex is increased, as the equilibrium constant for the second interaction event is ~10-fold greater than that for the first event. It is interesting to note that interaction between Dist-A and the quadruplex is entropically

driven, being the -TΔS value higher than ΔH value. These data confirm once again the groove binding of Dist-A, in fact, a recent analysis made on a number of DNA binding agents revealed that groove-binding is entropically driven, while intercalation is enthalpically driven.[37]

As shortly mentioned in the introduction, Dist-A binds in very different way the quadruplexes $[d(TGGGGT)]_4$[8] and $[d(TAGGGTTA)]_4$ (and its analogues)[9]. As demonstrated in this paper, Dist-A binds as dimers in two of the four grooves of the quadruplex $[d(TGGGGT)]_4$, whereas, in the case of $[d(TAGGGTTA)]_4$, two molecules of distamycin A extend over each of the two G-tetrad planes in a 4:1 binding mode. Interestingly, as suggested by the NMR titration profiles, Dist-A seems to possess an higher affinity towards $[d(TGGGGT)]_4$ than $[d(TAGGGTTA)]_4$. In fact, throughout the titration of $[d(TAGGGTTA)]_4$, the signals underwent only to a general change in chemical shift resonances and no appearance of a new set of DNA signals was observed. Apparently, there is no specific reason to justify this very different binding mode of Dist-A. Nevertheless, it is known[38] that when a parallel quadruplex structure contains a TA step at the 5' edge of the sequence, that portion of the complex is less structured. Interestingly, in the structure reported here, Dist-A dimers indeed are not perfectly centered into the quadruplex grooves, but they are slightly shifted towards the 5'-edge. This could mean that Dist-A preferentially recognizes that part of the quadruplex and that, in the case of $[d(TAGGGTT)]_4$ (and its analogues), it is not able to recognize the 5'-end, that is unstructured. An alternative explanation, could be that the interaction of Dist-A with the grooves of the quadruplex $[d(TAGGGTT)]_4$ could be prevented by motion of the unstructured TA moiety.

In summary, we have proved for the first time that the grooves of a quadruplex structure can be recognized by a small organic molecule. In fact we have demonstrated that distamycin A is able to bind the grooves of the quadruplex $[d(TGGGGT)]_4$. So far, groove binding has been a useful way to selectively recognize only duplex DNA. The results reported here may stimulate the design of new quadruplex groove binders hopefully capable to discriminate between duplex and quadruplex DNA grooves. Then, since groove dimensions vary according to the type of quadruplex,[39] groove binding could also offer the opportunity for obtaining increased selectivity for a particular quadruplex structure.

Notes and references

1. Neidle, S.; Parkinson, G.N. *Curr. Opin. Struct. Biol.* **2003**, 13, 275-283.
2. Keniry, M.A. *Biopolymers*. **2001**, 56(3), 123-126.
3. Kelland, L.R. *Eur. J. Cancer*. **2005**, 41, 971-979.
4. Aboul-ela, F.; Murchie, A.I.H.; Norman, D.G.; Lilley, D.M.J. *J. Mol. Biol.* **1994**, 243, 458-471.
5. Laughlan, G.; Murchie, A.I.; Norman, D.G.; Moore, M.H.; Moody, P.C.; Lilley, D.M.; Luisi, B. *Science*. **1994,** 265, 520-524.
6. Pelton, J.G.; Wemmer, D.E. *J. Am. Chem. Soc.* **1990**, 112, 1393-1399.
7. Chen, X.; Ramakrishnan, B.; Sundaralingam, M. *J. Mol. Biol.* **1997**, 267, 1157-1170.
8. Randazzo, A.; Galeone, A.; Mayol, L. *Chem. Commun.* **2001**, 11, 1030-1031.
9. Cocco, M.J.; Hanakahi, L.A.; Huber, M.D.; Maizels, N. *Nucleic Acids Res.* **2003**, 31, 2944-2951.
10. Moore, M.J.B.; Cuenca, F.; Searcey, M.; Neidle, S. *Org.& Biomol.*

11. Di Micco, S.; Bassarello, C.; Bifulco, G.; Riccio, R.; Gomez-Paloma, L. *Angew. Chem. Int. Ed.* **2006**, 45, 224-228.
12. Kohda, K.; Tsunomoto, H.; Minoura, Y.; Tanabe, K.; Shibutani, S. *Chem. Res. Toxicol.* **1996**, 9, 1278-1284.
13. Hwang, T.L.; Shaka, A.J. *J. Magn. Res.* **1995**, A112, 275-279.
14. Dalvit, C.*J. Biomol. NMR.* **1998**, 11, 437-444.
15. Jeener, J.; Meier, B.; Bachmann, H.P.; Ernst, R.R. *J. Chem. Phys.* **1979**, 71, 4546-4553.
16. Piotto, M.; Saudek, V.; Sklenar, V.J. *J. Biomol. NMR.* **1992**, 2, 661-665.
17. Sklenar, V.; Piotto, M.; Leppik, R.; Saudek, V. *J. Magn. Reson.* **1993**, 102, 241-245.
18. Braunschweiler, L.; Ernst, R.R. *J. Magn. Reson.* **1983**, 53, 521-528.
19. Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1989**, 85, 393–399.
20. Weiner, S.J.; Kollman, P.A.; Case, D.A.; Singh, U.C.; Ghio, C.; Alagona, G.; Profeta, S.; Weimer, P. *J. Am. Chem. Soc.* **1984**, 106, 765-784
21. Martino, L.; Virno, A.; Randazzo, A.; Virgilio, A., Esposito, V.; Giancola, C.; Bucci, M., Cirino, G., Mayol, L. *Nucleic Acids Res.*, **2006**, 34(22), 6653-6662.
22. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. *J. Comput. Chem.* **2004**, 25(13), 1605-1612.
23. Michel, F.S. *J. Mol. Graphics Mod.* **1999**, 17, 57-61.
24. Baker, N.A.; Sept. D.; Joseph, S.; Holst. M.J.; McCammon, J.A. *Proc. Natl. Acad. Sci. USA* **2001**, 98, 10037-10041.
25. Sanner, M.F.; Spehner, J.C.; Olson, A.J. *Biopolymers.* **1996**, 38(3), 305-

320.

26. Freire, E.; Mayorga, O. L.; Straume, M. *Anal. Chem.* **1990**, 62, 950-959.
27. Mayer, M.; Meyer, B. *Angew. Chem. Int. Ed.*,**1999**, 38, 1784-1788.
28. Mayer, M.; Meyer, B. *J. Am. Chem. Soc.* **2001**, 123, 6108-6117.
29. Luedtke, N.W.; Liu, Q.; Tor, Y. *Chem. Eur. J.* **2005**, 11, 495-508.
30. Guo, Q.; Lu, M.; Marky, L.A.; Kallenbach, N.R. *Biochemistry*, **1992**, 31, 2451-2455.
31. Fernández, C.; Szyperski, T.; Ono, A.; Iwai, H.; Tate, S.; Kainosho, M.; Wüthrich, K. *J. Biomol. NMR.* **1998,** 12, 25-37.
32. Wüthrich, K. NMR of Protein and Nucleic Acids **1986**, Wiley, NY
33. Haq, I.; Ladbury, J. *J. Mol. Recognit.* **2000**, 13, 188-197.
34. Virgilio, A.; Esposito, V.; Randazzo, A.; Mayol, L.; Galeone, A. *Nucleic Acids Res.* **2005**, 33(19), 6188-6195.
35. Lavery, R.; Sklenar, H. *J. Biomol.Struct.Dyn.* **1988**, 6, 63-91.
36. Lavery, R.; Sklenar, H. *J. Biomol.Struct.Dyn.* **1989**, 6, 655-667.
37. Chaires, J.B. *Arch. Biochem. Biophys.* **2006**, 453, 26-31.
38. Patel, P.K.; Koti, A.S.R.; Hosur, R.V. *Nucleic Acids Res.* **1999**, 27, 3836-3843.
39. Parkinson, G.N. In *Quadruplex Nucleic Acids*, Neidle, S.; Balasubramanian, S. Eds.; RSC Publishing, London, **2006**; ch. 1, pp. 1-30.

2. Targeting DNA quadruplexes with distamycin A and its derivatives: an ITC and NMR study

Nucleic acids reversibly interact with a broad range of small organic molecules, such as intercalators or groove binders, which represent one of the most important lines of drug development and of current chemotherapy against cancer, viral, and some parasitic diseases [1]. Even though these compounds are known to bind to double-stranded DNA, several investigations involving the interaction of small ligands with unusual structures of nucleic acids have been reported [2], including quadruplex forms [3]. G-quadruplex structures comprise stacks of G-tetrads, which are the planar association of four guanines in a cyclic Hoogsteen hydrogen-bonding arrangement [4]. A multitude of proteins have been shown to interact with these structures which, particularly, are implicated in the molecular biology of telomeres [5]. Telomeres protect the end of the chromosome from damage and recombination and its shortening has been implicated in cellular senescence [6, 7]. Telomeric DNA consists of tandem repeats of simple sequences that are short and rich in guanine residues [8, 9]. In the presence of metal ions such as K^+ or Na^+, telomeric DNA can form quadruplex structures [10]. The telomerase, the enzyme which synthesizes the G-rich strand of telomeric DNA, is active in about 85% of tumors, leading the cancer cells to infinite lifetime, whereas telomerase activity is basically absent in somatic cells [11]. The inhibition of telomerase has become an attractive strategy for the anticancer approach [12] and, because telomerase requires a single-stranded telomeric primer [13], the formation of G-quadruplex

complexes by telomeric DNA inhibits the telomerase activity [14]. Furthermore, small molecules that stabilize G-quadruplex structures have been found to be effective telomerase inhibitors and, then, the use of drugs to target G-quadruplexes is emerging as a promising way to interfere with telomere replication in the tumors cells and to act as anticancer agents [15, 16].

Distamycin A (Dist-A) is a small molecule, which is known to bind A and T rich regions of duplex DNA in an antiparallel dimeric form [17]. Dist-A has also been shown to interact with four-stranded parallel DNA quadruplex containing oligonucleotides of different sequence in K^+ buffer solutions [18-20]. Two opposite models have been proposed for the distamycin-quadruplex complex: the first, proposed by some of us, suggests that distamycin molecules bind as dimers in two opposite grooves of quadruplex $[d(TGGGGT)]_4$ [18, 19]; the second one, proposed by Maizel and coworkers [20], suggests that two distamycin molecules stack on the terminal G-tetrad planes of the quadruplexes $[d(TAGGGTTA)]_4$, $[d(TAGGGGTT)]_4$ and $[d(TAGGGGGT)]_4$. The different binding behaviours observed indicate that a sequence-dependent interaction takes place in the recognition process.

All these findings have stimulated other investigations aimed to shed light onto the recognition processes between small organic molecules and quadruplex structures. Since derivatives of Dist-A have been reported to be inhibitors of the human telomerase enzyme [21], and since some distamycin analogues have been found to interact with a DNA quadruplex of sequence $d[GGG(TTAGGG)_3]$ [22], we have focused our attention on the interaction of Dist-A and its two derivatives (compounds 1 and 2, Fig. 1), with quadruple helices of different sequence and molecularity. Compounds 1 and 2 are two carbamoyl

analogues of distamycin A, containing four and five pyrrole units, respectively.

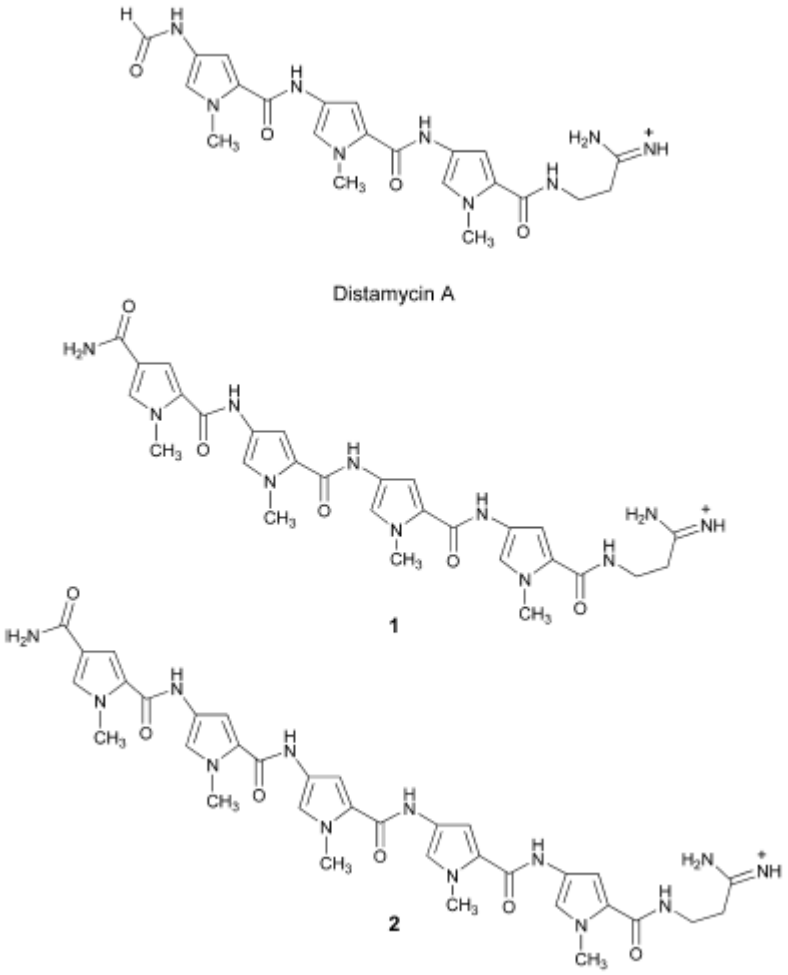

Fig. 1. Chemical structures of the compounds used in this study.

We have already reported a very preliminary ITC study dealing with the interaction of Dist-A and of compound 1 with the parallel quadruplex structure $[d(TGGGGT)]_4$ in Na^+ containing solution [23]. Surprisingly, Dist-A displayed a very different stoichiometry and titration behaviour with respect to K^+ buffer condition. These results strongly indicate that, as also observed in other study [24], the nature of the ions present in solution does affect the drug/quadruplex recognition processes.

Therefore, in order to get further insight into the influence of the ions

on the binding behaviour of Dist-A and its derivatives, here we report an NMR study on the interaction of the quadruplex $[d(TGGGGT)]_4$ with Dist-A in Na^+ buffer, and with compounds 1 and 2 in both Na^+ and K^+ buffers. The study is also completed by the ITC analysis of the binding of 1 to $[d(TGGGGT)]_4$ in K^+ buffer. Finally, we also report ITC data about the interaction of Dist-A, 1 and 2 with the human telomeric sequence $d[AG_3(T_2AG_3)_3]$ in both K^+ and Na^+ buffers.

.Materials and methods

The oligonucleotides d(TGGGGT) and $d[AG_3(T_2AG_3)_3]$ were synthesized on a Millipore Cyclone Plus DNA synthesizer using solid phase β-cyanoethyl phosphoramidite chemistry at 15 μmol scale. The oligomers were detached from the support and deprotected by treatment with concentrated aqueous ammonia at 55°C for 12 h. The combined filtrates and washings were concentrated under reduced pressure, redissolved in H_2O, analyzed and purified by high-performance liquid chromatography (HPLC) on a Nucleogel SAX column (Macherey–Nagel, 1000-8/46); using buffer A: 20 mM KH_2PO_4/K_2HPO_4 aqueous solution (pH 7.0), containing 20% (v/v) CH_3CN; buffer B: 1 M KCl, 20 mM KH_2PO_4/K_2HPO_4 aqueous solution (pH 7.0), containing 20% (v/v) CH_3CN; a linear gradient from 0 to 100% B for 30 min and flow rate 1 ml/min were used. The fractions of the oligomer were collected and successively desalted by Sep-pak cartridges (C-18). The isolated oligomers proved to be >98% pure by NMR.

DNA solutions were prepared by dissolving solid lyophilized oligonucleotides in buffered solutions containing either K^+ or Na^+. The K^+ buffer used was 10 mM potassium phosphate, 70 mM KCl, 0.1 mM EDTA; the Na^+ buffer was 10 mM sodium phosphate, 70 mM NaCl, 0.1

mM EDTA; both buffer solutions were at pH= 7.0. DNA quadruplexes were formed by heating the solutions to 90 °C for 5 min. The solutions were then cooled slowly to room temperature and equilibrated for one day at 4 °C. The oligonucleotides concentrations were determined by UV adsorption measurements at 90°C using molar extinction coefficient values $\varepsilon_{(260\ nm)}$ of 57800 and 234500 $M^{-1}cm^{-1}$ for d(TGGGGT) and d[$AG_3(T_2AG_3)_3$], respectively. The molar extinction coefficients were calculated by the nearest neighbour model [25].

Compound 1 and 2 have been synthesized as reported in the literature [26], while Dist-A has been purchased from Sigma-Aldrich (St. Louis, MO). The purity of Dist-A and its derivatives has been checked by 1D ^{1}H-NMR and all the samples turned out to be more than 95% pure. Ligand solutions were freshly prepared every time before binding experiments by dissolving the drugs in the same K^+ or Na^+ buffer used for DNA quadruplexes.

Isothermal Titration Calorimetry

ITC experiments were performed on a CSC 4200 Calorimeter from Calorimetry Science Corporation (Lindon, Utah) with a cell volume of 1.3 mL. Calorimetric titrations were carried out at 25 °C, by injecting 15μL aliquots of a 300-940 μM ligand solution into a 30-50 μM quadruplex solution at 400 seconds intervals, with stirring at 297 rpm, for a total of 16 injections.

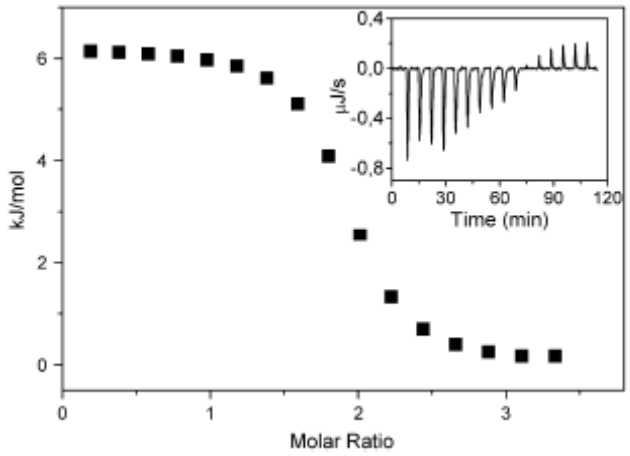

Fig. 2. Raw ITC data (inset) and binding isotherm for the interaction of compound 1 with $[d(TGGGGT)]_4$ using the buffered solution containing K^+ ions.

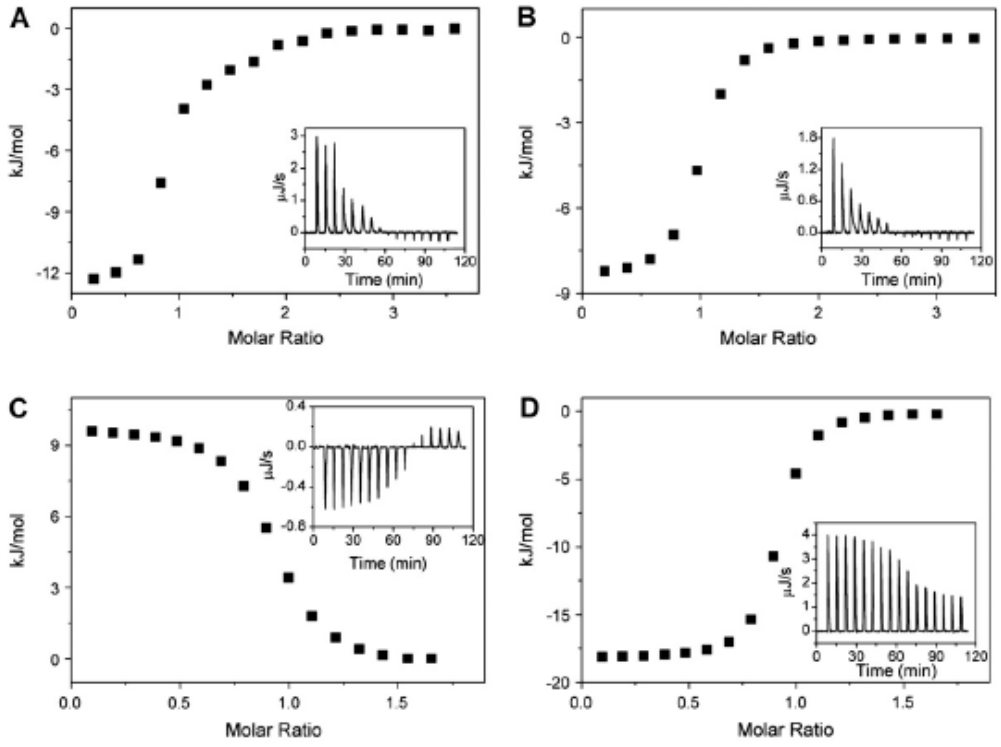

Fig. 3. ITC data curves (insets) and binding isotherms for titration of $d[AG_3(T_2AG_3)_3]$ with Dist-A in K^+ (A) and Na^+ (B) buffer and with compound 1 in K^+ (C) and Na^+ (D) buffer.

ITC measures heat generated or absorbed upon binding, and provides the values of the binding constant (K_b), the stoichiometry (n) and the enthalpy change ($\Delta H°$) [2]. The K_b value then permits calculation of the

change in Gibbs energy (ΔG°), which together with ΔH° allows the calculation of the entropic term T ·ΔS°. The separation of the free energy of binding into enthalpic and entropic components allows to reveal the nature of the forces that drive the binding reaction.

Inserts in the Figures 2 and 3 show typical ITC profiles for the binding of ligands to the quadruplexes. Each of the heat burst curves in the inserts corresponds to a single drug injection. The areas under these heat burst curves were determined by integration to yield the associated injection heats. The interaction heats were corrected for the heat of dilution associated with each ligand, independently determined by injecting ligand solution into the buffers. The heat of dilution for the distamycin was found to be slightly endothermic, while the dilution of the derivatives was moderately exothermic.

The corrected heat values are plotted as a function of the molar ratio (Fig. 2 and 3), to give the corresponding binding isotherms. The resulting isotherms were then fitted to a single set of identical sites model using the Bindwork program supplied with the instrument, to give the binding enthalpy (ΔH°), equilibrium binding constant (K_b), and stoichiometry (n). The remaining thermodynamic parameters, ΔG° and TΔS°, were derived using the standard relationships $\Delta G° = -RT \ln K_b$ and $T\Delta S° = \Delta H° - \Delta G°$.

Nuclear magnetic resonance experiments

NMR samples were prepared at a concentration of 2 mM, in 0.6 ml (H_2O/D_2O 9:1) buffer solution having 10 mM KH_2PO_4/K_2HPO_4, 70 mM KCl, 0.1 mM EDTA, pH 7.0 or 10 mM NaH_2PO_4/Na_2HPO_4, 70 mM NaCl, 0.1 mM EDTA, pH 7.0. NMR spectra were recorded with Varian ^{Unity}INOVA 700 MHz spectrometer. ^{1}H chemical shifts were

referenced relative to external sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS). 1D proton spectra of samples were recorded using pulsed-field gradient DPFGSE for H_2O suppression [27, 28].

Results

Titration of $[d(TGGGGT)]_4$ with distamycin A

The interaction between Dist-A and $[d(TGGGGT)]_4$ in the K^+ containing solution has already been studied using both NMR and ITC techniques [19]. Interestingly, below 2:1 ligand:quadruplex stoichiometry, Dist-A binds the quadruplex to form short-lived complexes on the NMR time scale. At higher drug:DNA ratios, Dist-A turned out to tightly and specifically bind the quadruplex, to give a 4:1 complex. All results unequivocally indicated that, in the 4:1 complex, two antiparallel distamycin dimers (as observed with duplex DNA [17]) bind simultaneously two opposite grooves of the quadruplex. The thermodynamic parameters determined by ITC confirm these finding and additionally indicate that Dist-A/$[d(TGGGGT)]_4$ interactions are entropically driven processes (Table 1).

When the same titration is performed in a Na^+ containing buffer (using the same ionic strength), the binding behaviour turned out to be surprisingly different. In fact, the already reported thermodynamic binding parameters [23], revealed an association constant of 2×10^5 M^{-1} that is lower than the constants observed in K^+ buffer and a complex stoichiometry of 1:1 with respect to 4:1 observed in K^+ buffer (Table 1).

Table 1 Thermodynamic parameters for the interaction of Dist-A and 1 with $[d(TGGGGT)]_4$ determined by ITC at 25 °C and pH 7.0

Buffer	Ligand	n	K_b ($\times 10^6$ M^{-1})	$\Delta_b H°$ (kJ/mol)	$T\Delta_b S°$ (kJ/mol)	$\Delta_g G°$ (kJ/mol)
K^+	Dist-A[a,b]	1.8 ± 0.2	0.4 ± 0.3	−8.0 ± 1.0	24 ± 2	−32 ± 2
		4.2 ± 0.2	4.0 ± 3.0	−10.0 ± 1.0	27 ± 2	−37 ± 2
	1	1.9 ± 0.1	2.0 ± 1.0	7.0 ± 0.5	43 ± 1	−36 ± 1
Na^+	Dist-A[c]	1.0 ± 0.1	0.2 ± 0.1	−14.0 ± 0.5	16 ± 1	−30 ± 1
	1[c]	0.9 ± 0.1	2.3 ± 1.0	10.0 ± 0.5	46 ± 1	−36 ± 1

[a] Two binding events.

[b] From Ref. [19].

[c] From Ref. [23].

These data are here corroborated by the NMR data obtained following, step by step, the titration in Na^+ buffer solution (Fig. 4).

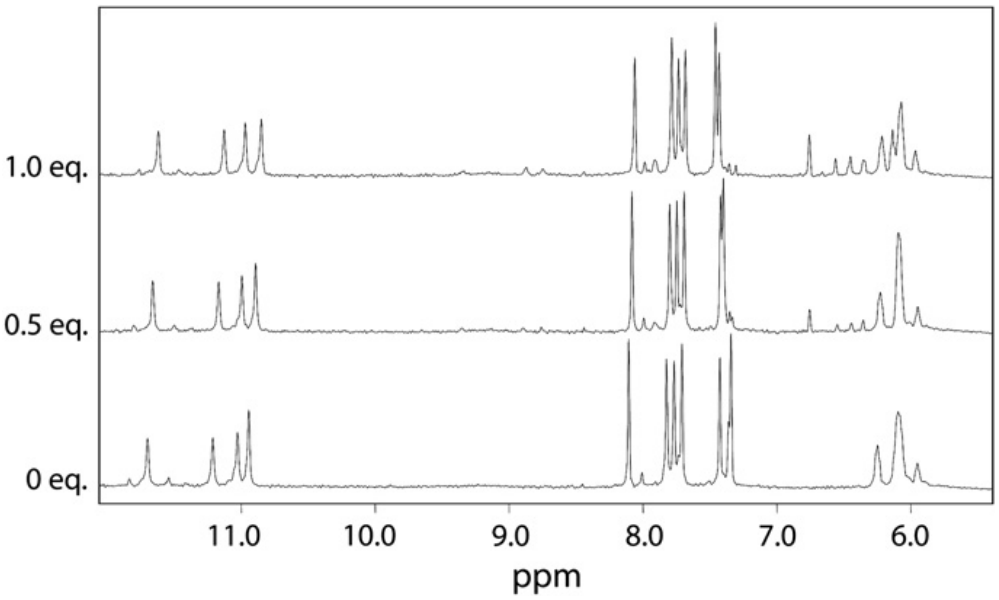

Fig. 4. Titration with Dist-A of $[d(TGGGGT)]_4$ containing 2 mM NaH_2PO_4, 0.2 mM EDTA and 70 mM NaCl, pH 7.0, in 9:1 H_2O/D_2O at 25°C (700 MHz). The drug equivalents are shown along the side of the spectra.

Thus, the addition of Dist-A to $[d(TGGGGT)]_4$ caused gradual changes in the chemical shift of DNA proton resonances. At ligand:DNA ratio of 1:1 the titration was virtually completed. The four strands resulted to be

magnetically equivalent throughout the titration, and no splitting of resonances was observed at any stage. This behaviour can be explained assuming that the ligand is in fast exchange on the NMR time scale with its binding sites on the quadruplex. This fast exchange behaviour could not be changed to a slow or intermediate regime by altering the temperature of the system.

As observed for the titration performed with a K^+ containing buffer, a single set of signals was present for Dist-A protons, which only grew in intensity and did not show any significant change in chemical shift values by increasing drug concentration.

Titration of $[d(TGGGGT)]_4$ with 1

The addition of 1 to the quadruplex (Fig. 5) caused gradual changes in the chemical shift and a broadening of DNA proton resonances in the ^{1}H-NMR spectra, using both K^+ or Na^+ buffer solutions. The titrations were virtually completed at ligand:DNA ratio of 2:1 and 1:1 for K^+ or Na^+ buffer solutions, respectively. The four strands resulted to be magnetically equivalent throughout in both titrations, since no splitting of resonances was observed at any stage.

As far as the ITC measurements are concerned, the raw data for the titration of compound 1 with $[d(TGGGGT)]_4$ in the K^+ containing solution (Fig. 2) indicate an endothermic interaction, based on the negative values observed for the peaks. With each injection of ligand, less and less heat uptake was observed until constant values were obtained (corresponding to the heat of dilution), reflecting a saturable process. The thermodynamic parameters determined for the interaction with $[d(TGGGGT)]_4$ are collected in Table 1, along with the already reported parameters for an easy comparison. The analysis of the data

shows that compound 1-$[d(TGGGGT)]_4$ interaction is characterized by a $K_b = 2 \times 10^6\ M^{-1}$. The binding stoichiometry of 2:1 (drug:quadruplex) has also been confirmed. Moreover, the binding of 1 to the $[d(TGGGGT)]_4$ quadruplex was characterized by a moderately large and positive enthalpy of binding (7.0 kJ mol^{-1}).

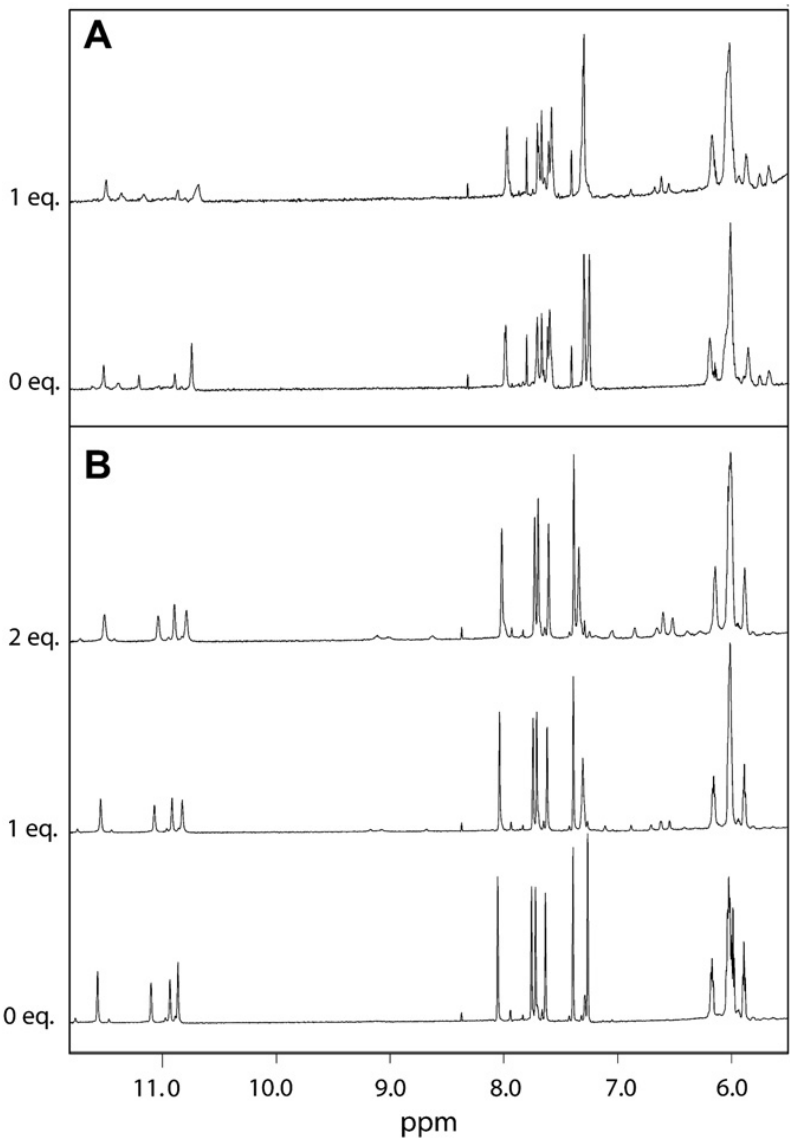

Fig. 5. Titration with 1 of $[d(TGGGGT)]_4$ containing 2 mM NaH_2PO_4 and 70 mM NaCl, pH 7.0, in 9:1 H_2O/D_2O at 25°C (700 MHz) (A) and containing 2 mM KH_2PO_4 and 70 mM KCl, pH 7.0, in 9:1 H_2O/D_2O at 25°C (700 MHz) (B). The drug equivalents are shown along the side of the spectra.

Except for stoichiometry, similar results were obtained for the

interaction of compound 1 with $[d(TGGGGT)]_4$ in Na^+ solution (Table 1). The positive heat of formation of the 1-$[d(TGGGGT)]_4$ complex indicates that this interaction is an enthalpically unfavourable (entropically driven) process in both solution conditions.

Titration of $[d(TGGGGT)]_4$ with 2

NMR titration of $[d(TGGGGT)]_4$ with compound 2 in both K^+ and Na^+ buffer solutions (Fig. S1, Supplementary Material) led only to a slight broadening of the NMR signal attributable to T6 H6. No drift of the resonances was observed, not even for T6 H6 signal. At the same way, in the ITC measurements, resolvable binding isotherms were never obtained for the compound 2 in both solution conditions (Na^+ and K^+) using any combination of reactant concentration.

Titration of $d[AG_3(T_2AG_3)_3]$ with distamycin A, 1 and 2

Figure 3 shows the results of calorimetric titrations of distamycin and compound 1 with the $d[AG_3(T_2AG_3)_3]$ quadruplex in both K^+ and Na^+ buffers. As for $[d(TGGGGT)]_4$, resolvable binding isotherms were never obtained for compound 2, suggesting, also in this case, a poor affinity of the molecule for the quadruplex.

The values of the binding constants and the Gibbs energy changes from ITC measurements (Table 2) indicate that the associations between the Dist-A and compound 1 with the $d[AG_3(T_2AG_3)_3]$ quadruplex are favoured, at 25 °C, from thermodynamic point of view. Although the binding stoichiometry is 1:1 for all interactions, the binding behaviours in the Na^+ solution were found to be considerably different compared to the equivalent K^+ solutions, with a different ligand affinity and/or enthalpy of binding. Indeed, although the affinity of compound 1 for the

quadruplex is higher than that of Dist-A, in both solution conditions, the equilibrium constants for the interactions of Dist-A and 1 with d[$AG_3(T_2AG_3)_3$] in Na^+ buffer (K_b = 1 x 10^6 M^{-1} and 6 x 10^6 M^{-1}, respectively) are higher than the corresponding constants in the K^+ buffer (K_b = 6 x 10^5 M^{-1} and 2 x 10^6 M^{-1}, respectively). The values of ΔH° and ΔS° show that, in all cases, the binding processes are entropically driven; however, the interaction of distamycin with d[$AG_3(T_2AG_3)_3$] is always associated with a favourable binding enthalpy (-9 and -14 kJ mol^{-1} for Na^+ and K^+ solutions, respectively), whereas the binding of 1 is exothermic (ΔH° = -18 kJ mol^{-1}) in the Na^+ buffer and moderately endothermic (ΔH° = 9 kJ mol^{-1}) in the K^+ solution.

Table 2 Thermodynamic parameters for the interaction of Dist-A and 1 with d[AG3(T2AG3)3] determined by ITC at 25 _C and pH 7.0

Buffer	Ligand	n	K_b ($\times 10^6$ M^{-1})	$\Delta_b H°$ (kJ/mol)	$T\Delta_b S°$ (kJ/mol)	$\Delta_g G°$ (kJ/mol)
K^+	Dist-A	0.9 ± 0.1	0.6 ± 0.3	−14.0 ± 2.0	19 ± 2	−33 ± 1
	1	1.0 ± 0.1	2.0 ± 1.0	9.0 ± 1.0	45 ± 1	−36 ± 1
Na^+	Dist-A	1.0 ± 0.1	1.0 ± 0.5	−9.0 ± 1.0	25 ± 1	−34 ± 1
	1	0.9 ± 0.1	6.0 ± 3.0	−18.0 ± 1.0	21 ± 1	−39 ± 1

Unfortunately the NMR analysis of these complexes is complicated from the fact that the ^{1}H-NMR spectrum of d[$AG_3(T_2AG_3)_3$] acquired using K^+ solution is characterized by a severe line broadening that prevent the detection of any structural information. Furthermore, since the question about the structure adopted by d[$AG_3(T_2AG_3)_3$] in a K^+ containing solution is still open [29-34], with the probability that in K^+ solution d[$AG_3(T_2AG_3)_3$] exists as a mixture of a mixed-parallel/antiparallel quadruplex and another structure, we have preferred to postpone this analysis to a dedicated study, where all the known structures adopted in solution by the human telomeric sequence will be taken in due account.

Discussion

Numerous structural and thermodynamic studies have been carried out in recent years in order to characterize the binding of quadruplex-interactive compounds [24, 35-38]. Besides those, here we report the calorimetric and NMR studies of the interaction of distamycin A and its derivatives 1 and 2 with $[d(TGGGGT)]_4$ and $d[AG_3(T_2AG_3)_3]$ quadruplexes. Experiments reveal that distamycin A and compound 1 bind the investigated quadruplexes; conversely the compound 2 appears to have a very poor affinity in any case.

The thermodynamic parameters determined by ITC (Tables 1 and 2) indicate that the interactions of both Dist-A and 1 with quadruplexes are entropically driven processes in both solution conditions. However, calorimetric and NMR experiments reveal a notable alteration in the binding behaviour of the drugs upon switching from the K^+ to Na^+ buffer.

Our recent study revealed that the interaction of distamycin with $[d(TGGGGT)]_4$ in the K^+ buffer give rise to a biphasic binding isotherm with a curvature typically observed in the presence of two binding sites with different affinities [19]. Indeed, the ITC data were best fit using a two-site model whose thermodynamic parameters are listed in Table 1. The binding stoichiometry is 2:1 and 4:1 (drug/quadruplex) for the first and second binding event, respectively. Interestingly, the affinity between Dist-A and $[d(TGGGGT)]_4$ in the K^+ buffer is enhanced when the ratio of distamycin to the quadruplex is increased, as the equilibrium constant for the second interaction event is ~10-fold greater than that for the first event. In the NMR experiments the affinity of distamycin toward $[d(TGGGGT)]_4$ follows the same behaviour observed by ITC.

The binding behaviour of distamycin in the Na^+ containing solution was

found to be surprisingly different compared to the equivalent K^+ containing solution. In this case, ITC and NMR experiments show that there is only one binding event with a stoichiometry of 1:1 and that the affinity of Dist-A is lower than the ones observed in K^+ buffer.

In the case of the interaction of compound 1 with $[d(TGGGGT)]_4$, similar results were obtained in both solution conditions, except for stoichiometry. Indeed, the ITC and NMR results obtained in K^+ solution reveal a stoichiometry of 2:1 (drug/quadruplex), whereas in the presence of Na^+ the stoichiometry is 1:1.

Interestingly, as shown in Table 1, the association processes of both ligands with $[d(TGGGGT)]_4$ are always entropically driven, even if the direct ITC measurements of the binding enthalpies indicate that the enthalpic contribution for the binding of Dist-A favours the interaction, whereas for compound 1 the major entropic binding forces are countered by a small unfavourable enthalpic term.

In general, the binding of an intercalator to DNA is entirely driven by a large favourable enthalpy change but with an unfavourable entropy decrease, and the binding of a groove binder to DNA is driven by a large favourable increase in entropy [39]. This observation has been further inferred by J. B. Chaires, that have recently reported the enthalpic and entropic contributions to the DNA binding of representative DNA binding compounds [40]. Interestingly, the entropically driven interactions of compound 1 in Na^+ and K^+ buffer solutions and of distamycin A in Na^+ buffer suggest a groove-binding mechanism also in these cases, as observed for distamycin A in K^+ condition. Nevertheless, it is not possible to demonstrate with certainty this binding mode from the available calorimetric and NMR titration data alone, and, at this stage, we cannot rule out the possibility that the analyzed molecules

interacts also with the terminal G-planes, as described by Maizel and coworkers [20], or with a mixed groove/G-quartet stacking binding mode, as described by Neidle and coworkers [22].

Both of these drugs show a different behaviour in the binding of $[d(TGGGGT)]_4$ when the interaction occurs in the presence of K^+ or Na^+. Apparently, there is no specific reason to justify this disparity because the same parallel-stranded conformation is found in solution for the $[d(TGGGGT)]_4$ independently from the ions present [41-45].

However, a difference between the K^+ and Na^+ structures of $[d(TGGGGT)]_4$ involves the coordination of the cations. Indeed, the larger K^+ ions are placed exactly at the midpoint between two G-tetrads; whereas the smaller Na^+ ions are usually placed closer to the plane of each G-tetrad rather than between two G-tetrads [43, 46-48]. Additionally, the smaller Na^+ ions are quite mobile within the central cavity of the quadruplex and can even enter or exit from the quadruplex core. On the other hand, the larger K^+ ions remain caged in the central position between G-tetrads [49, 50]. Reasonably, the different cations coordination could alter the recognition process of Dist-A. In fact, the presence of cations coordinated at the top and at the bottom of the cavity at the carbonyl oxygens of the G-tetrads, or their absence, could also influence the organization of the terminal thymines that (i) can occupy an external disordered position or (ii) form interstrand T-T base pairs or (iii) can even penetrate into the grooves forming van der Waals interactions and/or hydrogen bonds, as in some cases found in presence of Na^+ and Tl^+, preventing the binding of the ligands [44].

These effects on the binding may also partly reflect inherent differences in hydration for the quadruplex structures and/or binding-induced release of counterions. Indeed, the different ionic radii between the two

cations could led to a different extent of ions' distribution all around the surface of the quadruplex [51]. In particular, the Na^+ ions should fit better into the grooves of the quadruplex, preventing or at least altering the binding of the groove ligands.

The calorimetric data for the interactions of distamycin and compound 1 with the $d[AG_3(T_2AG_3)_3]$ reveal that, also in this case, the binding behaviour of the two drugs in the Na^+ solution is different compared to the equivalent K^+ solution. Although the overall stoichiometries are unaffected, different ligand affinity and/or enthalpy of binding were found. In particular, both ligands show a higher affinity for the quadruplex in the Na^+ buffer. Additionally, the binding of 1 in this buffer was characterized by a moderately large and positive enthalpy of binding, in contrast with the behaviour observed in the K^+ buffer.

This different behaviour in the two buffers could be more easily justified because the $d[AG_3(T_2AG_3)_3]$ forms two completely different structures in Na^+ or K^+ solution [31, 52], both containing three stacked G-tetrad planes. Indeed, it is well known that the human telomeric sequence forms an intramolecular anti-parallel arrangement in Na^+ solution, with a diagonal TTA loop at one end and two lateral TTA loops at the other end. The guanines adopt a *syn-syn-anti-anti* glycosidic conformation in each tetrad. On the other hand, in K^+ solution, this sequence adopts a (3+1) hybrid-type G-quadruplex folding topology, in which three strands are oriented in one direction and the fourth is in opposite direction. Interestingly, one of the outer G-tetrads is characterized by an *anti-syn-syn-syn* guanine glycosidic conformation, whereas the other two G-tetrads are characterized by *syn-anti-anti-anti* guanine glycosidic torsion angles. Furthermore, it has been shown that this sequence in K^+ buffer adopt two distinct intramolecular folding containing the same (3+1) G-tetrad

core, but different loop arrangements both with one strand-reversal and two lateral loops [32].

The major entropic contribution for the binding of the ligands to the human telomeric quadruplexes suggests a groove-binding mechanism also for these interactions. However, the human telomeric quadruplexes used in this study differ from $[d(TGGGGT)]_4$ for the structure (grooves size and presence of lateral loops), and for the number of stacked G-tetrad planes ($[d(TGGGGT)]_4$ and $d[AG_3(T_2AG_3)_3]$ contain four and three planes, respectively). This could give explanation for the different stoichiometries in the binding of the $[d(TGGGGT)]_4$ and $d[AG_3(T_2AG_3)_3]$ quadruplexes.

Very interestingly, this ITC and NMR study shows that the structural modifications of the compounds 1 and 2 influence the energetic of interaction with quadruplexes, increasing (for compound 1) or decreasing (for compound 2) the affinity of the ligands toward the quadruplex molecules.

The data reported in this article, along with other studies dealing with the interactions of the distamycin and its derivatives with quadruplex structures [19, 20, 22, 23, 53], demonstrate that the interaction between these drugs and DNA quadruplex structures is regulated by many factors. In particular, cation-induced rearrangements and any structural factor involving accessibility of drugs to the quadruplexes are important. All these findings encourage the design and the investigation of new quadruplex ligands that could be capable to discriminate among several quadruplex scaffolds.

Notes and references

1. Hurley L.H., Boyd F.L., Trends Pharm. Sci. 9 (1988) 402-407.

2. Haq I., Ladbury J., J. Mol. Recognit. 13 (2000) 188-197.
3. Pagano B., Giancola C., Curr. Cancer Drug Targets 7 (2007) 520-540.
4. Gellert M., Lipsett M.N., Davies D.R., Proc. Natl. Acad. Sci. USA 48 (1962) 2013-2018.
5. Fry M., Front Biosci. 12 (2007) 4336-4351.
6. de Lange T., Oncogene 21 (2002) 532-540.
7. Zou Y., Sfeir A., Gryaznov S.M., Shay J.W., Wright W.E., Mol. Biol. Cell 15 (2004) 3709-3718.
8. Klobutcher L.A., Swanton M.T., Donini P., Prescott D.M., Proc. Natl. Acad. Sci. USA 78 (1981) 3015-3019.
9. Moyzis R.K., Buckingham J.M., Cram L.S., Dani M., Deaven L.L., Jones M.D., Meyne J., Ratliff R.L., Wu J.R., Proc. Natl. Acad. Sci. USA 85 (1988) 6622-6626.
10. Williamson J.R., Annu. Rev. Biophys. Biomol. Struct. 23 (1994) 703-730.
11. 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., Science 266 (1994) 2011-2015.
12. Shay J., Wright W., Nat. Rev. Drug Discov. 5 (2006) 577-584.
13. Nugent C.I., Lundblad V., Genes Dev. 12 (1998) 1073-1085.
14. Zahler A.M., Williamson J.R., Cech T.R., Prescott D.M., Nature 350 (1991) 718-720.
15. Alberti P., Lacroix L., Guittat L., Helene C., Mergny J.L., Mini Rev. Med. Chem. 3 (2003) 23-36.
16. Riou J.F., Curr. Med. Chem. Anticancer Agents 4 (2004) 439-443.
17. Pelton J.G., Wemmer D.E., J. Am. Chem. Soc. 112 (1990) 1393-1399.
18. Randazzo A., Galeone A., Mayol L., [1] Chem. Commun. 11 (2001) 1030-1031.

19. Martino L., Virno A., Pagano B., Virgilio A., Di Micco S., Galeone A., Giancola C., Bifulco G., Mayol L., Randazzo A., J. Am. Chem. Soc. 129 (2007) 16048-16056.
20. Cocco M.J., Hanakahi L.A., Huber M.D., Maizels N., Nucl. Acids Res. 31 (2003) 2944-2951.
21. Zaffaroni N., Lualdi S., Villa R., Bellarosa D., Cermele C., Felicetti P., Rossi C., Orlandi L., Daidone M.G., Eur. J. Cancer 38 (2002) 1792-1801.
22. Moore M.J., Cuenca F., Searcey M., Neidle S., Org. Biomol. Chem. 4 (2006) 3479-3488.
23. Pagano B., Mattia C.A., Virno A., Randazzo A., Mayol L., Giancola C., Nucleosides Nucleotides Nucleic Acids 26 (2007) 761-765.
24. Haq I., Trent J.O., Chowdhry B.Z., Jenkins T.C., J. Am. Chem. Soc. 121 (1999) 1768-1779.
25. Cantor C.R., Warshow M.M., Shapiro H., Biopolymers 9 (1970) 1059-1077.
26. Animati F., Arcamone F.M., Conte M.R., Felicetti P., Galeone A., Lombardi P., Mayol L., Paloma L.G., Rossi C., J. Med. Chem. 38 (1995) 1140-1149.
27. Hwang T.L., Shaka A.J., J. Magn. Res. A112 (1995) 275-279.
28. Dalvit C., J. Biomol. NMR 11 (1998) 437-444.
29. Parkinson G.N., Lee M.P., Neidle S., Nature 417 (2002) 876-880.
30. Ambrus A., Chen D., Dai J., Bialis T., Jones R.A., Yang D., Nucleic Acids Res. 34 (2006) 2723-2735.
31. Luu K.N., Phan A.T., Kuryavyi V., Lacroix L., Patel D.J., J. Am. Chem. Soc. 128 (2006) 9963-9970.
32. Phan A.T., Luu K.N., Patel D.J., Nucleic Acids Res. 34 (2006) 5715-5719.

33. Xu Y., Noguchi Y., Sugiyama H., T, Bioorg. Med. Chem. 14 (2006) 5584-5591.
34. Dai J., Carver M., Punchihewa C., Jones R.A., Yang D., Nucl. Acids Res. 35 (2007) 4927-4940.
35. Haider S.M., Parkinson G.N., Neidle S., J. Mol. Biol. 326 (2003) 117-125.
36. Parkinson G.N., Ghosh R., Neidle S., Biochemistry 46 (2007) 2390-2397.
37. Gavathiotis E., Heald R.A., Stevens M.F., Searle M.S., J. Mol. Biol. 334 (2003) 25-36.
38. White E.W., Tanious F., Ismail M.A., Reszka A.P., Neidle S., Boykin D.W., Wilson W.D., Biophys. Chem. 126 (2007) 140-153.
39. Haq I., Arch. Biochem. Biophys. 403 (2002) 1-15.
40. Chaires J.B., Arch. Biochem. Biophys. 453 (2006) 24-29.
41. Aboul-ela F., Murchie A.I., Norman D.G., Lilley D.M., J. Mol. Biol. 243 (1994) 458-471.
42. Laughlan G., Murchie A.I., Norman D.G., Moore M.H., Moody P.C., Lilley D.M., Luisi B., Science 265 (1994) 520-524.
43. Phillips K., Dauter Z., Murchie A.I., Lilley D.M., Luisi B., J. Mol. Biol. 273 (1997) 171-182.
44. Caceres C., Wright G., Gouyette C., Parkinson G., Subirana J.A., Nucleic Acids Res. 32 (2004) 1097-1102.
45. Creze C., Rinaldi B., Haser R., Bouvet P., Gouet P., Acta Crystallogr. D Biol. Crystallogr. D63 (2007) 682-688.
46. Horvath M.P., Schultz S.C., J. Mol. Biol. 310 (2001) 367-377.
47. Sundquist W.I., Klug A., Nature 342 (1989) 825-829.
48. Haider S., Parkinson G.N., Neidle S., J. Mol. Biol. 320 (2002) 189-200.

49. Chowdhury S., Bansal M., J. Biomol. Struct. Dyn. 18 (2001) 647-669.
50. van Mourik T., Dingley A.J., Chemistry 11 (2005) 6064-6079.
51. Chowdhury S., Bansal M., J. Phys. Chem. B 105 (2001) 7572-7578.
52. Wang Y., Patel D.J., Structure 1 (1993) 263-282.
53. David W.M., Brodbelt J., Kerwin S.M., Thomas P.W., Anal. Chem. 74 (2002) 2029-2033.

"The important thing is not to stop questioning."

Albert Einstein