J|A|C|S
ARTICLES

Published on Web 04/15/2010

Structural and Conformational Requisites in DNA Quadruplex Groove Binding: Another Piece to the Puzzle

Sandro Cosconati,[†] Luciana Marinelli,[†] Roberta Trotta,[‡] Ada Virno,[‡] Stefano De Tito,[†] Romeo Romagnoli,[§] Bruno Pagano,[||] Vittorio Limongelli,[†] Concetta Giancola,[⊥] Pier Giovanni Baraldi,[§] Luciano Mayol,[‡] Ettore Novellino,[†] and Antonio Randazzo*[,‡]

Dipartimento di Chimica Farmaceutica e Tossicologica, Università degli Studi di Napoli "Federico II", via D. Montesano 49, I-80131 Napoli, Italy, Dipartimento di Chimica delle Sostanze Naturali, Università degli Studi di Napoli "Federico II", via D. Montesano 49, I-80131 Napoli, Italy, Dipartimento di Scienze Farmaceutiche, Università di Ferrara, via Fossato di Mortara 17/19, 44100 Ferrara, Italy, Dipartimento di Scienze Farmaceutiche, Università di Salerno, via Ponte Don Melillo, 84084 Fisciano (SA), Italy, and Dipartimento di Chimica "P. Corradini", Università degli Studi di Napoli "Federico II", via Cintia, I-80126, Napoli, Italy

Received January 15, 2010; E-mail: antonio.randazzo@unina.it

Abstract: The study of DNA G-quadruplex stabilizers has enjoyed a great momentum in the late years due to their application as anticancer agents. The recognition of the grooves of these structural motifs is expected to result in a higher degree of selectivity over other DNA structures. Therefore, to achieve an enhanced knowledge on the structural and conformational requisites for quadruplex groove recognition, distamycin A, the only compound for which a pure groove binding has been proven, has been chemically modified. Surprisingly, structural and thermodynamic studies revealed that the absence of Coulombic interactions results in an unprecedented binding position in which both the groove and the 3′ end of the DNA are occupied. This further contribution adds another piece to the so far elusive puzzle of the recognition between ligands and DNA quadruplexes and will serve as a platform for a rational design of new groove binders.

Introduction

Molecular recognition is a very important phenomenon in biology and chemistry, and the molecular basis of interaction between small ligands and biological macromolecules is the subject of numerous investigations aimed at the rational design of molecules with specific biological activities. In particular, during the last years, a great number of manuscripts have appeared in the literature dealing with the development of low molecular weight agents specifically interacting with DNA.[1]

Since now, most of the investigations were focused on targeting duplex DNA. Only recently, upon the identification of G-quadruplex motifs as biologically crucial structures, a certain number of studies about ligand/DNA-quadruplex interactions have been reported.[2] G-quadruplexes are four-stranded helical structures, comprising stacks of G-tetrads, which are the planar association of four guanines in a cyclic Hoogsteen hydrogen-bonding arrangement.[3] This structural motif seems to play a critical role in a number of biological processes and they are widespread in genomic regions like gene promoters and telomeres.[4] Telomeres are ensembles of proteins and specialized noncoding DNA sequences that protect the ends of the chromosomes 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, which, therefore, can form G-quadruplexes.[4] Interestingly, telomerase, the enzyme which elongates the G-rich strand of telomeric DNA, is active in about 85% of tumors, leading cancer cells to infinite lifetime. Therefore, the inhibition of telomerase has become an attractive and promising strategy for the anticancer therapy. Since telomerase requires a single-stranded telomeric primer, the formation of G-quadruplex structures by telomeric DNA inhibits the telomerase activity. Moreover, small molecules able to stabilize G-quadruplexes have been shown to be effective telomerase inhibitors.[5,6] These findings have brought a great momentum in the study of G-quadruplex binders.

So far, most of the reported G-quadruplex binding agents bind to DNA by interacting with the wide π-surface of the G-tetrads at the edges of the quadruplex. Groove-binding mode to quadruplex structure was first proposed by our research group

[†] Dipartimento di Chimica Farmaceutica e Tossicologica, Università degli Studi di Napoli "Federico II".
[‡] Dipartimento di Chimica delle Sostanze Naturali, Università degli Studi di Napoli "Federico II".
[§] Università di Ferrara.
[||] Università di Salerno.
[⊥] Dipartimento di Chimica "P. Corradini", Università degli Studi di Napoli "Federico II".

(1) Nelson, S. M.; Ferguson, L. R.; Denny, W. A. *Mutat. Res.* **2007**, *623*, 24–40.
(2) Ou, T. M.; Lu, Y. J.; Tan, J. H.; Huang, Z. S.; Wong, K. Y.; Gu, L. Q. *ChemMedChem* **2008**, *3*, 690–713.
(3) Neidle, S. *Curr. Opin. Struct. Biol.* **2009**, *19*, 239–250.
(4) Lipps, H. J.; Rhodes, D. *Trends Cell Biol.* **2009**, *19*, 414–422.
(5) Neidle, S. *FEBS J.* **2010**, *277*, 1118–1125.
(6) Pagano, B.; Giancola, C. *Curr. Cancer Drug Targets* **2007**, *7*, 520–540.

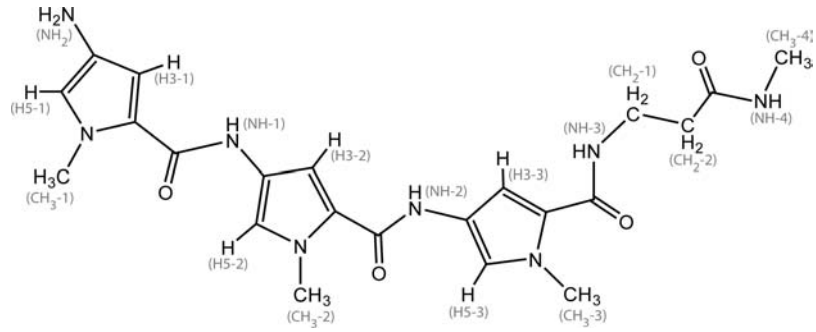


Figure 1. Chemical structure of derivative **3**.

for the distamycin A,[7] and up to now, this is the only molecular entity for which a pure groove binding mode has been experimentally proven.

The interest in the groove binding recognition is mainly fueled by the chemical and conformational differences existing between quadruplex and duplex grooves. Unfortunately, the understanding of the chemical, structural, and conformational features responsible for G-quadruplex groove binding is still in its infancy. In this scenario, discovery of new distamycin A analogues could greatly contribute to linking structural features and groove binding properties.

Analysis of the distamycin A/[d(TGGGGT)]$_4$ interaction patterns revealed that this ligand binds, as antiparallel dimer, the quadruplex groove by adopting a crescent shape and by establishing four H-bonds with the G bases.[7] Moreover, strong Coulombic interactions are established between positively charged amidinium moiety of the ligand and the backbone phosphate groups of the quadruplex. In a previous study, we investigated the importance of the crescent shape extension by varying the pyrrole units number in distamycin A.[8] Experiments reveal that the presence of one additional pyrrole unit affects the binding of the molecule and varies its stoichiometry, whereas the addition of two pyrrole units leads to a total loss of affinity of the new derivative for [d(TGGGGT)]$_4$.

Here, we report a distamycin A (**3**) analogue (Figure 1), where the major change into the structure is the replacement of amidinium group by an uncharged N-methyl amide moiety, to probe the real importance of the unique Coulombic interaction in the distamycin A/[d(TGGGGT)]$_4$ complex formation.

Particularly, the synthesis, the NMR titration profile, and the three-dimensional structure of **3**/[d(TGGGGT)]$_4$ complex are reported. To thoroughly characterize the binding of the distamycin A derivative with the G-quadruplex DNA motif, a calorimetric study was also performed.

Experimental Section

Oligonucleotide Synthesis. The oligonucleotide d(TGGGGT) was synthesized on a Millipore Cyclone Plus DNA synthesizer using solid phase β-cyanoethyl phosphoramidite chemistry at 15 μmol scale. Commercially available 5′-DMT-aminoprotected-8-bromodeoxyguanosine-3′-phosphoramidite was used for the preparation of the modified oligonucleotide d(TGG^{Br}GGT). The latter was assembled using the standard solid phase β-cyanoethylphosphoramidite chemistry. 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.

Synthesis of *N*-[5-({[5-({[3-Amino-3-(*N*-methylcarbamoyl)-ethyl]amino}carbonyl)-1-methyl-1*H*-pyrrol-3-yl]amino}carbonyl)-1-methyl-1*H*-pyrrol-3-yl]-4-(formylamino)-1-methyl-1*H*-pyrrole-2-carboxamide (2). A solution of 3-amino-*N*-methylpropanamide hydrochloride (123 mg, 1.2 mmol) in dry DMF (8 mL) was cooled to 5 °C and *N*, *N*′-diisopropylethylamine (206 mL, 1.2 mmol) was added. After 10 min, the acid **1** (412 mg, 1 mmol.) and then EDCI (384 mg, 2 mmol) were added. The reaction mixture was stirred at room temperature for 18 h, the solvent evaporated *in vacuo*, and the crude residue purified by flash chromatography. The crude product purified by chromatography on a silica gel column, using DCM/MeOH (8:2) as eluent, furnished the compound **2** as a white solid (367 mg, 74% yield). Mp = 156−158 °C; IR (KBr): 3420, 3310, 1665, 1554, 1312, and 1212 cm^{-1}; ^{1}H NMR (DMSO-d_6) δ: 2.42 (m, 2H), 2.77 (m, 3H), 3.34 (m, 2H), 3.78 (s, 3H), 3.82 (s, 3H), 3.83 (s, 3H), 6.65 (d, J = 1.6 Hz, 1H), 6.83 (d, J = 1.6 Hz, 1H), 6.82 (d, J = 1.6 Hz, 1H), 7.01 (d, J = 1.6 Hz, 1H), 7.17 (d, J = 1.6 Hz, 1H), 7.22 (d, J = 1.6 Hz, 1H), 7.89 (bs, 1H), 8.08 (s, 1H), 8.13 (s, 1H), 8.18 (s, 1H); 9.91 (s, 1H), 9.94 (s, 1H). FAB-MS (MALDI-TOF): 497.8 [M + 1]$^+$.

Synthesis of 4-Amino-*N*-[5-({[5-({[3-amino-3-(*N*-methylcarbamoyl)ethyl]amino}carbonyl)-1-methyl-1*H*-pyrrol-3-yl]amino}carbonyl)-1-methyl-1*H*-pyrrol-3-yl]-1-methyl-1*H*-pyrrole-2-carboxamide (3). To 5 mL of a solution obtained by the addition of 24 mL of ethanol (95%) to 1 mL of 36% HCl in water was added compound **2** (248 mg, 0.5 mmol), and the mixture stirred at room temperature for 24 h. After this time, the reaction was filtered, the residue washed with dry ethyl ether (5 mL), and then dried on P_2O_5, to furnish compound **3** as brown solid. Mp = 262−264 °C; ^{1}H NMR (DMSO-d_6) δ: 2.41 (m, 2H), 2.56 (m, 3H), 3.24 (m, 2H), 3.78 (s, 3H), 3.80 (s, 3H), 3.81 (s, 3H), 6.75 (d, J = 1.6 Hz, 1H), 6.82 (d, J = 1.6 Hz, 1H), 6.84 (d, J = 1.6 Hz, 1H), 6.96 (d, J = 1.6 Hz, 1H), 7.04 (d, J = 1.6 Hz, 1H), 7.20 (d, J = 1.6 Hz, 1H), 7.88 (bs, 1H), 8.12 (s, 1H), 9.92 (s, 1H), 9.96 (s, 1H), 10.20 (bs, 3H). FAB-MS (MALDI-TOF): 469.4 [M + 1]$^+$.

Nuclear Magnetic Resonance Experiments. The quadruplex NMR sample was prepared at a concentration of 2 mM (8 mM single strand concentration), 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

(7) 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.* **2007**, *129*, 16048–16056.
(8) Pagano, B.; Virno, A.; Mattia, C. A.; Mayol, L.; Randazzo, A.; Giancola, C. *Biochimie* **2008**, *90*, 1224–1232.

^{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 the sample in H_2O were recorded using pulsed-field gradient DPFGSE[9,10] for H_2O suppression. Phase-sensitive NOESY spectra[11] were recorded with mixing times of 100 and 200 ms (T = 25 °C). Pulsed-field gradient DPFGSE[9,10] sequence was used for NOESY experiments in H_2O. TOCSY[12] and zTOCSY[13] spectra with mixing times of 100 ms were recorded with D_2O solution.

All experiments were recorded using STATES-TPPI[14] procedure for quadrature detection. In all 2D experiments, the time domain data consisted of 2048 complex points in t2 (16K in the zTOCSY) and 400−512 fids in t1 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) and on iMAC running iNMR software (www.inmr.net).

Isothermal Titration Calorimetry. ITC measurements were conducted at 25 °C on a CSC 4200 Calorimeter from Calorimetry Science Corporation (Lindon, UT). In these experiments, 10 μL aliquots of a 1250 μM ligand solution were injected into a sample cell containing 1300 mL of 40 μM quadruplex solution, with stirring at 297 rpm, for a total of 25 injections. The delay between injections was 400 s. Each ligand injection produced a heat burst curve. The areas under these heat burst curves were determined by integration to yield the associated injection heats. The heats of dilution were determined in parallel experiments by injecting a ligand solution of the same concentration in the same buffer. The heat of dilution was subtracted from the interaction heats prior to curve fitting. The corrected heat values are plotted as a function of the molar ratio, 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 ($\Delta H°$), equilibrium binding constant (K_b), and stoichiometry (n). The remaining thermodynamic parameters, $\Delta G°$ and $T\Delta S°$, were derived using the standard relationships $\Delta G° = -RT \ln K_b$ and $T\Delta S° = \Delta H° - \Delta G°$. The buffer conditions used for ITC measurements were the same of NMR experiments.

Structure Calculations. Restrained Simulated Annealing (SA) calculations were performed using the AMBER 10.0 package,[15] with DNA described by the latest AMBER force field for nucleic acids (ff99bsc0),[16,17] while compound **3** was parametrized using the General AMBER Force Field.[18] The Jaguar software package (Schrodinger) was used to calculate atomic charges using DFT with an HF/6-31G** basis set at the B3LYP level.

For the initial AMBER model, compound **3** was manually docked into two opposite grooves of the $[d(TGGGGT)]_4$ using as a reference structure the distamycin A/$[d(TGGGGT)]_4$ solution structure (PDB code 2JT7).[7] In fact, given the presence of many head to tail ligand−ligand NOEs, as happened for distamycin A, also **3** is expected to bind the quadruplex structure in a dimeric antiparallel form. Each **3** dimer was then placed into the quadruplex groove consistently with what recorded in NMR experiments.

For annealing simulations, the General Born solvation (igb = 2) with monovalent salt concentration corresponding to 0.1 M was used. The complex was heated to 900 K in the first 5 ps, cooled to 100 K for the next 13 ps, and then cooled to 0 K for the last 2 ps. The temperature of the system was maintained with a varying time constant: 0.4 ps during heating, 4 ps during cooling to 100 K, 1 ps for the final cooling stage, and then reduced from 0.1−0.05 for the last picosecond. The force constants for NOE constraints were increased from 3 to 30 kcal mol^{-1} Å^{-2} during the first 5 ps and then maintained constant for the rest of the simulation. These force constants were applied in the form of a parabolic, flat-well energy term where r is the model distance or torsion angle and k is the respective force constant.

$$E_{\text{constraint}} = k(r_2 - r)^2 \qquad r_1 \le r < r_2$$

$$E_{\text{constraint}} = 0 \qquad r_2 \le r \le r_3$$

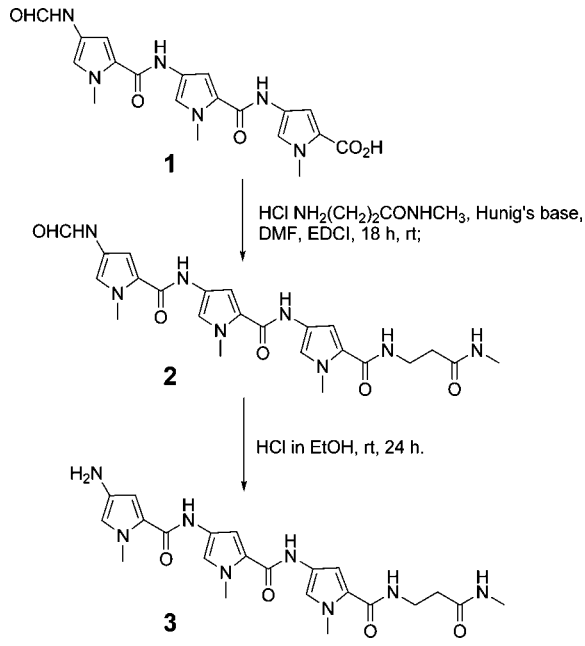
$$E_{\text{constraint}} = k(r_3 - r)^2 \qquad r_3 \le r < r_4$$

The values for r_1 and r_4 represent upper and lower distance bounds, defining the linear energetic penalty before and after the flat-well energy term. The applied distance NOE constraints were retrieved by cross-peak volume integrations performed with the program FELIX 98 (Accelrys, San Diego, CA), 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. The NOE restraints were generated with three distance classifications as follows: strong NOEs (r_2 = 1.0 Å; r_3 = 3.5 Å), medium NOEs (r_2 = 3.0 Å; r_3 = 4.5 Å), and weak NOEs (r_2 = 4.0 Å; r_3 = 6.0 Å). Distance constraints between base pairs, set at half the NOE distance constraint strength, were applied to maintain hydrogen bonding (as indicated by the observation of imino proton resonances). Hydrogen bonds constraints were used in the range of 1.7 Å (r_2)/2.3 Å (r_3). Planarity force constraints of 200 kcal mol^{-1} rad^{-2} were applied to G bases throughout the simulations. In agreement with NMR data (see Results and Discussion), the backbone torsion angles α, β, γ, δ, and ε were restrained in the range −150°/−30°, −230°/−110°, 20°/100°, 95°/175°, and −230°/−110°, respectively.[19] Further, glycosidic torsion angles χ were fixed in the anti domain (−155°/−75°). Interactions within the system were calculated. An unrestrained energy minimization step completed the simulated annealing run. This simulated annealing/energy minimization procedure was repeated 100 times. Ten best models were selected based on the value of the overall potential energy and NMR restraint violations for further analysis. After charge neutralization by the addition of 20 K^+ ions, the complexes were solvated with 16 758 water molecules in a TIP3P pre-equilibrated box.[20] Several equilibration steps were performed comprising minimization of the solvent molecules with the DNA and ligand fixed, minimization of the whole system, and slow heating to 300 K with weak positional restraints on DNA and drug atoms under constant-volume conditions. The following 5 ns production runs were applied in the NPT ensemble. The particle mesh Ewald method[21] was used to evaluate the electrostatic interactions with a direct space sum cutoff of 10 Å. With the bond lengths involving hydrogen atoms kept fixed with the SHAKE algorithm, a time step of 2 fs was employed.[22]

Related conformational substates populated during the molecular dynamics simulation were analyzed with the AMBERS's PTRAJ

(9) Hwang, T. L.; Shaka, A. J. *J. Magn. Reson.* **1995**, *A112*, 275–279.
(10) Dalvit, C. *J. Biomol. NMR* **1998**, *11*, 437–444.
(11) Jeener, J.; Meier, B.; Bachmann, H. P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546–4553.
(12) Braunschweiler, L.; Ernst, R. R. *J. Magn. Reson.* **1983**, *53*, 521–528.
(13) Trippleton, M. J.; Keeler, J. *Angew. Chem., Int. Ed.* **2003**, *42*, 3938–3941.
(14) Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1989**, *85*, 393–399.
(15) Case, D. A. *AMBER*; 10th ed.; University of California: San Francisco, CA, 2008.
(16) Wang, J. M.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049–1074.
(17) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., 3rd; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817–3829.
(18) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
(19) Kim, S.; Lin, L.; Reid, B. R. *Biochemistry* **1992**, *31*, 3564–3574.
(20) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
(21) (a) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092. (b) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
(22) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311–1327.

Scheme 1. Chemical Synthesis of Derivative **3**



module.[23] For the 10 trajectories, the average configuration was taken as a reference for subsequent mass-weighted rmsd calculations considering all atoms excluding terminal T bases. The single snapshot of each trajectory with the lowest rmsd was taken as the representative dynamic structure of that simulation.

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

Results and Discussions

Synthesis of Derivative 3. The novel derivative **3** was obtained following the procedure reported in the Scheme 1. Starting from derivative **1**, prepared by a methodology previously reported,[24] it was condensed with 3-amino-*N*-methylpropanamide hydrochloride[25] using an excess (2 equiv) of 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide hydrochloride (EDCI) as coupling agent, in DMF and in presence of Hunig's base at room temperature for 18 h, to obtain the N-methyl carbamoyl derivative **2** in good yield after purification by silica gel flash-chromatography. The amido modified distamycin A derivative **2** was then transformed in the corresponding desformyl derivative **3** by treatment of HCl in a mixture of ethanol and water. The reason for this last synthetic step was to obtain the hydrochloride derivative **3**, that turned out to be more soluble in pure H_2O than derivative **2** at the concentration used for the preparation of the NMR samples.

Nuclear Magnetic Resonance Experiments. To evaluate the binding properties of derivative **3**, and to perform a direct comparison with the binding behavior of Dist-A, $[d(TGGGGT)]_4$ has been titrated with **3** at the same experimental conditions (buffer, temperature, DNA concentration) used for Dist-A.[7]

As far as distamycin A is concerned, below 2:1 ligand/quadruplex stoichiometry, the addition of Dist-A to $[d(TGGGGT)]_4$ caused gradual changing in chemical shift of the signal of the quadruplex, whereas further addition of drug caused the appearance of a new set of proton signals, whose intensities 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−DNA.

On the other hand, the NMR titration profile with **3** turned out to be clearly different from that observed in the case of Dist-A (Figure 2). In fact, the addition of **3** to $[d(TGGGGT)]_4$ caused only gradual changes in the chemical shift of DNA proton resonances. At ligand/DNA ratio of 4:1, 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. Moreover, it is noteworthy that, throughout the whole NMR titration, a single set of signals was present for derivative **3** protons, which only grew in intensity and did not show any significant change in chemical shift values by increasing ligand concentration. These observations suggest that derivative **3** bind the quadruplex in a fast process on the NMR time scale.

To preliminarily evaluate the binding site of derivative **3**, a comparison of resonances of some protons of the uncomplexed DNA and the complexed one has been done. In particular, we report the $\Delta\delta$ values (chemical shifts of the complex minus free DNA) of aromatic, methyl and imino protons in Figure 3A. Interestingly, all the analyzed resonances shifted. Nevertheless, the signal of the protons of T1 residue shifted the least, whereas the ones of residue T6 the more. In any case, a general shift of the aromatic and imino signals was observed also for the G2, G3, G4, and G5. This means that, basically, derivative **3** is able to recognize most of the molecule, even if it recognizes preferentially the 3′ edge of the quadruplex.

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 and TOCSY spectra (700 MHz, $T =$ 25 °C). In particular, as far as DNA is concerned, ^{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. The direct comparison of the intensities of the NOESY crosspeaks (700 MHz, $T = 25$ °C, mixing time 100 ms) 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, 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 conformation resembles closely that of the uncomplexed $[d(TGGGGT)]_4$ possessing a right-handed B-form helix structure.

Each exchangeable proton signal was then assigned to the pertinent hydrogen by an in-depth analysis of the NOESY spectra. In particular, the NOE contact between H1 proton of DNA at δ_H 11.34 and the methyl protons at δ_H 1.34 of the T1 residue led us to identify the imino protons belonging to the tetrad that is in proximity of the 5′ edge of the quadruplex. Similarly, the identification of the imino protons belonging to the tetrad in proximity of the 3′ edge was confirmed by the NOE between H1 proton at δ_H 10.64 and the methyl group at δ_H 1.66 ppm of the thymine T6. The other four H1 resonances of the remaining two central tetrads were identified by analyzing the NOE connections with the adjacent tetrads.

(23) Shao, J.; Tanner, S. W.; Thompson, N.; Cheatham, T. E., III. *J. Chem. Theory Comput.* **2007**, *3*, 2312–2334.
(24) Grehn, L.; Ragnarsson, U. *J. Org. Chem.* **1981**, *46*, 3492–3497.
(25) Juaristi, E.; Quintana, D.; Lamatsch, B.; Seebach, D. *J. Org. Chem.* **1991**, *56*, 2553–2557.

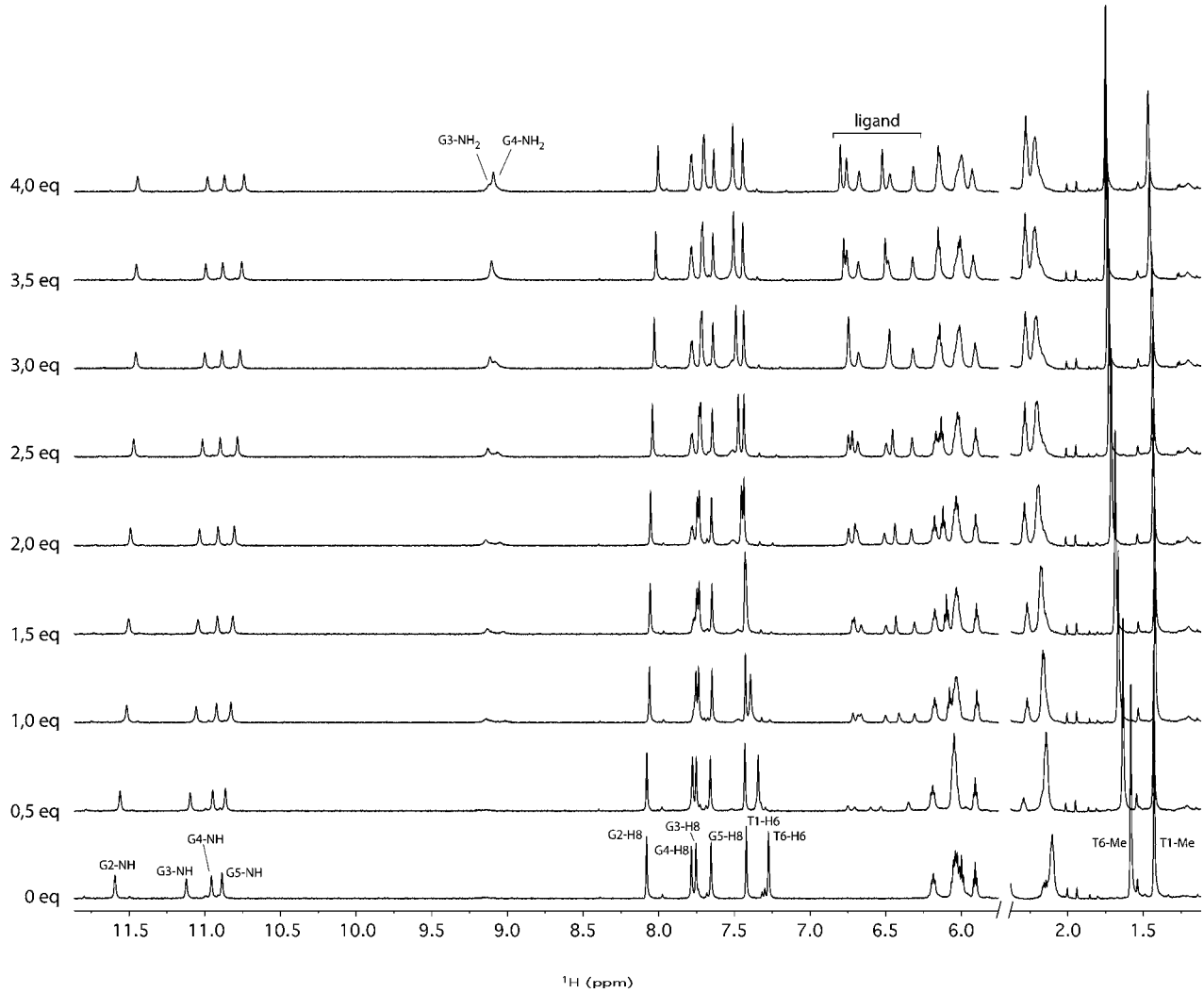


Figure 2. NMR titration of $[d(TGGGGT)]_4$ with derivative **3** (700 MHz, $T = 25$ °C). Two millimolar (8 mM single strand concentration), in 0.6 mL (H_2O/D_2O 9:1) buffer solution having 10 mM KH_2PO_4, 70 mM KCl, 0.2 mM EDTA, pH 7.0. Equivalents of the drug are reported on the left of each spectrum.

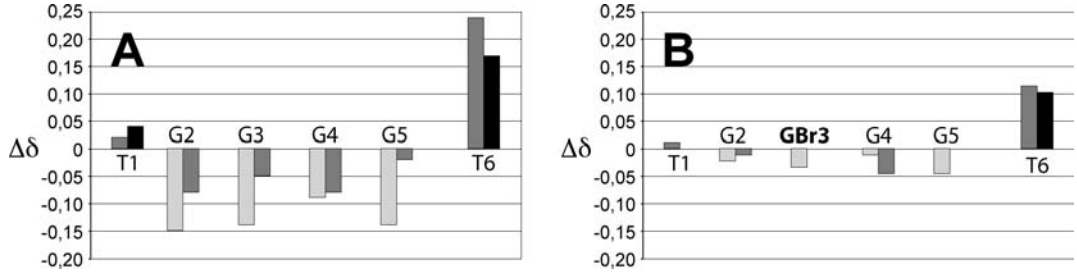


Figure 3. Difference in chemical shifts ($\Delta\delta$) of DNA upon binding of **3** (complex **3**:DNA minus DNA alone) to $[d(TGGGGT)]_4$ (A) and $[d(TGG^{Br}GGT)]_4$ (B). $\Delta\delta$ values are reported for imino (light gray), aromatic (dark gray) and methyl (black) protons. No aromatic hydrogens are present for brominated residues.

The resonances of derivative **3** were assigned, first, identifying the resonances belonging to the hydrogen on the convex side of the molecule. Particularly, the H5 pyrrole protons (δ_H 6.70, 6.66, and 6.57 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 3.52, 3.40, and 3.37 ppm for CH_3-1, CH_3-2, CH_3-3, respectively). As for the protons on the concave side of derivative **3**, the H3 pyrrole protons (δ_H 6.42, 6.37, and 6.22 for H3-1, H3-2 and H3-3, respectively) have been assigned by virtue of very small scalar coupling with H5 protons of the same pyrrole ring. Unfortunately, we were not able to unequivocally assign NH-1, NH-2 and NH-3 protons due to the presence of ambiguous NOEs with the adjacent H3s.

The NOESY spectrum of the complex contains a number of intermolecular ligand−ligand and ligand−DNA NOEs, in addition to intramolecular ones. Particularly, as for ligand−ligand contacts, 12 head to tail NOEs (for each molecule of **3**, 24 for each dimer) were clearly discernible in the NOESY spectra (Table 1), that unambiguously indicates that the ligand molecules

Table 1. Intermolecular Head to Tail Ligand−Ligand NOE Contacts

ligand proton	ligand proton
H3-1	CH_3-4, CH_3-3, H5-3
H5-1	CH_2-1, CH_3-4, CH_3-3
H3-2	CH_3-3, H3-3, H5-3
H5-2	CH_3-4
H3-3	CH_3-1
H5-3	CH_3-1

Table 2. Intermolecular Ligand−DNA NOE Contacts

ligand proton	DNA proton
H3-1	T6:H2′, T6:H1′,G5:H2′, G3:NH2
H5-1	G5:H1′, G5:H2′, G4:H2′
CH3-1	T6:H1′
H3-2	T6:H1′
H5-2	G5:H2′
H3-3	T6:H6, G5:H1′, G4:H2′
H5-3	G4:H2′, G5:H1′

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 for the binding of Dist-A to both duplex and quadruplex DNA).[7,26]

In addition, 14 NOEs were observed between each molecule of derivative **3** and $[d(TGGGGT)]_4$. In particular, derivative **3** displayed NOE contacts predominantly with residues G4, G5, and T6 of DNA. Analogously to that observed for Dist-A, also in this case, there are no contacts between derivative **3** and the imino protons of the quadruplex (since these point inside the quadruplex core) (Table 2). All this suggests that derivative **3** actually does not stack on the edges of the quadruplex, even if it prefers to interact with the 3′ edge of the quadruplex.

Synthesis and NMR Titration of the Quadruplexes Containing Bromine Atoms. To further confirm that **3** is able to bind the grooves of the quadruplex, as its analogue Dist-A, we have designed and synthesized a modified oligonucleotide, namely, $d(TGG^{Br}GGT)$, where dG^{Br} is 8-bromo-2′-deoxyguanosine, potentially capable to form quadruplex structure and possessing a bulky group (bromine) at the very center of the grooves. Our intention was to titrate $d(TGG^{Br}GGT)$ with **3** and to analyze the titration profile. As a result, if **3** interacts with the portion of the grooves of the quadruplex where the bromine is present, this should prevent (or at least should limit) the insertion of molecules of **3** into the grooves. However, we first tested the capability of $d(TGG^{Br}GGT)$ to form a quadruplex structure. NMR sample of $d(TGG^{Br}GGT)$ 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 min at 80 °C and slowly cooled down to room temperature, then ^{1}H NMR spectra were recorded using DPFGSE pulse sequence for H_2O suppression.[9,10]

The ^{1}H NMR spectrum (700 MHz, $T = 25$ °C) of $d(TGG^{Br}GGT)$ shows the presence of four well-defined singlets in the region 11−12 ppm (data not shown), attributable to imino protons involved in Hoogsteen hydrogen bonds of G-quartets, and the presence of five signals (Figure 4), belonging to three guanine H8 and two thymine H6 protons in the aromatic region. These indicate that a single well-defined quadruplex species is present in solution, consisting of four G-tetrads and possessing

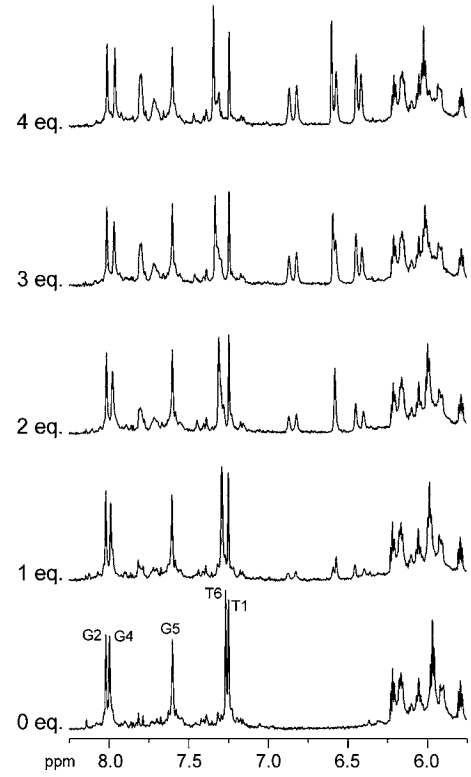


Figure 4. Expanded regions of NMR titration of $[d(TGG^{Br}GGT)]_4$ with derivative **3** (700 MHz, $T = 25$ °C). Two millimolar (8 mM single strand concentration), in 0.6 mL (H_2O/D_2O 9:1) buffer solution having 10 mM KH_2PO_4, 70 mM KCl, 0.2 mM EDTA, pH 7.0. Equivalents of the drug are reported on the left.

a 4-fold symmetry with all strands parallel to each other. In a quadruplex like this, each Br group of the quadruplex faces right into the grooves, pointing outward the quadruplex.

Thus, the quadruplex $[d(TGG^{Br}GGT)]_4$ has been titrated with **3** and the results are reported in Figure 4. The progressive increase of concentration of **3** up to 4 mol equiv caused only a slight drift of DNA signals and drug resonances to gradually grow in intensity (between 6 and 7 ppm). Further additions of drug did not lead to significant changes. As suggested also by the $\Delta\delta$ values of DNA signals (Figure 3B), the extent of chemical shift perturbation is lighter in comparison to the unmodified quadruplex. This strongly indicates that, actually, the presence of the bromines at the very center of the grooves does affect the recognition process, and that **3** binds the grooves of $[d(TGGGGT)]_4$. This also suggests that, in spite of **3** which does not show any NOE contacts with the 5′ edge of the groove, **3** is actually able to span the entire groove.

Structure Calculations. 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). A total of 278 experimental distance restraints 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 3).

Restraints for a number of backbone torsion angles were also used in the calculations. In particular, the combined analysis of the PE-COSY and zTOCSY experiments revealed that all measurable $J_{H1',H2'}$ and $J_{H1',H2''}$ ranged from 7 to 9 Hz and from

(26) Pelton, J. G.; Wemmer, D. E. *J. Am. Chem. Soc.* **1990**, *112*, 1393–1399.

Table 3. Structural Data for the NMR Restraints and Statistics for the 10 Best Structures

total no. of experimental NOE restraints	278
Quadruplex intraresidual restraints	96
Quadruplex interresidual restraints	68
Intermolecular ligand-quadruplex restraints	62
Intermolecular ligand−ligand restraints	52
Dihedral angles (α, β, γ, δ, ε and χ) restraints	96
Quadruplex H-bonds restraints	32
Planarity restraints for G bases	56
Average rmsd (all atoms excluding terminal T bases)	0.84 Å
Energy (kcal/mol)	
Mean amber energy	-1610.42 ± 62.29
NOE distance restraints violation energy	84.34 ± 4.82
Torsion angle restraints violation energy	184.10 ± 18.29
Restraint violations	
Distance (>0.3 Å)	7
Experimental dihedral angles (>10 Å)	2

5 to 6.5 Hz, respectively, and that a number of crosspeaks for H3′/H4′ correlations were missing. Furthermore, the measurable $J_{H3',H4'}$ values were very low. All this indicates a predominant S-type nature of sugar ring conformations. Therefore, the relative δ angles constraints have been consistently constrained into the range 95°/175°.[19] Very interesting information could be also retrieved from the analysis of the sum of H4′ couplings constant ($\sum J_{H4'} = J_{H4',H3'} + J_{H4',H5'} + J_{H4',H5''}$). All measured $\sum J_{H4'}$ turned out to be <10 Hz. This, along with the lack of intense NOEs between H8 and any H5′/H5″protons, suggested that γ angles could be constrained in the range 20°/100°.[19] As far as β and ε angles are concerned, their estimation have been performed measuring the scalar coupling in the 2D proton-detected heteronuclear $^1H-^{31}P$ COSY and by using the semiempirical Karplus equation.[19] Thus, since for all $J_{P,H5'}$ and $J_{P,H5''}$ were <8 Hz, β backbone torsion angles were restricted to −230°/−110°.[19] On the contrary, $J_{P,H3'}$ for residues 2, 3, 4, and 5 were larger than 8 Hz, and the corresponding ε angles were restricted to −230°/−110°.[19] Furthermore, since α (O3′−P−O5′−C5′) and ζ (C3′−O3′−P−O5′) dihedral angles do not involve any protons, it has been impossible to directly retrieve structural information for these angles. Nevertheless, since in a structured DNA α and ζ angles cannot be considered independent from β, γ, δ and ε, we thought opportune to restraint at least one of these angles; thus, α angles have been constrained in the wide range −150°/−30°. Finally, due to the presence of very weak NOEs between H8 and H1′ hydrogens, glycosidic torsion angles χ were fixed in the anti-domain (−155°/−75°).

Therefore, 3D structures which satisfy NOEs and dihedral angles were constructed by simulated annealing (SA) calculations. A total of 100 structures was generated, and among them, the 10 best models were selected based on the value of the overall potential energy and NMR restraint violations for further simulations. To probe the thermodynamic stability of these structures and to further refine the models, these were all subjected to a 5 ns MD simulation in explicit solvent (see Experimental Section). All the 10 MD runs demonstrated a good stability of the whole structure as depicted in Figure 3S in Supporting Information reporting the time versus rmsd plot calculated for all atoms excluding terminal T bases. Moreover, the aforementioned simulations converged toward a well-defined ligand/DNA conformation as depicted in Figure 5 reporting the representative conformations (namely, the ones with the lowest rmsd value with respect to the average structure calculated over the whole production run).

As expected, differently from what demonstrated for Dist-A, compound **3** dimers are shifted toward the 3′ end of the quadruplex. The two staggered antiparallel **3** molecules overlap for their total length with each *N*-methylpyrrole ring of one molecule contacting the terminal amide bond of the other, thus, making synchronized twisting to easily fit the curvature of the DNA quadruplex grooves. Moreover, one molecule of the dimer of **3** is more solvent exposed, while the other is sandwiched between the first one and the quadruplex structure (herein referred as **3ex** and **3in**, respectively). The preference for such an orientation with respect to the G-quadruplex could be ascribable to the absence of a positively charged anchoring point that in distamicyn A stabilizes each monomer so as to project the methyl groups toward the solvent and the opposite amide protons toward the crevice of the groove.

As aforementioned, the two molecules of compound **3** interact with the quadruplex in an antiparallel fashion, as its analogue distamycin A. This is a very interesting result; in fact, in theory, two possible orientations could be predicted for a dimer of **3**: parallel and antiparallel. This behavior could be imputed to the charge-transfer $\pi-\pi$ interactions and to the energetically favorable dipole−dipole interactions established between the monomers as demonstrated by the dipole moments of the two staggered structures which are pointing in opposite directions (see Figure 4S in Supporting Information).

The crescent shape of the ligand allows the formation of favorable interactions with the quadruplex by following the curvature of the grooves in the G4−G5 region and taking contact with T6 at the 3′ end. A comparison with the experimental Dist-A/$[d(TGGGGT)]_4$ complex reveals that in both cases a crescent shape is adopted by the ligand. Nevertheless, such a conformation alone does not seem to guarantee a pure groove binding mode in the $[d(TGGGGT)]_4$. This indicates that the concurrence of the crucial anchoring positively charged point (amidinium group) and the crescent shape conformation in Dist-A are both responsible for a pure groove binding mode.

Calculation of the Delphi electrostatic potential surface reveals that the terminal methylcarboxamide groups of the **3** dimer are hosted in rather polar regions of the quadruplex while the terminal pyrrole ring is exposed to the solvent in the G4 region in **3ex** or embedded in a less hydrophilic environment in the T6 region in **3in** (Figure 5S in Supporting Information).

A close-up view of the best binding pose resulting from NMR constrained SA experiments reveals that the aforementioned terminal groups of **3** are involved in a network of H-bonds with the quadruplex (Figure 6). Most precisely, the flexible carboxamide terminal group of the **3in** molecule is kept in place by an intramolecular H-bond between its NH-4 proton and the carbonyl oxygen attached to the third pyrrole ring and by an intermolecular one established between the terminal carbonyl oxygen of **3in** and the terminal NH_2 group of **3ex**. Moreover, the G5 base is contacted by the dimer through two H-bonds between its O4′ and its N3 atoms with the **3ex** NH_2 and **3in** NH-4 groups, respectively. While polar contacts mainly govern the recognition between the **3** dimer and the center of the quadruplex groove, rather hydrophobic interactions are established with the 3′ end of the $[d(TGGGGT)]_4$ structure. Notably, the terminal pyrrole ring of the **3in** molecule is embedded in a sort of aromatic cage formed by the two T6 bases of the groove and the third pyrrole of the **3ex** molecule.

Isothermal Titration Calorimetry Measurements. Thanks to the great advances in the sensitivity and reliability of the calorimeter, ITC has become an important tool for the direct

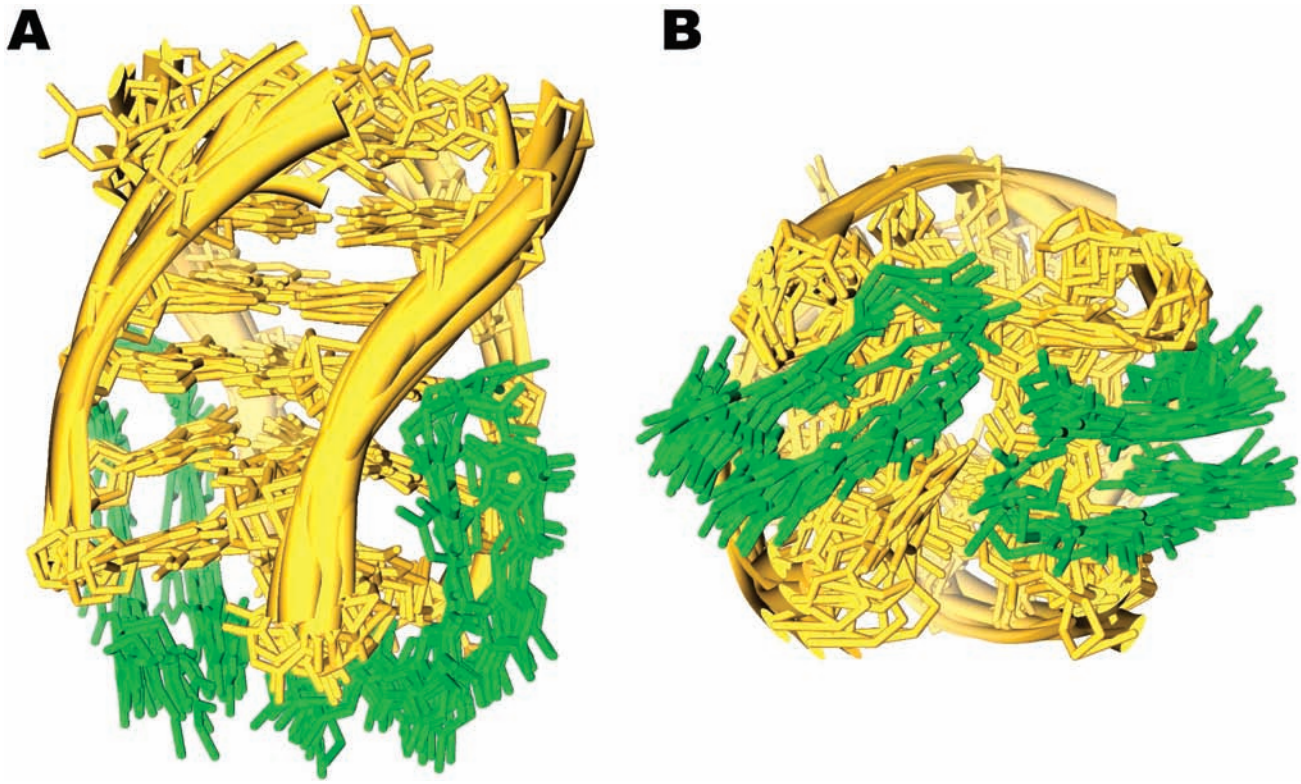


Figure 5. Side view (A) and bottom view (B) of the superimposition of the 10 best structures of the 4:1 **3**/[d(TGGGGT)]$_4$ complex. Derivative **3** is reported in green; DNA is colored in yellow.

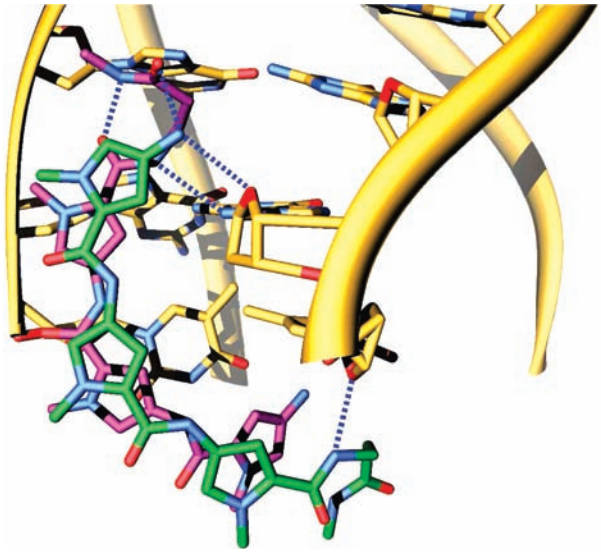


Figure 6. Side view representation of the 4:1 complex **3**/[d(TGGGGT)]$_4$. **3in**, **3ex**, and [d(TGGGGT)]$_4$ are depicted in magenta, green, and yellow sticks, respectively.

measurement of thermodynamic parameters in several biological interactions. Moreover, it has been proven to be useful to study the energetic aspects of interactions between G-quadruplexes and other molecules, including small ligands.[27]

Therefore, we have utilized ITC to characterize the thermodynamics of binding of derivative **3** to [d(TGGGGT)]$_4$ quadruplex. Figure 7 shows the binding isotherm resulting from the integration of raw data after correction for the heat of ligand dilution. The calorimetric data indicate an endothermic interaction. After a number of injections of ligand, increasingly less heat uptake was observed until constant values were obtained (corresponding to the heat of dilution), reflecting a saturable process. The values of K_b and $\Delta H°$ derived by ITC enable us to complete the thermodynamic binding profiles by calculating corresponding values of $T\Delta S°$ and $\Delta G°$. The resulting thermodynamic parameters are showed in Table 4. The binding affinity measured from ITC is 9×10^5 M^{-1} at 25 °C. The binding enthalpy was found to be positive ($\Delta H° = 6$ kJ mol^{-1}) and the stoichiometry of interaction was found to be 4:1 ligand−quadruplex. Inspection of the data reveals that the entropic contribution ($T\Delta S° = 40$ kJ mol^{-1}) to the binding of derivative **3** to the quadruplex provides the driving force for the ligand−quadruplex

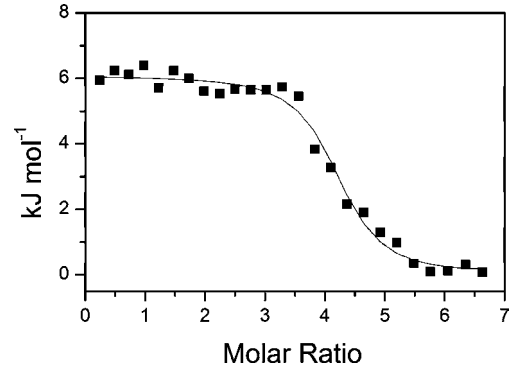


Figure 7. Normalized heat of interaction between compound **3** and [d(TGGGGT)]$_4$. The squares represent the experimental data obtained by integrating the raw data and subtracting the heat of ligand dilution into the buffer. The line represent the best fit obtained by a nonlinear least-squares procedure based on a single set of identical sites model.

(27) Pagano, B.; Mattia, C. A.; Giancola, C. *Int. J. Mol. Sci.* **2009**, *10*, 2935–2957.

Table 4. Thermodynamic Parameters for the Interaction of Compound **3** with $[d(TGGGGT)]_4$ Determined by ITC at 25 °C[a]

drug	n	K_b (M^{-1})	ΔH° (kJ mol^{-1})	$T\Delta S^\circ$ (kJ mol^{-1})	ΔG°_{298K} (kJ mol^{-1})
3	4	9×10^5	6	40	−34

[a] The experimental error on each thermodynamic property is <10%.

binding reaction, while the positive heat of formation of the complex indicates that this interaction is an enthalpically unfavorable process.

The results from the calorimetric titration data alone do not allow us to demonstrate with certainty the binding mode of this ligand. Usually, small molecules bind to duplex DNA by two predominant binding modes, intercalation and groove binding. Calorimetric studies have determined the enthalpic and entropic contributions to the DNA binding of representative DNA binding compounds. Generally, the binding of an intercalator to DNA is driven by a large favorable enthalpy change and by an unfavorable entropy decrease, while the binding of a groove binder to DNA is characterized by a large favorable increase of entropy and a small favorable or unfavorable enthalpy change.[28] In this case, the thermodynamic data determined by ITC for the binding of derivative **3** to $[d(TGGGGT)]_4$ (large favorable increase of entropy and small unfavorable enthalpy change) are in agreement with the groove binding profile and are confirmed by the structure of the complex.

The thermodynamic parameters for the interaction of derivative **3** are slightly different from the ones determined by ITC for the binding of distamycin A to the same quadruplex molecule. In both cases, the driving force for the binding reaction is the entropy change. However, the binding of derivative **3** shows a small unfavorable enthalpy change, while in the case of distamycin, a small favorable enthalpy change was observed. This interesting difference is probably due to structural features of the two ligands. Indeed, the positively charged amidinium moiety of distamycin interacts with the phosphate groups of the quadruplex, providing a favorable (although small) enthalpy contribution, while in the derivative **3**, the amidinium group is replaced by the uncharged N-methyl amide moiety that can not give this contribution.

In summary, the structural data herein reported unveil the role of the Coulombic interaction engaged by distamicyn A and the quadruplex. Such an interaction not only influences the strength of ligand/quadruplex formation, but surprisingly governs the orientation of the ligand with respect to the DNA. In fact, the removal of the positively charged terminal group results in an unprecedented ligand binding position in which both the groove and the 3′ end of the DNA are occupied. Furthermore, the lack of charge in the ligand does not affect the relative orientation of two molecules of **3** forming a dimer. All this opens up the tempting opportunity for a fine-tuning of the drug−quadruplex interaction mode. Moreover, with this further contribution, we provided the scientific community with an enhanced knowledge of the structural and conformational demands of the quadruplex groove that will serve as a platform for a rational design of novel groove binding agents.

Acknowledgment. This work is supported by Italian Institute of Technology (IIT), Italian M.U.R.S.T. (P.R.I.N. 2007), Regione Campania (L.41, L.5). The authors are grateful to "Centro di Servizio Interdipartimentale di Analisi Strumentale", C.S.I.A.S., for supplying NMR facilities.

Supporting Information Available: Tables reporting ^{1}H NMR assignment and enlarged NOESY spectra; time-dependent rmsd for molecular dynamic calculations; additional figures. This material is available free of charge via the Internet at http://pubs.acs.org.

JA1003872

(28) Chaires, J. B. *Arch. Biochem. Biophys.* **2006**, *453*, 26–31.