Theoretical Study of the Guanine → 6-Thioguanine Substitution in Duplexes, Triplexes, and Tetraplexes

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · NOVEMBER 2004

Impact Factor: 12.11 · DOI: 10.1021/ja0468628 · Source: PubMed

4 AUTHORS, INCLUDING:

Elena Cubero

31 PUBLICATIONS **1,147** CITATIONS

SEE PROFILE

J|A|C|S
ARTICLES
Published on Web 10/16/2004

Theoretical Study of the Guanine → 6-Thioguanine Substitution in Duplexes, Triplexes, and Tetraplexes

Nad'a Špačková,[†] Elena Cubero,[‡,§] Jiří Šponer,*[,†] and Modesto Orozco*[,‡,§]

Contribution from the Institute of Biophysics, Academy of Sciences of the Czech Republic, Kralovopolska 135, 612 65 Brno, Czech Republic, Departament de Bioquímica i Biologia Molecular, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1, Barcelona 08028, Spain, and Molecular Modeling and Bioinformatic Unit, Institut de Recerca Biomèdica, Parc Científic de Barcelona, Josep Samitier 1-5, Barcelona 08028, Spain

Received May 27, 2004; E-mail: modesto@mmb.pcb.ub.es; sponer@ncbr.chemi.muni.cz

Abstract: Molecular dynamics and thermodynamic integration calculations have been carried out on a set of G-rich single-strand, duplex, triplex, and quadruplex DNAs to study the structural and stability changes connected with the guanine → 6-thioguanine (G → ***S***) mutation. The presence of 6-thioguanine leads to a shift of the geometry from the B/A intermediate to the pure B-form in duplex DNA. The G → ***S*** mutation does not largely affect the structure of the antiparallel triplex when it is located at the reverse-Hoogsteen position, but leads to a non-negligible local distortion in the structure when it is located at the Watson−Crick position. The G → ***S*** mutation leads to destabilization of all studied structures: the lowest effect has been observed for the G → ***S*** mutation in the reverse-Hoogsteen strand of the triplex, a medium effect has been observed in the Watson−Crick strand of the triplex and duplex, and the highest influence of the G → ***S*** mutation has been found for the quadruplex structures.

Introduction

6-Thioguanine (abbreviated as ***S*** throughout this paper) is one of the most important therapeutic agents used in the clinical treatment of acute childhood lymphoblastic leukemia, inflammatory bowel disease, Crohn's disease, AIDS, and some other pathologies.[1−11] Very recently, the anti-angiogenic properties of 6-thioguanine have been proved, which might even increase the range of therapeutic actions of this fascinating molecule.[8] The therapeutic effects of ***S*** are believed to be related to its incorporation into the DNA,[12−16] but not many details regarding structural and functional consequences of such incorporation are known. Some authors have suggested that the main effect of ***S*** might be related to its ability to block quadruplex formation, interfering then in telomere structure.[17] On the contrary, others have suggested that the inhibition of RNase H and/or topoisomerase II is the main mechanism of cytotoxicity of ***S***.[3−5] In any case, it is clear that all these effects should be related to the ability of this molecule to distort normal or unusual nucleic acid structures.

Experimental structural information on thioguanine-containing nucleic acids is limited to DNA duplexes containing this nucleobase paired to cytosine or thymine.[17−19] All NMR data suggest that the G → ***S*** substitution leads to only moderate changes in the stability of the DNA duplex,[17−19] but discrepancy exists in the magnitude of the structural change induced by the presence of 6-thioguanine in the duplex. Thus, Marathias et al.[17] suggested that the G → ***S*** substitution largely distorts the structure of the DNA duplex, while Somerville et al.[18] and Bohon et al.[19] suggested that it has only a small and very local impact on DNA structure. All NMR data suggest that the presence of ***S*** leads to partial opening of the ***S***·C Watson−Crick base pair into the major groove, probably as a consequence of

[†] Academy of Sciences of the Czech Republic.
[‡] Universitat de Barcelona.
[§] Parc Científic de Barcelona.

(1) Inamochi, H.; Higashigawa, M.; Shimono, Y.; Nagata, T.; Cao, D. C.; Mao, X. Y.; M'soka, T.; Hori, H.; Kawasaki, H.; Sakurai, M. *J. Exp. Clin. Cancer Res.* **1999**, *18*, 417−424.
(2) Glaab, W. E.; Risinger, J. I.; Umar, A.; Barrett, J. C.; Kundel, T. A.; Tindall, K. R. *Carcinogenesis* **1998**, *19*, 1931−1937.
(3) Krynetskaia, N. F.; Krynetski, E. Y.; Evans, W. E. *Mol. Pharmacol.* **1999**, *56*, 841−848.
(4) Krynetskaia, N. F.; Cai, X. J.; Nitiss, J. L.; Krynetski, E. Y.; Relling, M. V. *FASEB J.* **2000**, *14*, 2339−2344.
(5) Krynetskaia, N. F.; Feng, J. Y.; Krynetski, E. Y.; Garcia, J. V.; Panetta, J. C.; Anderson, K. S.; Evans, W. E. *FASEB J.* **2001**, *15*, 1902−1908.
(6) Karran, P. *Carcinogenesis* **2001**, *22*, 1931−1937.
(7) Massey, A.; Xu, Y. Z.; Karran, P. *DNA Repair* **2002**, *1*, 275−286.
(8) Presta, M.; Belleri, M.; Vacca, A.; Ribatti, D. *Leukemia* **2002**, *16*, 1490−1499.
(9) Herrlinger, K. R.; Kreisel, W.; Schwab, M.; Schoelmerich, J.; Fleig, W. E.; Ruhl, A.; Reinshagen, M.; Deibert, P.; Fellermann, K.; Greinwald, R.; Stange, E. F. *Aliment. Pharmacol. Ther.* **2003**, *17*, 503−508.
(10) Dubinsky, M. C.; Feldman, E. J.; Abreu, M. T.; Targan, S. R.; Vasiliauskas, E. A. *Am. J. Gastroenterol.* **2003**, *98*, 1058−1063.
(11) Dervieux, T.; Blanco, J. G.; Krynetski, E. Y.; Vanin, E. F.; Roussel, M. F.; Relling, M. V. *Cancer Res.* **2001**, *61*, 5810−5816.
(12) Bostrom, B.; Erdmann, G. *Am. J. Pediatr. Hematol./Oncol.* **1993**, *15*, 80−86.
(13) Waters, T. R.; Swann, P. F. *Biochemistry* **1997**, *36*, 2501−2506.
(14) Karran, P.; Bignami, M. *Bioessays* **1994**, *16*, 833−839.
(15) Karran, P.; Bignami, M. *Chem. Biol.* **1996**, *3*, 875−879.
(16) Swann, P. F.; Waters, T. R.; Moulton, D. C.; Xu, Y. Z.; Zheng, Q. G.; Edwards, M.; Mace, R. *Science* **1996**, *273*, 1109−1111.
(17) Marathias, V. M.; Sawicki, M. J.; Bolton, P. H. *Nucleic Acids Res.* **1999**, *27*, 2860−2867.
(18) Somerville, L.; Krynetski, E. Y.; Krynetskaia, N. F.; Beger, R. D.; Zhang, W. X.; Marhefka, C. A.; Evans, W. E.; Kriwacki, R. W. *J. Biol. Chem.* **2003**, *278*, 1005−1011.
(19) Bohon, J.; de los Santos, C. R. *Nucleic Acids Res.* **2003**, *31*, 1331−1338.

10.1021/ja0468628 CCC: $27.50

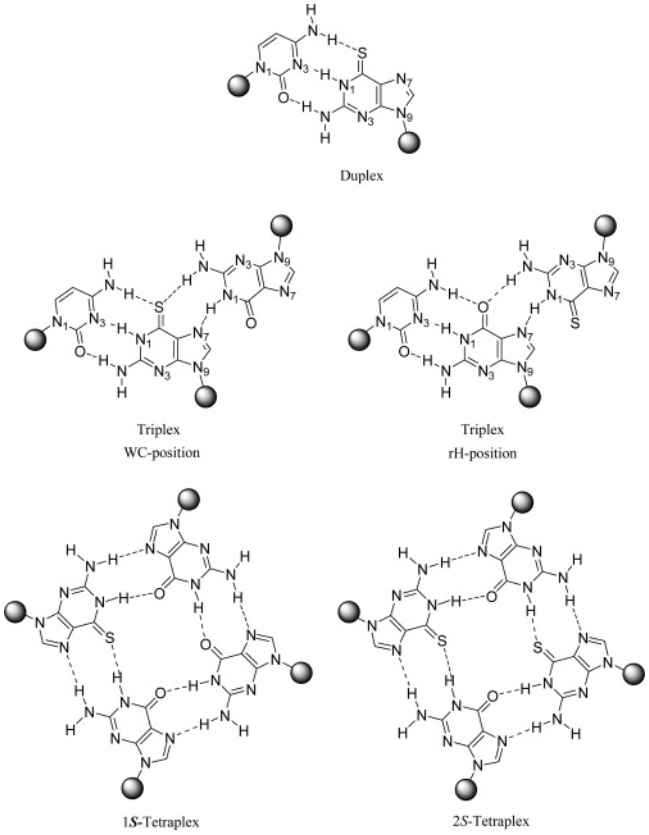


Figure 1. Structure of duplexes, triplexes, and quadruplexes containing 6-thioguanine studied in this work.

the larger size of the sulfur atom.[17−19] Particularly, NMR-exchange measurements by Somerville et al.[18] indicated an anomalous fast proton exchange rate for the ***S***·C base pairs, corresponding to a ca. 80-fold decrease of the base pair lifetime,[18] which might suggest a lower stability of the ***S***·C pair compared with the G·C pair.

Recent experimental studies suggest that the presence of thioguanine (see Figure 1) largely destabilizes quadruplexes, with a small impact on the stability of antiparallel triplexes.[17,20−22] This finding opens a wide range of possibilities for 6-thioguanine in anti-gene therapies (i.e., strategies based on the inhibition of pathological genes by triplex formation), since the generation of antiparallel triplexes based on the d(G#G·C) motif is in many cases handicapped (see Figure 2) by the formation of intramolecular quadruplexes in the triplex-forming oligonucleotide (TFO).[20−22]

The gas-phase properties and molecular interactions of 6-thioguanine have been studied by means of ab initio methods.[23] It was found that 6-thioguanine has tautomeric properties very similar to those of guanine, and that its electrostatic distribution is also similar, as noted in the dipole moment, oriented in the same direction and only slightly larger for 6-thioguanine than for guanine (7.8 D vs 6.6 D).[23] Ab initio calculations also demonstrated that, at least in the gas phase, 6-thioguanine can establish stable stacking interactions, and that

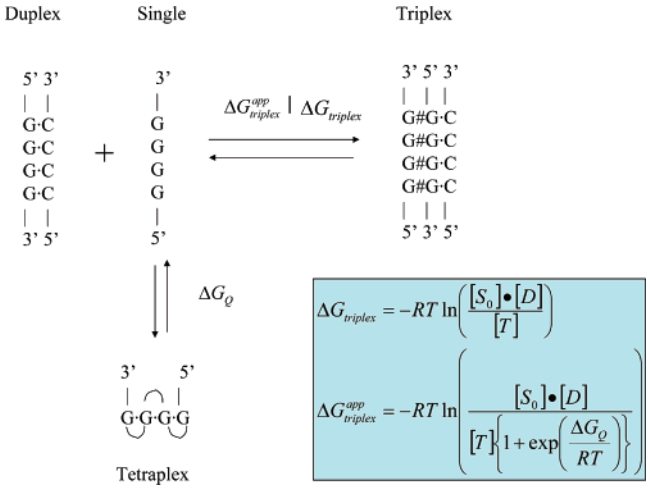


Figure 2. Schematic representation of the duplex ↔ triplex transition in the absence ($\Delta G_{triplex}$) or presence ($\Delta G_{triplex}^{app}$) of quadruplex competition. The corresponding dissociation constants are displayed at the left.

the ***S***·C Watson−Crick pair is of stability similar to that of the G·C pair (−23.4 kcal/mol vs −25.5 kcal/mol; from ref 23). Ab initio calculations in the gas phase suggested that the major differences between G and ***S*** were related to the ability of both purines to interact with cations and water molecules. Thus, it was suggested that the less polar S6 group reduces hydration of the ***S***·C pair in the major groove side compared with that of the G·C pair.[23] On the other hand, the presence of a sulfur atom increases the ability of 6-thioguanine to interact with soft metal cations, such as mercury and cadmium, compared to that of guanine.[24] It remains to be investigated to what extent they are affecting nucleic acids in solution.

To our knowledge only one extended molecular dynamics (MD) study on nucleic acids containing 6-thioguanine has been published.[25] The analysis of several trajectories of DNA quadruplexes of the types d(GGGG)$_4$, d(***SSSS***)$_4$, and d(G***S***GG)$_4$ suggested that tetrads containing four 6-thioguanines are unstable probably for steric reasons while the presence of one 6-thioguanine in a tetrad formed with three guanines is tolerated.

In the present paper we report an extended MD study of molecular interactions of 6-thioguanine in DNA duplexes, triplexes, and tetraplexes. MD coupled to thermodynamic integration (MD/TI) calculations provide for the first time quantitative estimates of the change in stability of duplexes, triplexes, and tetraplexes arising from the substitution of one guanine by one 6-thioguanine. The impact of the results in the design of more efficient triplex-based anti-gene strategies is discussed.

Methods

We have studied the effect of 6-thioguanine on a wide range of molecules, the majority of them containing stretches of guanines. All studied sequences are summarized in Table 1, with sites of G ↔ ***S*** mutation highlighted.

Starting Structures. The duplex structures were constructed in the standard B-like conformation using the Nucgen module of AMBER6.[26]

(20) Rao, T. S.; Durland, R. H.; Seth, D. M.; Myrick, M. A.; Bodepudi, V.; Revankar, G. R. *Biochemistry* **1995**, *34*, 765−772.
(21) Olivas, W. M.; Maher, L. J. *Nucleic Acids Res.* **1995**, *23*, 1936−1941.
(22) Gee, J. E.; Revankar, G. R.; Rao, T. S.; Hogan, M. E. *Biochemistry* **1995**, *34*, 2042−2048.
(23) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. A* **1997**, *101*, 9489−9495.
(24) Sponer, J.; Burda, J. V.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1999**, *17*, 61−77.
(25) Stefl, R.; Spackova, N.; Berger, I.; Koca, J.; Sponer, J. *Biophys. J.* **2001**, *80*, 455−468.
(26) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Ross, W. S.; Simmerling, C. L.; Darden, T. L.; Marz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER6*; University of California: San Francisco, 1999.

Table 1. Oligonucleotides Investigated in the Course of This Study

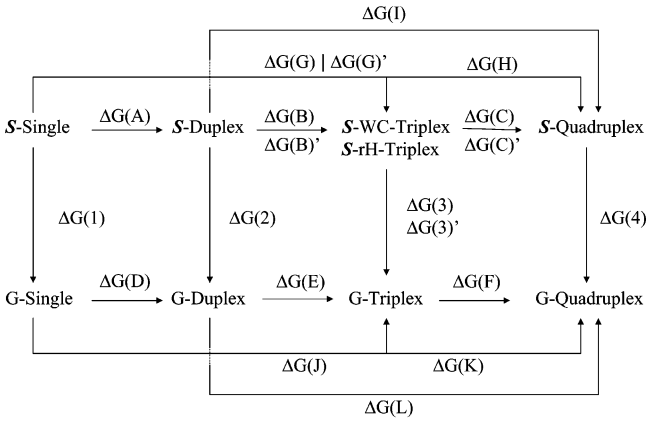
Number of strands	Simulation name	Sequences of all strands [a]	TI	Simulation time [b] (ns)	Additional information
1	Single strand	5′GG***S***GT3′	yes	-	
2	G-duplex	5′GTTTGGGGTTTG3′ 3′CAAACCCCAAAC5′	no	10	Reference simulation without presence of ***S***
2	***S***-duplex	5′GTTTGG***S***GTTTG3′ 3′CAAACCCCAAAC5′	yes	10	
2	4***S***-duplex	5′GTTT***SSSS***TTTG3′ 3′CAAACCCCAAAC5′	no	10	
2	Somerville	5′CTAAG***S***AAAG3′ 3′GATTCCTTTC5′	no	10	Structure based on the sequence studied by Somerville et al.
2	Marathias	5′GC***S***AATTC***S***CG3′ 3′CGCTTAAGCGC5′	no	10	Structure based on the sequences studied by Marathias et al. and Bohon et al.
3	G-triplex	5′GGGGGGGGGG3′ 3′CCCCCCCCCC5′ 3′GGGGGGGGGG5′	no	5	Reference simulation without presence of ***S***
3	***S***-WC-triplex	5′GGGG***S***GGGGG3′ 3′CCCCCCCCCC5′ 3′GGGGGGGGGG5′	yes	5	Mutation in the Watson-Crick strand of the triplex structure
3	***S***-rH-triplex	5′GGGGGGGGGG3′ 3′CCCCCCCCCC5′ 3′GGGG***S***GGGGG5′	yes	5	Mutation in the Hoogsteen strand of the triplex structure
4	1***S***-quadruplex	5′GG***S***G3′ 5′GGGG3′ 5′GGGG3′ 5′GGGG3′	yes	-	Parallel stranded quadruplex in presence of Na^+ - 1st mutation in the guanine quartet
4	1***S***-quadruplex-K	5′GG***S***G3′ 5′GGGG3′ 5′GGGG3′ 5′GGGG3′	yes	-	Parallel stranded quadruplex in presence of K^+ - 1st mutation in the guanine quartet
4	1***S***-quadruplex-a	5′GG***S***G3′ 3′GGGG5′ 5′GGGG3′ 3′GGGG5′	yes	-	Antiparallel stranded quadruplex in presence of Na^+ - 1st mutation in the guanine quartet
4	1***S***-quadruplex-aK	5′GG***S***G3′ 3′GGGG5′ 5′GGGG3′ 3′GGGG5′	yes	-	Antiparallel stranded quadruplex in presence of K^+ - 1st mutation in the guanine quartet
4	2***S***-quadruplex	5′GG***S***G3′ 5′GGGG3′ 5′GG***S***G3′ 5′GGGG3′	yes	-	2nd mutation in the same guanine quartet but in the opposite strand of the 1***S***-quadruplex (1st base remains mutated)

[a] When TI calculation was carried out, the sequence of the structure after mutation is presented. [b] The simulation time is mentioned only for simulations where structural analysis has been carried out.

For single-strand simulation needed to calculate the thermodynamics data (see below) only the central G_4T part of the duplex strand was used. The initial triplex geometry was generated using our previously MD equilibrated conformations[27] which were originally built on the basis of Patel's NMR structure of TAGGAGGAT.[28] In the case of quadruplexes, both parallel and antiparallel arrangements have been studied. Initial coordinates for the parallel-stranded quadruplex are based on the crystal structure of $d(TGGGGT)_4$, where all terminal thymines were omitted.[29] The structure of the antiparallel quadruplex was taken from the NMR structure of $d(GGGGTTTTGGGG)_2$.[30] Again, only the central guanine stem without the attached thymine loops has been used in the simulations. The quadruplex channel is in all cases occupied by three ions located in cavities between adjacent quartets.

Force Field. AMBER parm99[31,32] and TIP3P[33] parameters were used to describe the interactions between standard residues of DNA. Atomic charges defining the electrostatic interactions of 6-thioguanine were obtained by fitting HF/6-31G(d) potentials following RESP methodology.[34] On the basis of ab initio calculations, the equilibrium length for

(27) Aviñó, A.; Cubero, E.; González, C.; Eritja, R.; Orozco, M. *J. Am. Chem. Soc.* **2003**, *125*, 16127−16138.
(28) Radhakrishnan, I.; Patel, D. J. *J. Am. Chem. Soc*. **1993**, *115*, 1615−1617.
(29) Phillips, K.; Dauter, Z.; Murchie, A. I. H.; Lilley, D. M. J.; Luisi, B. *J. Mol. Biol*. **1997**, *273*, 171−182.
(30) Schultze, P.; Smith, F. W.; Feigon, J. *Structure* **1994**, *2*, 221−233.
(31) Cheatham, T. E.; Cieplak, P., Kollman, P. A. *J. Biomol. Struct. Dyn*. **1999**, *16*, 845−862.
(32) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.: Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc*. **1995**, *117*, 5179−5197.
(33) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys*. **1983**, *79*, 926−935.
(34) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.



$\Delta G(A) - \Delta G(D) = \Delta G(1) - \Delta G(2)$ $\quad$ $\Delta G(G) - \Delta G(J) = \Delta G(1) - \Delta G(3)$

$\Delta G(B) - \Delta G(E) = \Delta G(2) - \Delta G(3)$ $\quad$ $\Delta G(H) - \Delta G(K) = \Delta G(1) - \Delta G(4)$

$\Delta G(C) - \Delta G(F) = \Delta G(3) - \Delta G(4)$ $\quad$ $\Delta G(I) - \Delta G(L) = \Delta G(2) - \Delta G(4)$

Figure 3. Thermodynamic cycles characterizing the impact of G → ***S*** substitution in the transition of different multimeric forms of DNA. The vertical terms ($\Delta G(1)$, $\Delta G(1)$, $\Delta G(2)$, $\Delta G(3)$, $\Delta G(3)'$, and $\Delta G(4)$) are those computed by MD/TI calculations.

the C−S bond was set to 1.6589 Å and the stretching force constant was 570 kcal/(mol·Å^2). The van der Waals parameters of the sulfur atom of 6-thioguanine were fitted to extensive ab initio data for H-bonding and stacking of thioguanine.[23] Final values were 2.0 Å for the van der Waals radius and 0.5 kcal/mol for the potential well depth. The final parameters (see Table S1 in the Supporting Information) reproduce very well MP/6-31G*(0.25) results for stacking and hydrogen bond interaction energies (see Table S2 in the Supporting Information). For example, the best stacking energy (for a twist angle of 150°) for the S−S dimer is −12.5 kcal/mol (AMBER) and −12.7 kcal/mol (MP/6-31G*(0.25)). The hydrogen bond energies of the G·C and S·C dimers are −21.7 and −20.6 kcal/mol, respectively, values which compare well with MP/6-31G*(0.25) calculations: −23.4 and −22.5 kcal/mol, respectively.

Molecular Dynamics. Starting geometries of the different nucleic acids were immersed in large boxes of TIP3P water molecules extending around 12 Å in each direction from the solute. The systems were then neutralized by a minimum number of sodium or potassium ions, placed initially in the regions of the most electronegative potential. The systems were then thermalized and equilibrated using standard protocols.[35] Finally, the equilibrated systems were subjected to unrestrained MD simulations for 1 ns. These structures were used as starting points for MD/TI simulations (see below). Some of the trajectories were extended as noted in Table 1 to collect structural data. All simulations were performed using periodic boundary conditions and the particle mesh Ewald method.[36] SHAKE[37] was used to maintain all the bonds at their equilibrium lengths, which allowed us to use 2 fs for integration of Newton's equations.

Thermodynamic Integration. TIs were carried out using a very lengthy and conservative protocol.[38] Accordingly, each perturbation was done both in the G → ***S*** direction and in the ***S*** → G direction. To even improve the statistical quality of the simulation, and to reduce the noise, each of the individual mutations was done twice: starting from independent configurations and running for 420 and 840 ps using in all cases 21 windows. Every window was divided into two halves, each used to obtain an independent estimate of the free energy change. Thus, for each mutation eight independent numbers were averaged to obtain each free energy value. The stabilizing effect of the G ↔ ***S*** substitution on different structures was explored using standard thermodynamic cycles (see Figure 3) from mutation data. All nonbonded interactions (including intragroup contributions) were considered dependent on the mutation coordinate, which implies that all the free energy values needed to be corrected by the gas-phase values.

Analysis. Trajectories have been analyzed using the Carnal and Ptraj modules of AMBER6 as well as specific *in-house* software. Helical base pair and base pair step parameters have been calculated by the 3DNA program.[39] Helical parameters have been calculated for all trajectory frames as well as for MD-averaged structures. Hydration and ion distribution have been analyzed by evaluation of water density maps and molecular interaction potential (MIP)[40] profiles. Residence times were evaluated following the standard procedure in the Carnal module of AMBER.[26] Accordingly, waters around selected polar atoms of DNA were assigned to the nearest atoms in consecutive snapshots, analyzing the period of time which the water occupies a given binding position.

Results and Discussion

Structural Analysis of Duplexes. On the basis of 3−5 ns trajectories, Stefl et al.[25] described the structural impact of the presence of 6-thioguanine in DNA tetraplexes, suggesting that tetrads formed by four 6-thioguanines are unstable due to steric hindrance of the sulfur atoms, which leads to the expulsion of ions from the central channel, resulting in major distortion and likely collapse of the structure. The same authors also concluded that the presence of only one 6-thioguanine in a tetrad is tolerated without any major distortions in the structure of the G-DNA.[25] We will complete the analysis here by inspecting the structural effects of the introduction of ***S*** in duplex and triplexes (at both the Watson−Crick and reverse-Hoogsteen positions).

We first analyzed the structure of three duplexes, d(GTTT-GGGGTTTG), d(GTTTGG***S***GTTTG), and d(GTTT***SSSS***TTTG), noted as G-, ***S***-, and 4***S***-duplexes in the following, using 10 ns samplings (detailed structural results are shown in Table S3 of the Supporting Information). Average RMSDs from the canonical B-form (base pairs at both ends of the molecule were omitted) are 3.6 Å (G-duplex), 2.9 Å (***S***-duplex), and 2.0 Å (4***S***-duplex). The RMSDs reduce to 1.8 Å (G-duplex), 1.7 Å (***S***-duplex), and 1.2 Å (4***S***-duplex) if only the central four base pairs are considered. As a reference, the RMSDs of the central four steps (central quartet) with respect to the canonical A-form are 2.0 Å (G-duplex), 2.2 Å (***S***-duplex), and 2.4 Å (4***S***-duplex). The large RMSD with respect to the B-form of the G-duplex, and the fact that for the central quartet of the G-duplex RMSD(A-form) is equal to RMSD(B-form), confirms that G-tracts have an intermediate B/A character.[41−45] This is clear in a positive roll of ca. 5°, a negative slide up to −1.5 Å, a base pair inclination of 10°, and a low helical twist (around 30°) for the pure G-track (G-duplex; see Table S3 in the Supporting Information).

The introduction of ***S*** produces a displacement of the structure of the duplex toward the B-form. Thus, RMSd(B) decreases by

(35) Shields, G. C.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc*. **1997**, *119*, 7463−7469.
(36) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys*. **1993**, *98*, 10089−10092.
(37) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(38) (a) Blas, J. R.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2004**, *126*, 154−164. (b) Cubero, E.; Laughton, C. A.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2000**, *122*, 6891−6899.
(39) Lu, X. J.; Olson, W. K. *Nucleic Acids Res.* **2003**, *31*, 5108−5121.
(40) Gelpi, J. L.; Kalko, S. G.; Barril, X.; Cirera, J.; De la Cruz, X.; Luque, F. J.; Orozco, M. *Proteins: Struct., Funct., Genet.* **2001**, *45*, 428−437.
(41) Dornberger, U.; Spackova, N.; Walter, A.; Gollmick, F. A.; Sponer, J.; Fritzsche, H. *J. Biomol. Struct. Dyn.* **2001**, *19*, 159−174.
(42) Sponer, J.; Florian, J.; Ng, H.-L.; Sponer, J. E.; Spackova, N. *Nucleic Acids Res*. **2000**, *28*, 4893−4902.
(43) Cheatham, T. E.; Brooks, B. R. *Theor. Chem. Acc*. **1998**, *99*, 279−288.
(44) Lankas, F.; Sponer, J.; Hobza, P.; Langowski, J. *J. Mol. Biol*. **2000**, *299*, 695−709.
(45) Ng, H.-L.; Dickerson, R. E. *Nucleic Acids Res*. **2002**, *30*, 4061−4067.

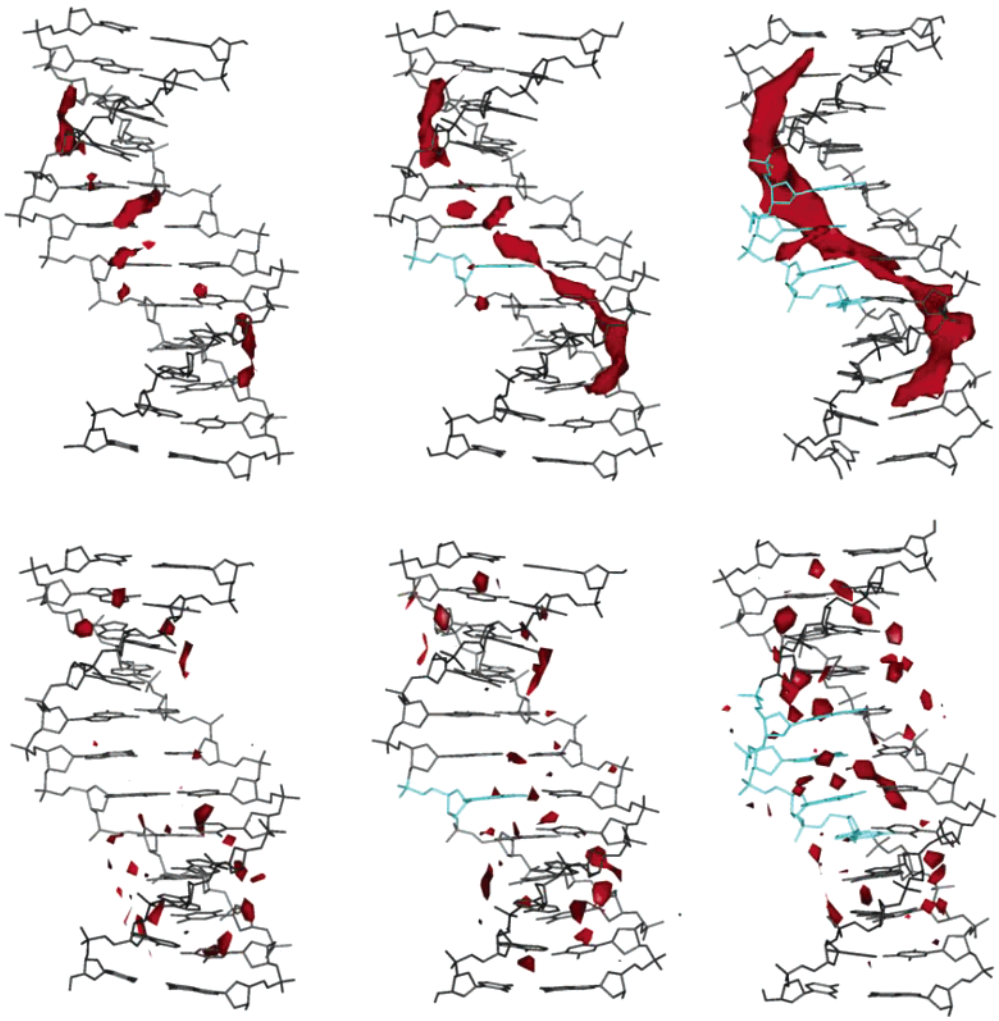


Figure 4. MIP profiles (top) and solvation maps (bottom) for (left to right) G-, ***S***-, and 4***S***-duplexes. The probe molecule for MIP calculations was O^+, and the contour level was −8.0 kcal/mol. A density contour of 2.0 g/mL is shown in the solvation maps.

0.6 Å while RMSd(A) increases by 0.3 Å when four 6-thioguanines are introduced. Thus, single G → ***S*** mutation leads to subtle changes in the structure of the G-tract, especially at the mutation site: an increase of 2−3° in the helical twist, a reduction of around 2° in the roll, a reduction of around 0.4 Å in the negative slide, and finally a reduction of around 3° in inclination (see Table S3). The local structural changes are more important when four 6-thioguanines are introduced (4***S***-duplex), since there is a substantial increase of the helical twist (around 5°), the average roll and slide are close to 0, and inclination is small and negative. In summary, the introduction of four 6-thioguanines drives the structure from the B/A to the pure B conformation. Not surprisingly, sugar puckerings change from 124° in the G-duplex to 134° in the ***S***-duplex and 143° in the 4***S***-duplex, signaling then the B/A → B transition upon 6-thioguanine introduction. Clearly, all these local changes are mostly related to the steric hindrance generated by the presence of the sulfur group at position 6, which leads to enlargement in both intra- and interstrand nucleobase−nucleobase distances (see Table S3).

6-Thioguanine narrows the minor groove by almost 2 Å (widths are measured as the shortest distance between phosphates on opposite strands following the 3DNA procedure), which leads to more electronegative minor grooves, as shown in MIP distributions in Figure 4, and in the depth of the MIP minima (differences of 1 and 2 kcal/mol between the G- and ***S***- and 4***S***-duplexes). The change in MIP, and the recovery of the large electronegative region in the minor groove upon introduction of ***S***, agrees with the displacement toward pure B-forms induced by 6-thioguanine.

The hydration shape in the major groove of the G-tract of the G-duplex is highly ordered, with the best hydration point located in the plane of the G·C pair. Regions of large water density are located in both major and minor grooves (Figures 4 and 5). These sites are occupied by water molecules with moderate residence times (around 0.1 ns). Three hydration sites are located in the minor groove and three others in the major groove of the G·C pairs (see Figure 5). The presence of 6-thioguanine leads to distortions of the hydration pattern, which are especially important in the major groove, and which are mostly related to the poorer hydration of the sulfur S6 atom compared to the oxygen O6 atom. This is especially clear from inspection of the solvation map of the 4***S***-duplex, where the density of water around the S6 atom is clearly reduced compared to that around the O6 atom (see Figure 5). This result confirms the earlier suggestions based on gas-phase ab initio calculations.[23] No dramatic changes are found in the average residence times of waters bound to G·C and ***S***·C pairs.

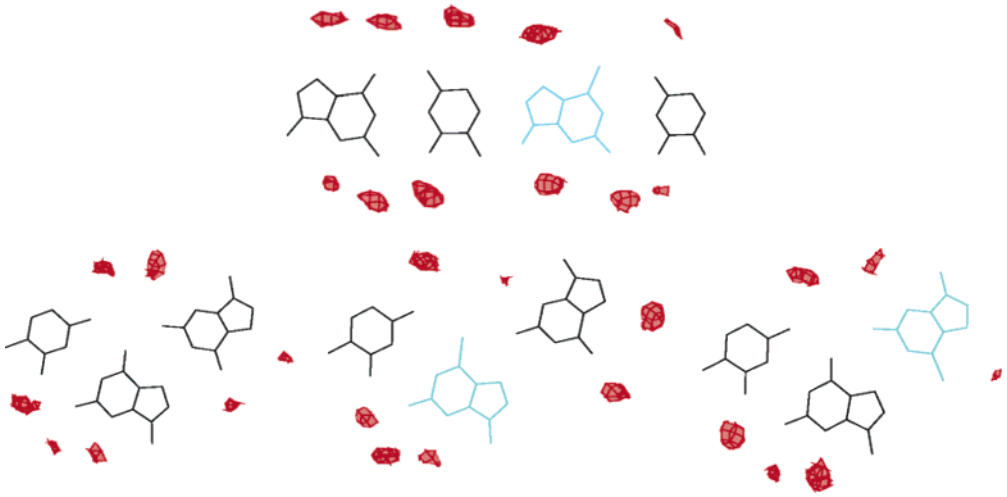


Figure 5. Details of solvation in the plane of dimers and trimers containing guanine or 6-thioguanine (in blue). The solvation contour corresponds to a density of 3.6 g/mL.

In this paper we are mostly interested in the impact of G → ***S*** substitution in the G-tract, but we considered it also interesting to analyze the impact of this mutation in duplexes where the substitution site is not surrounded by other guanines. The first corresponds to the sequence studied by Somerville,[18] and the second corresponds to Marathias's duplex with one mutated base in the Cp***S***pA step and the other in the Cp***S***pC step (see Table 1), mimicking the type of mutation studied by Bohon et al.[19] Simulations were performed using the same procedure as for G-tracts and demonstrated that most effects caused by the G → ***S*** substitution in the non-G-tract segments were similar to those for the G-tracts. All trajectories were stable on the nanosecond time scale, with RMSD(B) around 2.8−2.9 Å (common RMSD range for MD simulations of duplexes), suggesting that also here the presence of 6-thioguanine does not introduce major alterations to the structure. It is worth noting that MD simulations do not support the existence of anomalous helical parameters for unmodified (d(CTAAGGAAAG)·d(CTTTCCTTAG)) and modified (d(CTAAG***S***AAAG)·d(CTTTCCTTAG)) duplexes, as Somerville's data suggested.

Structural Analysis of Triplexes. We have analyzed the structural impact of 6-thioguanine in the Watson−Crick (WC) or reverse-Hoogsteen (rH) positions of antiparallel triplexes (i.e., d(G#***S***·C) and d(***S***#G·C); see Figure 1). For this purpose we built three triplexes with sequences $d(G)_{10}$#$d(G)_{10}$·$d(C)_{10}$ (G-triplex), $d(G)_{10}$#$d(G_4$-***S***-$G_5)$·$d(C)_{10}$ (***S***-WC-triplex), and $d(G_4$-***S***-$G_5)$#$d(G)_{10}$·$d(C)_{10}$ (***S***-rH-triplex), and analyzed them using molecular dynamics (see Table 1 and the Methods). Control simulations (5 ns each) were performed using a 1 M concentration of NaCl, but the results found were very similar to those obtained using a minimum amount of ions (which were needed to neutralize the system) and are not shown here (they are available upon request).

The three trajectories sample conformational regions close to the canonical antiparallel triplex[27] generated from the NMR structure,[28] as noted in the average RMSD (from the canonical structure) ranging from 2.1 Å (G-triplex) to 2.3 Å (***S***-rH-triplex). The average RMSD between the G-triplex trajectory and the MD-averaged conformation is 1.7 ± 0.5 Å, while it reduces to 1.3 ± 0.4 Å for triplexes containing ***S***. This suggests that the presence of 6-thioguanine leads to a small restriction in the flexibility of the helix. Interestingly, the cross RMSDs (obtained by comparing one trajectory with the MD-averaged structure of another) range from 1.8 to 2.3 Å, demonstrating that there are no dramatic changes in the structure of the triplex induced by the presence of 6-thioguanine.

Helical analysis of the triplexes (for technical reasons restricted always to the WC strands) confirms that no dramatic general changes in the structure occur when ***S*** is present (see Table S4 in the Supporting Information) either in the Watson−Crick or reverse-Hoogsteen positions. The only remarkable variations are related to a certain loss of planarity and partial opening effects in the d(G#***S***·C) triad. These distortions are related to small changes in the pattern of hydrogen bonds in the d(G#***S***·C) triad. Thus, the G-triplex canonical Watson−Crick and reverse-Hoogsteen hydrogen bonds were present for 99.8% and 95.6% of the simulation time, the loss of hydrogen bonding being related to local and fast atomic fluctuations and not to real openings of the triad. When ***S*** is at the Watson−Crick position, there is only a small reduction of the WC hydrogen bonding (a total of 98.9% of the time the WC hydrogen bonds are detected) while the loss of reverse-Hoogsteen hydrogen bonding is more visible (from 95.6% to 92.5%). All these modest effects can be rationalized by the steric repulsion of the sulfur atom, which produces fast atomic movements leading to partial opening movement at the Watson−Crick portion of the triad, and a slight displacement of the third nucleobase (guanine) apart from the Watson−Crick dimer (see Figure 6). Interestingly, the presence of 6-thioguanine in the reverse-Hoogsteen position does not modify at all the population of either Watson−Crick (99.6% population) or reverse-Hoogsteen (93.4% population) hydrogen bonds (see Figure 6).

In summary, structural analysis clearly suggests that the impact of the 6-thioguanine in the structure of the antiparallel triplexes is generally small and localized at the mutation site. Only when 6-thioguanine is present in the Watson−Crick position some local distortions of the triad are detected. However, we cannot ignore that even small conformational changes might induce large alterations in the ability of the triplex to interact with other molecules. Thus, to clarify this point, we computed the MIP profiles for the three antiparallel triplexes considered. As noted before,[27] for d(G#G·C) antiparallel triplexes, the minor part of the major groove (mM groove) is the best region for interaction with small cations (see Figure 7)

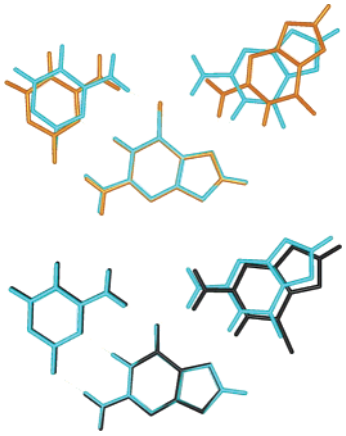


Figure 6. MD-averaged structure of triads containing 6-thioguanine compared with the unmodified d(G#G·C) triad: light blue, d(G#G·C); orange, d(G#***S***·C); black, d(***S***#G·C).

but favorable interactions are also found in the major part of the major groove (MM groove). When 6-thioguanine is present (at either the Watson−Crick or reverse-Hoogsteen positions), there is a decrease in the electronegativity of the MM groove but that located in the mM groove is mostly unaltered (see Figure 7).

Water density analysis (see Figures 5 and 7) shows a reasonable degree of hydration in all the triplexes, with large regions with high water density located in the minor groove (located between triads and making hydrogen bonds with N3 (G) and N2 (G) of stacked guanines) and smaller regions placed in the MM groove.[27] Quite surprisingly, despite its large electronegativity, the mM groove does not appear especially well solvated. The insertion of 6-thioguanine does not largely disturb the hydration pattern in the triplex, which can be understood considering that the thio group remains quite hidden in the triplex. The only remarkable change in hydration pattern is a marginal loss of highly structured waters when ***S*** is inserted into either Watson−Crick or reverse-Hoogsteen positions (see Figures 5 and 7).

Thermodynamic Calculations. One of the main advantages of computational techniques is their ability to directly relate the structural data with the energetics. Thus, MD/TI calculations were performed to analyze the impact of the G → ***S*** substitution on the stability of different multiplex structures (see the Methods). All simulations reported here appear well converged, as noted in the similarity of the free energy estimates obtained from different trajectories. Mutation profiles were smooth and without apparent discontinuities, and forward and reverse simulations yield very similar estimates, showing that hysteresis is not a major problem in our simulations. Clearly, the moderate magnitude of the mutation, the change from a C═O bond to a C═S bond, and our conservative approach (we have simulated

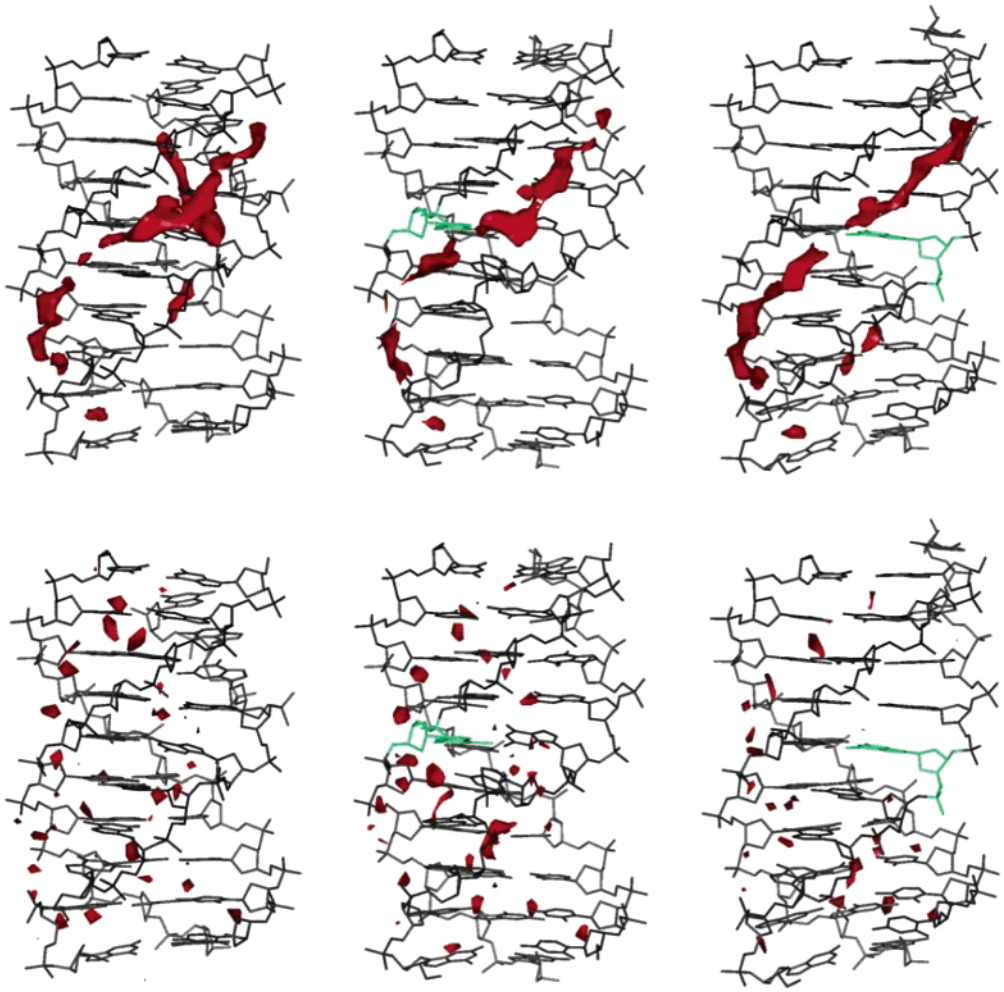


Figure 7. MIP profiles (top) and solvation maps (bottom) for (left to right) G-, ***S***-WC-, and ***S***-rH-triplexes. The probe molecule for MIP calculations was O^+, and the contour level was −7.0 kcal/mol. A density contour of 3.0 g/mL is shown in the solvation maps.

Table 2. Differences in the Free Energy Changes ($\Delta\Delta G$) Associated with G → ***S*** Mutation[a]

simulation name	free energy	$\Delta\Delta G$ (kcal/mol)	std dev (kcal/mol)
S-duplex	ΔG(A) − ΔG(D)	2.1	0.4
S-WC-triplex	ΔG(G) − ΔG(J)	2.4	0.3
S-rH-triplex	ΔG(G) − ΔG(J)	0.7	0.5
S-rH-triplex[b]	ΔG(G) − ΔG(J)	0.4	0.5
1***S***-quadruplex	ΔG(H) − ΔG(L)	4.4	0.3
1***S***-quadruplex[c]	ΔG(H) − ΔG(L)	3.9	0.2
1***S***-quadruplex-antiparallel[c]	ΔG(H) − ΔG(L)	3.9	0.6
2***S***-quadruplex[d]	ΔG(H) − ΔG(L)	3.4	0.3

[a] $\Delta\Delta G$ values are calculated for the single strand → multiplex (duplex, triplex, quadruplex) formation process using the thermodynamic cycles described in Figure 2. A positive value means that ***S*** destabilizes the multiplex structure with respect to the single strand and thus also relatively with respect to the multiplex with the natural base. The rest of the individual free energy differences in Figure 3 can be easily derived by taking advantage of the state function nature of the free energy from data in this table. [b] Calculation performed using 1 M NaCl solution. [c] Calculations performed using K^+ as the counterion. [d] The free energy value corresponds to the inclusion of the second 6-thioguanine in a central d(***S***G_3) tetrad of a parallel quadruplex.

only systems in which the mutation was not expected to destroy the structure of the multiplex) justify the good statistical quality of our theoretical estimates.

The G → ***S*** mutation destabilizes all the multiplex structures relative to the corresponding single-stranded oligonucleotides (see Table 2). The duplex is strongly destabilized (around 2.1 kcal/mol) by the G → ***S*** substitution. There is not a direct experimental measure of the change in stability/melting free energy of the duplex induced by the presence of 6-thioguanine, but a reasonable estimate can be obtained by using regression equations between melting free energies and temperatures obtained in duplexes of similar size and structure. Thus, taking data from Kool's group,[47] a regression equation can be obtained, $\Delta G = -0.254\Delta T_m + 0.339$ ($r^2 = 0.971$), relating differences in melting temperatures and stability free energies. Using this equation, we can predict that the G → ***S*** mutation should destabilize the duplex by around 6.9 °C, which agrees very well with rough estimates by Somerville et al. (around 6 °C in ref 18) and Marathias et al. (around 7 °C in ref 17). Analysis of H-bonding and stacking interactions shows that duplex destabilization (see Table 3) by 6-thioguanine originates in the lower stability of its base pairing (around 3 kcal/mol per substitution) in the substituted site, a fact that was already detected in previous gas-phase ab initio calculations.[23] The loss of stability is not compensated for by improved stacking interactions occurring as a consequence of the G → ***S*** substitution (see Table 3).

The profound impact of 6-thioguanine in the stability of the duplex strongly contrasts with its moderate structural effect on the duplex detected in our simulations and in previous experimental studies (see above). Clearly, this behavior, which is common to other modified nucleobases such as the apolar isosteres of nucleobases,[47] implies that under normal experimental conditions the impact on the structure of very stable duplexes will be negligible while for duplexes of moderate stability the presence of 6-thioguanine will lead to the total unfolding of the structure. Note that, although the global minimum molecular structure of the duplex is not altered significantly, the free energy change might well facilitate the base pair opening (or breathing).

MD/TI simulations (Table 2) suggest that the triplex is strongly destabilized (2.4 kcal/mol) with respect to the single-stranded structure when ***S*** is inserted into the Watson−Crick position of a triplex. This impact in stability is similar (difference of only 0.3 kcal/mol in free energy) to that of inserting ***S*** into a duplex, suggesting that the duplex → triplex transition is mostly unaltered by the insertion of 6-thioguanine into the parent duplex. This is a quite unexpected result, since the presence of ***S*** in the WC position should disrupt both Watson−Crick G·C and reverse-Hoogsteen G#G hydrogen bonds (as the energy analysis in Table 3 confirms), leading then to an extra destabilization of the triplex compared with the duplex. However, as noted in Table 3, the loss of stability in the reverse-Hoogsteen hydrogen bond is largely compensated for by the gain in stacking with the central 6-thioguanine. The final result is a net ***S***-induced destabilization for the triplex of 2.7 kcal/mol, not far from the destabilization of 1.5 kcal/mol found for the duplex (see the energy values in Table 3 and free energy values in Table 2).

The G → ***S*** substitution has a moderate impact on stability when taking place in the reverse-Hoogsteen position of the triplex with respect to the single-stranded oligonucleotide ($\Delta\Delta G = 0.7$ kcal/mol; from Table 2). To verify if this behavior was dependent on ionic strength, we equilibrated the structure of triplexes (normal and ***S***-modified) for 1 ns in 1 M NaCl. The results (see Table 2) are almost identical to those obtained using a minimum amount of Na^+ counterions to obtain electroneutrality, proving that the impact of an ionic atmosphere is not crucial in our MD/TI simulations. Overall, MD/TI calculations strongly suggest that the formation of antiparallel triplexes using a TFO containing 6-thioguanine is not much more difficult ($\Delta\Delta G = 0.4$ kcal/mol) compared with utilization of an unmodified TFO. This quantitative finding agrees very well with experimental qualitative estimates obtained by Maher's group,[21] who reported no significant changes in the duplex + TFO → triplex transition, in normal and ***S***-containing TFOs. Energy analysis (Table 3) shows that the introduction of 6-thioguanine in the reverse-Hoogsteen position does not alter the hydrogen bond stability and reduces only slightly the stacking energy (around 1 kcal/mol), mostly in the reverse-Hoogsteen intramolecular stacking.

The impact of a single ***S*** on the stability of a parallel tetraplex is dramatic, as noted in the difference of 4.4 kcal/mol between the quadruplex and single-strand free energies (see Table 2). To verify if this value was dependent on the nature of the counterion placed in the central hole of the tetraplex, we repeated the calculations for equilibrated (1 ns) quadruplexes (normal and ***S***-modified) containing K^+ in the central channel (see Table 1). Results are not dramatically different ($\Delta\Delta G = 3.9$ kcal/mol; see Table 2). Furthermore, to gain extra confidence in our estimates, we repeated the simulation, but now from 1 ns equilibrated quadruplexes (normal and ***S***-modified) with an antiparallel arrangement of the G-strands (and K^+ as the counterion). Once again (see Table 2), the results are not altered ($\Delta\Delta G = 3.9$ kcal/mol), which convinced us that, independently of the strand topology or counterion, the presence of 6-thioguanine very strongly destabilizes the tetraplex with respect to the single strand. Interestingly, the destabilizing effect of

(46) Schneider, B.; Berman, H. M. *Biophys. J.* **1995**, *69*, 2661−2669.
(47) Kool, E. T.; Morales, J. C.; Guckian, K. M. *Angew. Chem., Int. Ed.* **2000**, *39*, 990−1009.

Table 3. Nucleobase−Nucleobase and Nucleobase−Ion Interaction Energies and Standard Deviations (Total, H-Bond, and Stacking) for the Central Three Base Pairs of G-, ***S***-, and 4***S***-Duplexes, for the Central Three Triads of G-, ***S***-WC-, and ***S***-rH-Triplexes, and for the Central Three Tetrads[a] of Parallel G- and 1***S***-Tetraplexes with Na^+ as the Counterion[b]

	G-duplex	***S***-duplex	4***S***-duplex	G-triplex	***S***-WC-triplex	***S***-rH-triplex	G-quadruplex	1***S***-quadruplex
H-bond	−81 ± 2.8	−77 ± 2.8	−70 ± 6.3	−131.9 ± 4.9 (−80.3, −51.6)	−125.3 ± 4.4 (−77.1, −48.2)	−131.5 ± 4.1 (−81.0, −51.4)	−99.7 ± 13.2	−110.0 ± 15.2
stacking	−23 ± 2.8	−24 ± 3.1	−27 ± 3.7	−48.8 ± 5.5	−52.7 ± 3.7	−47.8 ± 5.1	−23.0 ± 6.2	−29.6 ± 6.9
nucleobase−ion							−186.9 ± 12.2	−159.9 ± 20.3
total	−103 ± 2.9	−102 ± 3.2	−97 ± 6.3	−180.7 ± 7.4	−178.0 ± 5.7	−179.2 ± 6.6	−309.6 ± 19.0	−299.5 ± 26.3

[a] The total number of quartets is four, so in the case of three quartets one tetrad is outer. [b] See Table 1. Only the ions placed in the central channel around the mutation site are considered. Similar results are obtained considering other topologies of the tetraplexes and other central counterions. All values are in kilocalories per mole. Values in parentheses correspond to Watson−Crick and reverse-Hoogsteen hydrogen bond interactions.

6-thioguanine on tetraplexes seems to be additive, since the presence of a second ***S*** in the tetrad further destabilizes the tetraplex by 3.4 kcal/mol more (Table 2; simulations done for a parallel tetraplex with Na^+ as the counterion). MD/TI calculations of the free energy change induced by the insertion of additional 6-thioguanines are precluded by the fact that perturbations are not reversible due to the spontaneous unfolding of the tetraplex induced by the presence of three or more 6-thioguanines in the tetrad.

There are no experimental quantitative estimates of the destabilization of tetraplexes induced by 6-thioguanine, but all the available experimental data (refs 17 and 21 and references therein) clearly show that this modified purine inhibits tetraplex formation even under Na^+ or K^+ concentrations that lead to very stable (melting temperatures above 60 °C) unmodified tetraplexes. Our theoretical calculations provide then direct quantitative evidence supporting qualitative experimental data.

Energy analysis (Table 3) shows that the strong destabilizing effect of 6-thioguanine is not related to poorer nucleobase−nucleobase interactions, since both hydrogen bond and stacking interactions are improved in tetrads containing 6-thioguanine with respect to the reference ones. On the contrary, the destabilization of tetraplexes induced by 6-thioguanine seems to be related to the conflict between the bulky sulfur atom and the ions placed in the central channel (see Table 3). It is clear that as the number of 6-thioguanines in the tetrad increases the interaction with the central ion will become more difficult, leading to the expulsion of the ion from the central channel, triggering then the unfolding of the tetraplex.[25]

We can suggest that the good antiparallel triplex formation reported when ***S*** was present in the TFO^{20-22} is due to the combination of (i) very moderate ***S***- induced destabilization of the triplex and (ii) very strong destabilization of the intramolecular TFO−tetraplex induced by 6-thioguanine. The effect of ***S*** is then to decrease the apparent dissociation constant by reducing the inhibitory effect of tetraplex competition (see Figure 2 and refs 20−22), leading then to a net increase in the quantity of antiparallel triplex formed.

Conclusions

(i) The presence of the 6-thioguanine in DNA duplexes leads to a clear displacement of the duplex to the canonical B-form-like geometry, especially for G-tract sequences that are known to adopt B/A intermediate structure. No visible perturbation of the duplex structure is noticed; however, the 6-thioguanine alters the major groove hydration pattern. Further, reduction of the base pairing stabilization energies leads to a sizable loss of stability in the duplex.

(ii) The insertion of 6-thioguanine into an antiparallel triplex is well tolerated from structural and energetic points of view. Thus, when it is incorporated into the Watson−Crick strand, the structure is not destabilized preferentially from duplexes. Similarly, the presence of 6-thioguanine in the reverse-Hoogsteen strand is not disfavored with respect to the single strand. In summary, the free energy associated with the folding process TFO + duplex → triplex is mostly unaltered by the presence of ***S*** in either the duplex or TFO position.

(iii) The presence of 6-thioguanine is very destabilizing in all studied quadruplexes, primarily due to the existence of unfavorable interactions between the bulky sulfur group and the cations in the central channel. The destabilization is not caused by electronic structure effects, i.e., the nucleobase−cation interaction term.

(iv) Combining the small impact of 6-thioguanine on the structure and stability of triplexes with its tremendous impact on tetraplexes strongly supports the use of 6-thioguanine in TFOs designed for antiparallel triplex formation as an excellent approach to reduce the competition between tetraplex and triplex formation. Single oligonucleotides containing 6-thioguanine should be then more effective as triplex-forming oligonucleotides than parent unmodified oligonucleotides.

Acknowledgment. We are indebted to Prof. F. J. Luque for many helpful discussions. This work has been supported by the Centre de Parallelisme de Barcelona (CEPBA), the Centre de Supercomputació de Catalunya (CESCA; Molecular Recognition Project), the European Training Mobility Program (Grant HPRI-CT-1999-00071), the Spanish Ministry of Science and Technology (Grant BIO2003-06848), a Wellcome Trust International Senior Research Fellowship in Biomedical Science in Central Europe (Grant GR067507), MSMT CR (Grant LN00A016), and the Supercomputer Center, Brno. Further, the Institute of Biophysics is supported by Project VZ Z5004920.

Supporting Information Available: Tables listing the force field parameters defining 6-thioguanosine in MD calculations, base stacking energies for different mutual positions characterized by the interbase twist angle, interaction energies for H-bonding in ***S***·C and G·C Watson−Crick base pairs, and averaged values and standard errors of selected helical parameters of duplexes calculated for inner G- or ***S***-tract and adjacent T·A and G·C base pairs using the 3DNA program. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0468628