J|A|C|S
ARTICLES
Published on Web 11/27/2003

Antiparallel Triple Helices. Structural Characteristics and Stabilization by 8-Amino Derivatives

Anna Aviñó,[†] Elena Cubero,[‡] Carlos González,[§] Ramon Eritja,*[,†] and Modesto Orozco*[,‡]

Contribution from the Institut de Biologia Molecular de Barcelona, CSIC, C/Jordi Girona 18-26, E-08034 Barcelona, Institut de Recerca Biomédica de Barcelona, Parc Científic de Barcelona, C/Josep Samitier 1-5, E-08028 Barcelona, Departament de Bioquímica i Biología Molecular, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1, E-08028 Barcelona, and Instituto de Química-Física Rocasolano, CSIC, C/Serrano 119, E-28006 Madrid, Spain

Received March 7, 2003; E-mail: recgma@cid.csic.es; modesto@mmb.pcb.ub.es

Abstract: The structural, dynamical, and recognition properties of antiparallel DNA triplexes formed by the antiparallel d(G#G·C), d(A#A·T), and d(T#A·T) motifs (the pound sign and dot mean reverse-Hoogsteen and Watson−Crick hydrogen bonds, respectively) are studied by means of "state of the art" molecular dynamics simulations. Once the characteristics of the helix are defined, molecular dynamics and thermodynamic integration calculations are used to determine the expected stabilization of the antiparallel triplex caused by the introduction of 8-aminopurines. Finally, oligonucleotides containing 8-aminopurine derivatives are synthesized and tested experimentally using several approaches in a variety of systems. A very large stabilization of the triplex is found experimentally, as predicted by simulations. These results open the possibility for the use of oligonucleotides carrying 8-aminopurines to bind single-stranded nucleic acids by formation of antiparallel triplexes.

Introduction

DNA is a largely polymorphic molecule, which in near-physiological conditions can adopt a variety of structures.[1−3] Triple helices are one of these minor conformations that appear when a DNA duplex containing a polypurine track interacts with a third strand by means of specific H-bonds in the major groove of the duplex. DNA triple helices were theoretically proposed in 1953 by Pauling and Corey,[4] and demonstrated experimentally by Rich and co-workers in 1957.[5] Triplexes have been since then the subject of intense research effort owing not only to their role in the cell cycle but also to their possible biomedical (the antigene strategy) and biotechnological applications.[6−12]

Depending on the orientation of the third strand with respect to the central polypurine Watson−Crick (WC) strand, triplexes are classified into two main categories: (i) parallel and (ii) antiparallel. The former (also named pyrimidine triplexes) are defined by three types of Hoogsteen triads (Figure 1): d(T-A·T), d(C-G·C), and d(G-G·C), where the first base refers to the Hoogsteen strand and the symbols dot and dash refer to Watson−Crick and Hoogsteen pairings, respectively. The antiparallel triplexes (also named purine triplexes) are based on three reverse-Hoogsteen triads (Figure 1): d(G#G·C), d(A#A·T), and d(T#A·T), where the pound sign refers to reverse-Hoogsteen hydrogen bonds.

Most structural studies on DNA triplexes have focused on parallel helices, which, under normal laboratory conditions, are more stable than the corresponding antiparallel conformations.[15−18] Accurate structural models of parallel triplexes have been derived from IR and NMR experiments[19] and molecular dynamics (MD) simulations.[20] This large amount of information about the structure, reactive properties, and flexibility of these triplexes has allowed the design and synthesis of new molecules for the stabilization of the structure in physiological conditions (for a review, see ref 21). Especially powerful are the 8-aminopurine derivatives developed by our groups, which are able to dramatically stabilize parallel triple helices built on the d(T-A·T) or d(C-G·C) triads.[22]

* To whom correspondence should be addressed.
[†] Institut de Biologia Molecular de Barcelona, CSIC.
[‡] Parc Científic de Barcelona and Universitat de Barcelona.
[§] Instituto de Química-Física Rocasolano, CSIC.

(1) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
(2) Bloomfield, V. A., Crothers, D. M., Tinoco, I., Eds. *Nucleic Acids: Structures, Properties and Functions*; University Science Books: Sausalito, CA, 2000.
(3) Blackburn, G. M., Gait, M. J., Eds. *Nucleic Acids in Chemistry and Biology*; IRL Press: Oxford, 1990.
(4) Pauling, L.; Corey, R. B. *Proc. Natl. Acad. Sci. U.S.A.* **1953**, *39*, 84.
(5) Felsenfeld, G.; Davis, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023.
(6) Soyfer, V. N.; Potaman, V. N. *Triple-Helical Nucleic Acids*; Springer-Verlag: New York, 1996.
(7) Malvy, C.; Harel-Bellan, A.; Pritchard, L. L. *Triple helix forming oligonucleotides*; Kluwer Academic: Dordrecht, The Netherlands, 1999.
(8) Fox, K. R. *Curr. Med. Chem.* **2000**, *7*, 17.
(9) Micklefield, J. *Curr. Med. Chem.* **2001**, *8*, 1157.
(10) Sun, J. S.; Garestier, T.; Hélène, C. *Curr. Opin. Struct. Biol.* **1996**, *6*, 327.
(11) Giovannangeli, C.; Hélène, C. *Nat. Biotechnol.* **2000**, *18*, 1245.
(12) Vasquez, K. M.; Narayanan, L.; Glazer, P. M. *Science* **2000**, *290*, 530.
(13) Cooney, M.; Czernuszewicz, G.; Postel, E. H.; Flint, S. J.; Hogan, M. E. *Science* **1988**, *241*, 456.
(14) Lyamichew, V. I.; Frank-Kamenetskii, M. D.; Soyfer, V. *Nature* **1990**, *344*, 568.
(15) Scaria, P. V.; Shafer, R. H. *Biochemistry* **1996**, *35*, 10985.
(16) Chandler, S. P.; Fox, K. R. *Biochemistry* **1996**, *35*, 15038.
(17) Washbrook, E.; Fox, K. R. *Biochem. J.* **1994**, *301*, 569.
(18) Cheng, Y. K.; Pettitt, B. M. *J. Am. Chem. Soc.* **1992**, *114*, 4465.

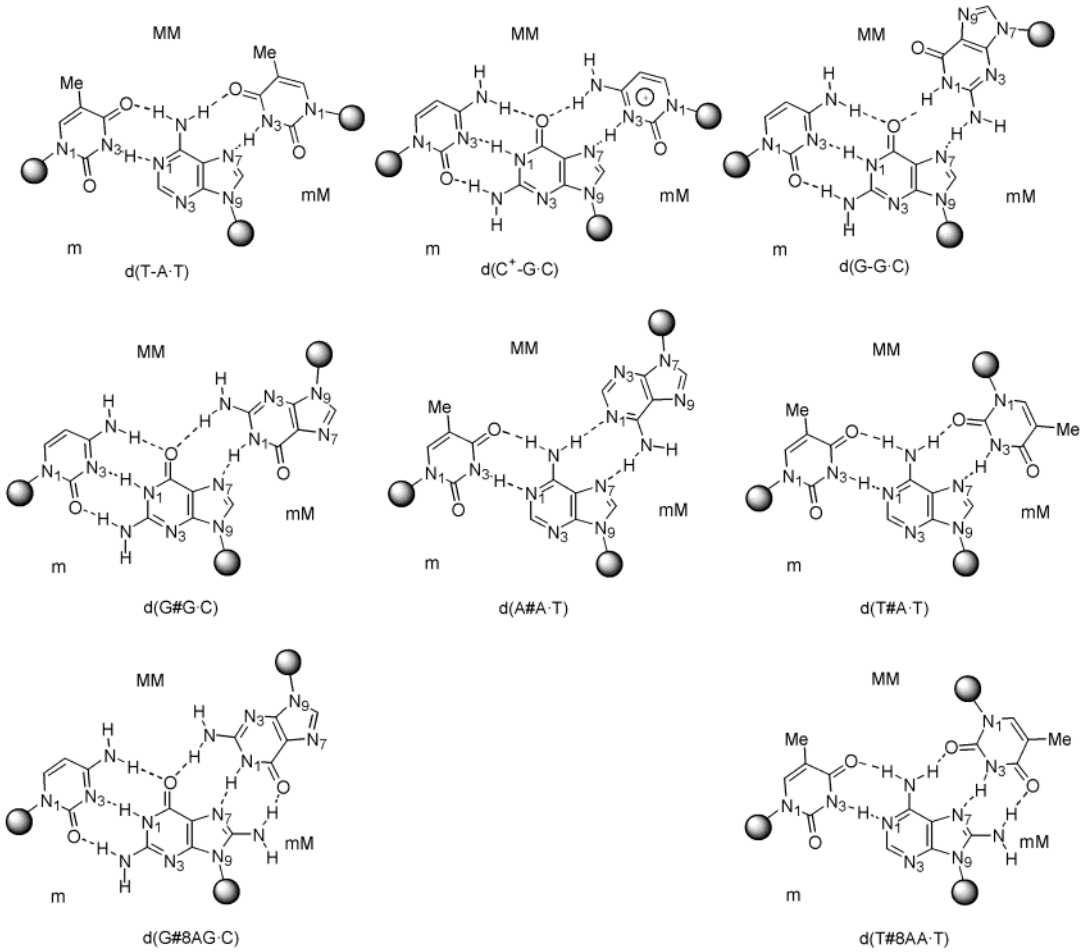


Figure 1. Schematic representation of distinct triads found in triplexes. First row: Hoogsteen pairings found in parallel triplexes. Second row: reverse-Hoogsteen pairings found in antiparallel triplexes. Third row: suggested reverse-Hoogsteen pairings involving 8-aminoadenine and 8-aminoguanine. The position of the minor (m), minor-major (mM), and major-major (MM) grooves is displayed for each triad.

Parallel triplexes are typically formed by binding a single-stranded oligonucleotide to a Watson−Crick duplex. However, they can be formed also by targeting a single-stranded DNA or RNA.[23] Thus, it has been demonstrated that linking Watson−Crick and Hoogsteen pyrimidine strands[23a,c,g−i] as well as Hoogsteen pyrimidine strands[23c,e] yields hairpins or clamps which bind their target single-stranded DNA with enhanced affinity for triplex formation. This effect is more pronounced using circular oligonucleotides,[23d] or when the Watson−Crick strand contains 8-aminopurine derivatives.[23g,h,i] Kool and co-workers[23e,j] have proved that the same strategy works to generate stable antiparallel triplexes using as templates purine-rich hairpins. In addition to the excellent binding properties of hairpins or clamps as templates for triplex formation, they present a priori several potential advantages which include better nuclease resistance and unlimited possibilities for functionalization in the loop region.

There is scarce high-resolution structural information on antiparallel triplexes,[21,23,24] which is probably because of their

(19) (a) Rajagopal, P.; Feigon, J. *Nature* **1989**, *22*, 637. (b) Rajagopal, P.; Feigon, J. *Biochemistry* **1989**, *28*, 7859. (c) Radhakrishnan, I.; Gao, X.; de los Santos, C.; Live, D.; Patel, D. J. *Biochemistry* **1991**, *30*, 9022. (d) Young, S. L.; Krawczyk, S. H.; Matteuci, M. D.; Toole, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10023. (e) Macaya, R. F.; Schultze, P.; Feigon, J. *J. Am. Chem. Soc.* **1992**, *114*, 781. (f) Raghunathan, G.; Miles, H. T.; Sasisekharan, V. *Biochemistry* **1993**, *32*, 455. (g) Radhakrishnan, I.; Patel, D. J. *Structure* **1994**, *2*, 395. (h) Radhakrishnan, I.; Patel, D. J. *J. Mol. Biol.* **1994**, *241*, 600. (i) Schultze, P.; Koshlap, K. M.; Feigon, J. *Biochemisty* **1997**, *36*, 2659. (j) Koshlap, K. M.; Schultze, P.; Brunar, H.; Dervan, P. D.; Feigon, J. *Biochemistry* **1997** *26*, 2659. (k) Bartley, J. P.; Brown, T.; Lane, A. N. *Biochemistry* **1997**, *36*, 14502. (l) Gotfredsen, C. H.; Schultze, P.; Feigon, J. *J. Am. Chem. Soc.* **1998**, *120*, 4281. (m) Asensio, J. L.; Brown, T.; Lane, A. N. *Nucleic Acids Res.* **1998**, *26*, 3677. (n) Tarkoy, M.; Phipps, A. K.; Schultze, P.; Feigon, J. *Biochemistry* **1998**, *37*, 5810. (n) Phipps, A. K.; Tarkoy, M.; Schultze, P.; Feigon, J. *Biochemistry* **1998**, *37*, 5820. (o) Asensio, J. L.; Dosanjh, H. S.; Jenkins, T. C.; Lane, A. N. *Biochemistry* **1999**, *37*, 1588. (p) Lane, A. N.; Jenkins, T. C. *Curr. Org. Chem.* **2001**, *5*, 845.

(20) (a) Shields, G. C.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc.* **1997**, *119*, 7463. (b) Soliva, R.; Laughton, C. A.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1998**, *120*, 1126. (c) Soliva, R.; Luque, F. J.; Orozco, M. *Nucleic Acids Res.* **1999**, *27*, 2248. (d) Orozco, M.; Pérez, A.; Noy, A.; Luque, F. J. *Chem. Soc. Rev.*, in press.

(21) Robles, J.; Grandas, A.; Pedroso, E.; Luque, F. J.; Eritja, R.; Orozco, M. *Curr. Org. Chem.* **1998**, *6*, 1333−1368.

(22) (a) Güimil-Garcia, R.; Ferrer, E.; Macías, M. J.; Eritja, R.; Orozco, M. *Nucleic Acids Res.* **1999**, *27*, 1991. (b) Cubero, E.; Güimil-Garcia, R.; Luque, F. J.; Eritja, R.; Orozco, M. *Nucleic Acids Res.* **2001**, *29*, 2522. (c) Soliva, R.; Güimil-Garcia, R.; Blas, J. R.; Eritja, R.; Asensio, J. L.; González, C.; Luque, F. J.; Orozco, M. *Nucleic Acids Res.* **2000**, *28*, 4531. (d) Güimil-García, R.; Bachi, A.; Eritja, R.; Luque, F. J.; Orozco, M. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 3011.

(23) (a) Giovannangeli, C.; Thuong, N. T.; Hélène, C. *J. Am. Chem. Soc.* **1991**, *113*, 7775. (b) Wang, S.; Kool, E. T. *J. Am. Chem. Soc.* **1994**, *116*, 8857. (c) Kandimalla, E. R.; Agrawal, S.; Venkataraman, G.; Sasisekharan, V. *J. Am. Chem. Soc.* **1995**, *117*, 6416. (d) Wang, S.; Kool, E. T. *Nucleic Acids Res.* **1995**, *23*, 1157. (e) Vo, T.; Wang, S.; Kool, E. T. *Nucleic Acids Res.* **1995**, *23*, 2937. (f) Kandimalla, E. R.; Agrawal, S. *Biochemistry* **1996**, *35*, 15332. (g) Aviñó, A.; Morales, J. C.; Frieden, M.; de la Torre, B. G.; Güimil Garcia, R.; Cubero, E.; Luque, F. J.; Orozco, M.; Azorín, F.; Eritja, R. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1761. (h) Cubero, E.; Aviñó, A.; de la Torre, B. G.; Frieden, M.; Eritja, R.; Luque, F. J.; González, C.; Orozco, M. *J. Am. Chem. Soc.* **2002**, *124*, 3133. (i) Aviñó, A.; Frieden, M.; Morales, J. C.; de la Torre, B. G.; Güimil García, R.; Azorín, F.; Gelpí, J. L.; Orozco, M.; González, C.; Eritja, R. *Nucleic Acids Res.* **2002**, *30*, 2609. (j) Kool, E. T. *Chem. Rev.* **1997**, *97*, 1473.

Table 1. Summary of Calculations Done with Antiparallel Triplexes Containing Normal Nucleobases and Their 8-Amino Derivatives[a]

oligo	name	simulation length (ns)	oligo	name	simulation length (ns)
d(GGGGGGGGGG) d(CCCCCCCCCC) d(GGGGGGGGGG)	T1	5	*d(GGGGGGGGGG)* d(CCCCCCCCCC) d(GGGGG^{N}GGGGG)	T1n	2
d(GGGGG^{N}GGGGG) d(CCCCCCCCCC) d(GGGGG^{N}GGGGG)	T1nn	2	d(CCTCCCTCTC) d(GGAGG^{N}GAGAG)	T1d[b]	2
d(AAAAAAAAAA) d(TTTTTTTTTT) d(AAAAAAAAAA)	T2	3	*d(AAAAAAAA)* d(TTTTTTTT) d(AAAAAAAA)	T5	2
d(TTTTTTTTTT) d(TTTTTTTTTT) d(AAAAAAAAAA)	T3	3	*d(TTTTTTTTT)* d(TTTTTTTTT) d(AAAAAAAAA)	T6	2
d(TTTTTTTTT) d(TTTTTTTTT) d(AAAAA^{N}AAAA)	T6n	2	d(TTTTTTTTT) d(AAAAA^{N}AAAA)	T6d	2
d(GTGTTTGTTG) d(CTCTTTCTTC) d(GAGAAAGAAG)	T4	3	*d(GTGTTTGTTG)* d(CTCTTTCTTC) d(GAGAA^{N}AGAAG)	T4n	2
d(CTCTTTCTTC) d(GAGAA^{N}AGAAG)	T4d	2			
d(TGGTGGT) d(TCCTCCT) d(AGGAGGA)	T7a	2	*d(TGGTGGT)* d(TCCTCCT) d(AGGAGGA)	T7b	2

[a] Simulations with duplexes containing the 8-amino derivatives are also displayed for completeness. The sequence of the third strand is shown in italics. [b] Note that the sequence of the duplex matches always that of the parent triplex except for this case, where some adenines replaced guanines to avoid an A-philic duplex.

reduced stability,[15−18] making the structures stable only in the presence of a high concentration of divalent ions.[19a,24−26] However, despite their reduced stability under laboratory conditions, antiparallel triplexes are more promising than parallel structures in the biomedical field, since the formation of the former is pH independent, while a parallel triplex requires in most cases an acidic pH, which is not always present inside the cell.[27] Accordingly, there is a clear need for more structural information on antiparallel triplexes, since this structural knowledge would help in the design of new strategies for the stabilization of this important family of triple helices.

Here we present a wide theoretical study of antiparallel triplexes using state of the art MD simulations. Structures built on the d(G#G·C), d(A#A·T), and d(T#A·T) triads were analyzed using nanosecond time scale simulations in aqueous solvent. The equilibrated structures were used to examine the impact of 8-aminopurine substitutions on the stability of distinct antiparallel triplexes by means of a combination of MD and thermodynamic integration (TI) simulations. Moreover, the 8-aminopurine derivatives were synthesized and for the first time incorporated into antiparallel triplexes. Experimental analysis allowed us to confirm experimentally that 8-aminopurine derivatives lead to extremely stable antiparallel triplexes, even in conditions that mimic physiological ones. The impact of this discovery on the design of new antigene and antisense strategies is discussed.

Methods

Molecular Dynamics Simulations. MD simulations were performed to analyze the structural, dynamic, and recognition properties of antiparallel duplexes that contain all types of antiparallel triads (Figure 1). For this purpose, starting models of the following triplexes were generated (Table 1): (i) a 10-mer poly[d(G#G·C)] triplex (named T1), (ii) a 10-mer poly[d(A#A·T)] triplex (T2), (iii) a 10-mer poly(T#A·T) triplex (T3), and (iv) a 10-mer triplex containing d(T#A·T) and d(G#G·C) triads with the reverse-Hoogsteen pairing (T4 in Table 1). To analyze the impact of the oligonucleotide size on the structure of the triplexes, additional simulations were performed for (v) an 8-mer d(A#A·T) triplex (T5), (vi) a 9-mer poly[d(T#A·T)] triplex (T6), and (vi) two 7-mer triplexes containing d(T#T·A) and d(G#G·C) triads with the reverse-Hoogsteen pairing (T7a and T7b simulations in Table 1).

Starting conformations for triplexes T1−T7a were generated from Patel's structure[24] of the triplex d(AGGAGGA), which contains d(A#A·T) and d(G#G·C) triads (PDB entry pdb134d). Sequences were modified when needed, and triplexes longer than sseven triads were extended using average helical parameters. For comparison, one simulation (T7b) was repeated using as starting conformation another triplex structure also deposited by Patel's group in PDB as entry pdb135d (the root mean square deviation (RMSD) between the two NMR structures for all atoms is 0.9 Å (for the duplex portion) and 1.2 Å (for the entire molecule)). This new trajectory was very similar to that found for T7a, demonstrating that MD simulations are not highly dependent on small structural changes in the starting configuration. All the structures were partially minimized to avoid incorrect geometries for 1000 steepest descent and 1000 conjugate gradient cycles. These relaxed systems were then surrounded by water (between 2200 and 2800 TIP3P[28] molecules) and Na^+ to achieve neutrality. They were then optimized, thermalized, and equilibrated using our standard multistage protocol, which extends for 200 ps.[20a,b] Equilibrated systems were then subject to production runs of 5 ns (T1), 3 ns (T2−T4), and 2 ns (T5−T7) of unrestrained MD simulation.

To analyze the structural impact of the introduction of 8-aminopurines, we defined additional triplexes that contained these derivatives in at least one position of the triplex (Table 1). Structures were created from the corresponding reference triplexes (see above), and were then hydrated, optimized, and equilibrated using a protocol identical to that described above. Simulations of triplexes containing 8-aminopurines were extended for 2 ns in all the cases. Finally, when needed for TI simulations (see above), Watson−Crick duplexes were generated using

(24) Radhakrishnan, I.; Patel, D. J. *Structure* **1993**, *1* 135.
(25) Ji, J.; Hogan, M. E.; Gao, X. *Structure* **1996**, *4*, 425.
(26) Bernués, J.; Beltran, R.; Casasnovas, J. M.; Azorin, F. *EMBO J.* **1989**, *8*, 2087.
(27) Washbrook, E.; Fox, K. R. *Biochem. J.* **1994**, *301*, 569.
(28) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

the standard fiber parameter,[29] hydrated, neutralized, optimized, heated, and equilibrated as described above. Unrestrained simulations for duplexes (both modified and unmodified) extended for 2 ns.

The MD simulations were carried out in the isothermic−isobaric ensemble using periodic boundary conditions and the particle mesh-Ewald technique to represent long-range electrostatic effects.[30] AMBER-99[31,32] and TIP3P force fields[28] were used to represent molecular interactions in the system. Parameters for 8-aminopurines were taken from our previous parametrization studies.[23h,33] SHAKE[34] was used to constrain all the bonds at their equilibrium distances, which allowed us to use a 2 fs time scale for integration of Newton's laws of motion. The AMBER-6.0 computer program was used for the MD simulations.[35] The simulations reported here correspond to up to 30 ns of unrestrained MD simulation of antiparallel triplexes (22 ns for unmodified triplexes and 8 ns for triplexes containing 8-aminopurine derivatives).

Energetic analysis of the trajectories was carried out using analysis modules in AMBER-6.0, as well as "in house" programs. Interaction maps were determined using our cMIP methodology.[20a,b,36] Accordingly MD-averaged structures were used to compute the interaction energy of a given triplex with a probe molecule (O^+ in this paper) located in a regular grid around the molecule. Interaction energies were determined as the addition of a van der Waals term to an electrostatic component determined using Poisson−Boltzmann molecular electrostatic potentials (inner dielectric 2 and outer dielectric 80) computed using AMBER force field parameters. Hydration maps were obtained by integrating solvent population during the dynamics using in house programs.[20a,b]

Free Energy Calculations. MD/TI calculations were performed to determine the gain in stability of antiparallel triplexes obtained by the substitution of purines by 8-aminopurines. For this purpose, and following standard thermodynamic cycles, mutations between 8-aminopurines and purines were done in distinct triplexes and duplexes (see the Results and Discussion). Mutations between oligonucleotides containing 8-aminoguanine (G^N) and 8-aminoadenine (A^N) and those containing parent nucleobases were performed using 21 or 41 windows of 10 ps each (independent free energy estimates were taken from the first and second halves of each window), leading to simulations of 420 and 840 ps. To gain statistical confidence in the results, mutations were carried out in the 8-aminopurine → purine and purine → 8-aminopurine directions. Thus, each free energy difference value was obtained by averaging eight independent estimates of the same process. In our experience this procedure leads to better converged results than those obtained by single, but more extended trajectories. Structures for the oligonucleotides containing 8-aminopurines were built from MD-averaged structures of the corresponding unmodified oligonucleotides and equilibrated for 2 ns of unrestrained MD. In all the cases the mutations were done in positions located in the center of the helix.

The impact of the 8-aminopurine substitution on the triplex stability was determined as the difference between the free energies associated with the mutations in the triplex and in the duplex (see eq 1). That is, MD/TI calculations provided a direct estimate of the change in the free energy of the triplex → duplex transition associated with the substitution of a purine (in the Watson−Crick position) for an 8-aminopurine.

$$\Delta\Delta G(Y \rightarrow Y^N) = \Delta G(Y \rightarrow Y^N)_{triplex} - \Delta G(Y \rightarrow Y^N)_{duplex} \quad (1)$$

Oligonucleotide Synthesis. Oligonucleotides were prepared on an automatic Applied Biosystems 392 DNA synthesizer. The phosphoramidites of 8-aminoadenine, 8-aminoguanine, and 8-aminohypoxanthine were prepared as described elsewhere.[22a−c,37−39] The phosphoramidite of protected 8-amino-2′-deoxyinosine was dissolved in dry dichloromethane to yield a 0.1 M solution. The remaining phosphoramidites were dissolved in dry acetonitrile (0.1 M solution). Oligonucleotides containing natural bases were prepared using commercially available chemicals and following standard protocols. After the assembly of the sequences, oligonucleotide supports were treated with 32% aqueous ammonia at 55 °C for 16 h except for oligonucleotides bearing 8-aminoguanine. In this case, a 0.1 M 2-mercaptoethanol solution in 32% aqueous ammonia was used, and the treatment was extended to 24 h at 55 °C.[8] Ammonia solutions were concentrated to dryness, and the products were purified by reversed-phase HPLC. Oligonucleotides were synthesized on a 0.2 μmol scale and with the last DMT group at the 5′ end (DMT on protocol) to facilitate reversed-phase purification. All purified products presented a major peak, which was collected. Yields (OD units at 260 nm after HPLC purification, 0.2 μmol) were between 5 and 10 OD. HPLC solutions were as follows: solvent A, 5% ACN in 100 mM triethylammonium acetate, pH 6.5; solvent B, 70% ACN in 100 mM triethylammonium acetate, pH 6.5. Other HPLC conditions were as follows: column, PRP-1 (Hamilton), 250 × 10 mm; flow rate, 3 mL/min; a 30 min linear gradient from 10% to 80% B (DMT on) or a 30 min linear gradient from 0% to 50% B (DMT off).

Helix−Coil Transitions and Thermodynamic Analysis. Melting experiments with duplex d(C3T4C3)·d($GG^NG^NA4G^NG^NG$) and triplex d(C3T4C3)·2[d($GG^NG^NA4G^NG^NG$)] were performed as described by Pilch et al.[40]

Melting experiments with triplexes were performed as follows. Solutions of equimolar amounts of hairpins and the target Watson−Crick pyrimidine strand (11-mer) were mixed in 10 mM sodium cacodylate, 50 mM $MgCl_2$, and 0.1 mM EDTA at pH 7.2. The DNA concentration was determined by UV absorbance measurements (260 nm) at 90 °C, using the following extinction coefficients for the DNA coil state: 7500, 8500, 12500, 12500, 15000, and 15000 M^{-1} cm^{-1} for C, T, G, 8-amino-G, A, and 8-amino-A, respectively. The solutions were heated to 90 °C, allowed to cool slowly to room temperature, and stored at 4 °C until UV analysis. UV absorption spectra and melting experiments (absorbance vs temperature) were recorded in 1 cm path length cells using a spectrophotometer, with a temperature controller and a programmed temperature increase rate of 0.5 °C/min. Melts were run on triplex concentrations of 3 μM (1−1.2 OD units at 260 nm).

The samples used for the thermodynamic studies were prepared in a similar way, but melting experiments were recorded at 260 nm and using 0.1, 0.5, and 1 cm path length cells.

Thermodynamic data were analyzed as described elsewhere[23i,40−42] Melting curves were obtained at concentrations ranging from 0.5 to 20 μM triplex. The melting temperatures (T_m) were measured at the maximum of the first derivative of the melting curve. The plot of $1/T_m$ vs ln C was linear. Linear regression of the data gave the slope and the y-intercept, from which ΔH and ΔS were obtained. The free energy value was obtained from the standard equation $\Delta G = \Delta H - T\Delta S$.

(29) Arnott, S.; Bond, P. J.; Selsing, E.; Smith, P. J. C. *Nucleic Acids Res.* **1976**, *133*, 1405.
(30) Darden, T. A.; York, D. M.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *98*, 10089.
(31) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Fergurson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(32) Cheatham, T. E.; Cieplak, P.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845.
(33) Cubero, E.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2001**, *123*, 12018.
(34) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977** *23*, 327.
(35) 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. *AMBER-6.0*; University of California: San Francisco, 1999.
(36) Gelpí, J. L.; Kalko, S.; de la Cruz, X.; Barril, X.; Cireram, J.; Luque, F. J.; Orozco, M. *Proteins* **2001**, *45*, 428.
(37) Hawai, K.; Saito, I.; Sugiyama, H. *Tetrahedron Lett.* **1998**, *39*, 5221.
(38) Rao, T. S.; Durland, R. H.; Revankar, G. R. *J. Heterocycl. Chem.* **1994**, *41*, 935.
(39) Rieger, R. A.; Iden, C. R.; Gonikberg, E.; Jonson, F. *Nucleosides Nucleotides* **1999**, *18*, 73.
(40) Pilch, D. S.; Levenson, C.; Shafer, R. H. *Biochemistry* **1991**, *30*, 6081.
(41) Xodo, L. E.; Manzini, G.; Quadrifolio, F.; van der marel, G. A.; van Boom, J. H. *Nucleic Acids Res.* **1991**, *19*, 5625.
(42) Manzini, G.; Xodo, L. E.; Gasparotto, D. *J. Mol. Biol.* **1990**, *213*, 833.

Table 2. RMSDs (Å) between the Triplexes Studied Here and Several Reference Structures[a]

triplex	MD-avd[b]	pdb134d[c]	pdb135d[c]
T1	1.7(0.5)	2.1(0.3)	2.1(0.4)
T1n	1.0(0.4)		
T1nn	1.1(0.3)		
T2	1.2(0.3)	2.1(0.2)	2.0(0.2)
T3	1.1(0.3)	2.0(0.1)	1.7(0.1)
T4	1.5(0.4)	2.3(0.2)	2.1(0.2)
T4n	1.3(0.3)		
T5	1.2(0.3)	2.0(0.2)	1.9(0.2)
T6	1.4(0.3)	2.2(0.2)	2.0(0.2)
T6n	1.4(0.3)		
T7a	0.9(0.2)	2.1(0.2)	2.2(0.2)
T7b	1.1(0.2)	2.1(0.2)	2.1(0.2)

[a] Standard deviations (Å) are shown in parentheses. [b] Values computed with respect to the respective MD-averaged conformation using all the atoms in the corresponding oligonucleotide. [c] Values computed using only backbone atoms (including C1′) for the common 7-mer sequence.

NMR Experiments. NMR experiments were performed to gain additional evidence on the formation of antiparallel triplexes. For this purpose, equimolar mixtures of the hairpins H26GT and H26GT($2G^N$) with the WC-11-mer were prepared in 200 μL of either D_2O or 9:1 H_2O/D_2O. The resulting solutions were buffered (25 mM sodium phosphate, 50 mM $MgCl_2$, pH 7) and had a triplex concentration of about 1 mM. NMR spectra were acquired in a Bruker DMX spectrometer operating at 600 MHz and processed with X-WINNMR software. TOCSY[43] and NOESY[44] spectra were acquired in D_2O and in H_2O/D_2O (9:1). All 2D experiments were carried out at 5 °C. NOESY spectra were acquired in 100 ms in H_2O and 200 ms in D_2O. In H_2O experiments, water suppression was achieved by employing jump-and-return pulse sequences.[45]

Results and Discussion

Structural Description of the Antiparallel Triplex Helix. Equilibrium trajectories of the triplexes were well equilibrated, indicating that a reasonable description of the structure can be obtained by nanosecond time scale MD simulations (see Table 2 and the Supporting Information) performed under low ionic strength conditions. The RMSDs with respect to the respective MD-averaged structures were very small (around 1 Å) in all cases, indicating that the trajectories were visiting a quite well-defined region of the configurational space, and that the flexibility of the triplex was quite reduced compared with that of duplexes with the same (Watson−Crick) sequence. The RMSDs with respect to the NMR structures in pdb134d and pdb135d (computed using all backbone atoms of the central 7-mer sequence) were also quite small, around 2 Å. For the T7a and T7b triplexes (those corresponding to the same oligonucleotides studied by Patel's group) the RMSDs between theoretical and NMR structures were around 2.2 Å (only backbone atoms) and 1.8 Å (all atoms), confirming that the MD trajectories were sampling regions of the configurational space similar to those detected in the NMR experiments.[24] Given the ability of MD simulations to escape from incorrect triplex conformations,[20a,46] the agreement between MD and NMR results cannot be considered fortuitous or related to incomplete sampling of the MD trajectory.

Analysis of the trajectories shows the independence of the results on the length of the oligonucleotide (in the range 7−10-mer) on the starting conformation for the simulation (pdb134d or pdb135d) and on the length of the trajectory (in the 1−5 ns range; see Table 2). The introduction of one or two 8-aminopurines led to negligible changes in the structures, as found previously for parallel triplexes.[22a−c] In all the cases studied here the RMSDs between the oligonucleotides containing 8-aminopurines and the parent oligonucleotides were below 1 Å (i.e., very close to the thermal noise of the simulation). In summary, the main characteristics of the antiparallel triplex were well defined on the nanosecond time scale, and are independent of the definition of the simulation model and of the presence of small chemical alterations in the structure. We can then safely qualitatively discuss the characteristics of the antiparallel triplexes containing the three reverse-Hoogsteen triads (see the Methods), d(T#A·T), d(G#G·C), and d(A#A·T), by analyzing only simulations for the larger unmodified oligonucleotides T1−T4. Caution is however necessary before a quantitative explanation since we cannot forget that individual MD simulations are performed on a short time scale and under conditions which are not identical to those considered experimentally, where for instance a nonnegligible amount of Mg^{2+} (not considered in our simulations) is added to gain stability in the structure.

Despite some distortions in the third strand of sequences containing adjacent d(T#A·T) triads (T3 and T4), the general shapes of the helices are well preserved during all simulations (Figure 2). Watson−Crick hydrogen bonds were present for 98−100% of the time for trajectories T1−T4 (the same values were obtained for T5−T7), but the reverse-Hoogsteen hydrogen bonds were more labile, especially for adjacent d(T#A·T) triads, as shown by the fact that around 27% of the T3 simulation and 32% of the T4 simulation showed disruption of T#A reverse-Hoogsteen hydrogen bonds because of breathing movements. Similar values were obtained for T6, while greater conservation of reverse-Hoogsteen hydrogen bonds (86% and 92%) was found for triplexes T7a and T7b, which contained d(T#A·T) triads, but not in contiguous positions. Analysis of trajectories T3, T4, and T6 shows that the partial disruption of reverse-Hoogsteen hydrogen bonds was due to a clear tendency of the reverse-Hoogsteen thymine in poly[d(T#A·T)] tracks to escape from the planarity of the Watson−Crick d(A·T) pair in a pseudo-propeller-twist movement. The lost or reverse-Hoogsteen hydrogen bonds were less important for triplexes containing only d(G#G·C) and d(A#A·T) triads (85% and 90% of reverse-Hoogsteen hydrogen bonds were present in simulations T1 and T2, respectively). Interestingly, the magnitude of breathing in antiparallel d(T#A·T) triplexes was much larger than that found for the parallel d(T#A·T) structures. This finding suggests that the parallel arrangement is clearly more stable for the d(T#A·T) triads, despite the similar stability of Hoogsteen and reverse-Hoogsteen hydrogen bonds.

The general structure of antiparallel triplexes is surprisingly similar to that of parallel arrangements,[22a−c] as shown by RMSDs in the range 1−2 Å between the Watson−Crick backbones of the MD-averaged parallel and antiparallel triplexes (Table 3). Interestingly, the general characteristics of the antiparallel triplexes were quite independent of the sequence, as noted in RMSDs, which were also in the range 1−2 Å between the Watson−Crick backbones of MD-averaged struc-

(43) Bax, A.; Davies, D. J. *J. Magn. Reson.* **1985**, *65*, 355−360.
(44) Kumar, A.; Ernst, R. R.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1980**, *95*, 1−6.
(45) Plateau, P.; Güeron, M. *J. Am. Chem. Soc.* **1982**, *104*, 7310−7311.
(46) Shields, G.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc.* **1998**, *120*, 5895.

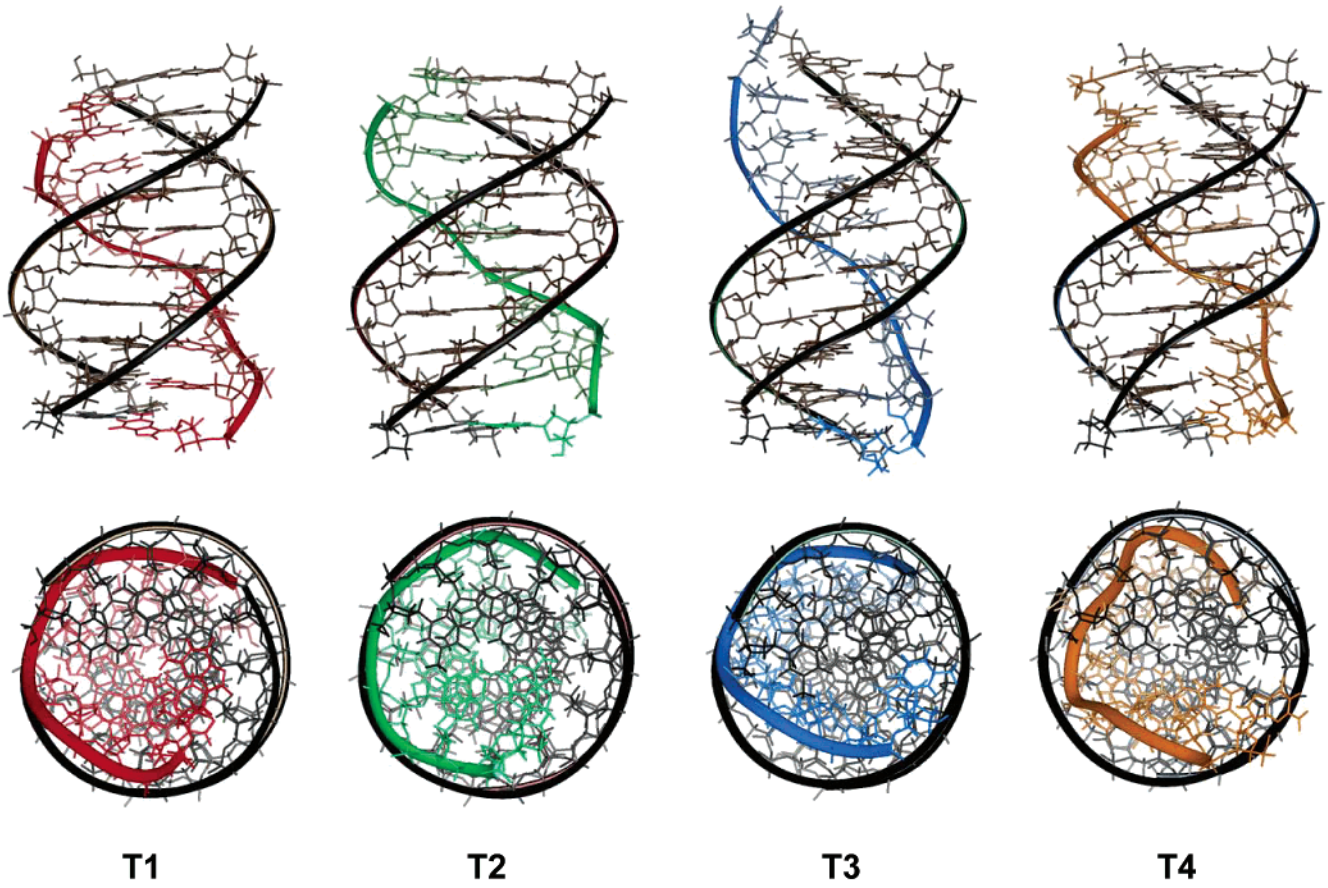


Figure 2. Schematic representation of the MD-averaged structure of the 10-mer antiparallel triplexes (T1−T4 from left to right) studied here. The third strand is shown in color (T1, red; T2, green; T3, blue; T4, orange), and the Watson−Crick duplex is in black.

Table 3. Backbone RMSDs (Å) between Different MD-Averaged Structures of Antiparallel Triplexes (T1, T2, T3, and T4) and the MD-Averaged Structure of the Poly[d(T-A·T)] Parallel Triplex in Reference 22[a]

	parallel	T1	T2	T3
T1	1.7			
T2	1.3	0.8/*2.2*		
T3	2.3	2.2/*3.0*	2.0/*2.8*	
T4	1.5	1.3/*2.3*	1.0/*2.1*	1.3/*2.2*

[a] Values in Roman font were obtained by fitting only the Watson−Crick strands, and values in italics were derived by fitting the three backbones.

Table 4. Helical Parameters (Watson−Crick Strands) of Antiparallel Triplexes T1−T4[a]

	T1	T2	T3	T4	NMR
x-displacement	−2.7(0.7)	−4.9(1)	−3.2(2)	−4.8(2)	−2.1/*−1.9*
inclination	−3.3(6)	10.9(7)	−0.1(5)	11.6(14)	−0.4/*−4.8*
rise	3.4(0.1)	3.4(0.2)	3.4(0.1)	3.4(0.1)	3.6/*3.7*
roll	2.9(2)	0.8(2)	−3.9(1)	−0.5(2)	1.8/*−0.6*
twist	30.1(1)	30.3(0.9)	32.1(0.7)	30.5(1)	30.0/*30.3*
propeller twist	−2.1(5)	−3.9(5)	−8.2(4)	−6.0(5)	−11/*−13*
phase	131(29)	128(32)	134(23)	131(27)	122/*121*[b]
m groove	12.6(0.5)	12.3(0.5)	10.5(0.4)	11.4(0.5)	12.7/*12.2*[b]
mM groove	10.8(0.6)	10.4(0.6)	11.7(0.4)	11.7(0.4)	9.0/*8.8*[b]
MM groove	12.8(2.5)	12.9(0.9)	12.9(0.7)	12.5(1)	14.5/*14.7*[b]

[a] Angular values are in degrees, and displacement values are in angstroms. Only the central 8-mer Watson−Crick strand of the triplex is considered for the analysis. Average NMR data are displayed for comparison (Roman data from pdb134d and dat in italics from pdb135d). [b] NMR data are obtained for a 7-mer oligonucleotide, which implies large uncertainties in the definition of the grooves.

tures of simulations T1, T2, T3, and T4 (Table 3). However, the introduction of the third strand in the calculation of the RMSD led to a dramatic increase of c.a. 1 Å, confirming[24] that the largest sequence-dependent changes are always located in the third strand, while the core of the triplex (defined by the Watson−Crick strands) is less sensitive to sequence effects. Cross-RMSD relationships suggest that triplexes containing only d(G#G·C) and/or d(A#A·T) triads (T1, T2) are very similar, and both differ slightly from triplexes containing d(T#A·T) triads (T3, T4), which are more distorted from an ideal helical conformation (Table 3). The reduced stability of the d(T#A·T) reverse-Hoogsteen triads can explain the differential structural properties of triplexes containing this type of triad. Additional distortions in the third strand in T4 (Figure 2) were probably due to the different sizes of d(T#A·T) and d(G#G·C) triads.

Triplexes T1−T4 showed quite standard helical values, with an average twist around 30° (which increases to 32° for T3), a rise around 3.4 Å, a small inclination, a roll and propeller twist, and small (in absolute terms) and negative *x*-displacement values (Table 4). On average, the sugars are in the south to southeast regions (Table 4 and Figure 3A) for all the triplexes, as expected for structures pertaining to the B-family. However, large differences were observed in the puckering population between Watson−Crick and reverse-Hoogsteen strands (Figure 3A). Thus, the vast majority of sugars in the Watson−Crick strands had phase angles in the range 90−180°. On the contrary, large populations of north puckerings were detected in the reverse-Hoogsteen strand, and for one of the triplexes (T3) north puckerings were more populated than south puckerings. In summary, as found in previous studies with parallel triplexes,[20a,b]

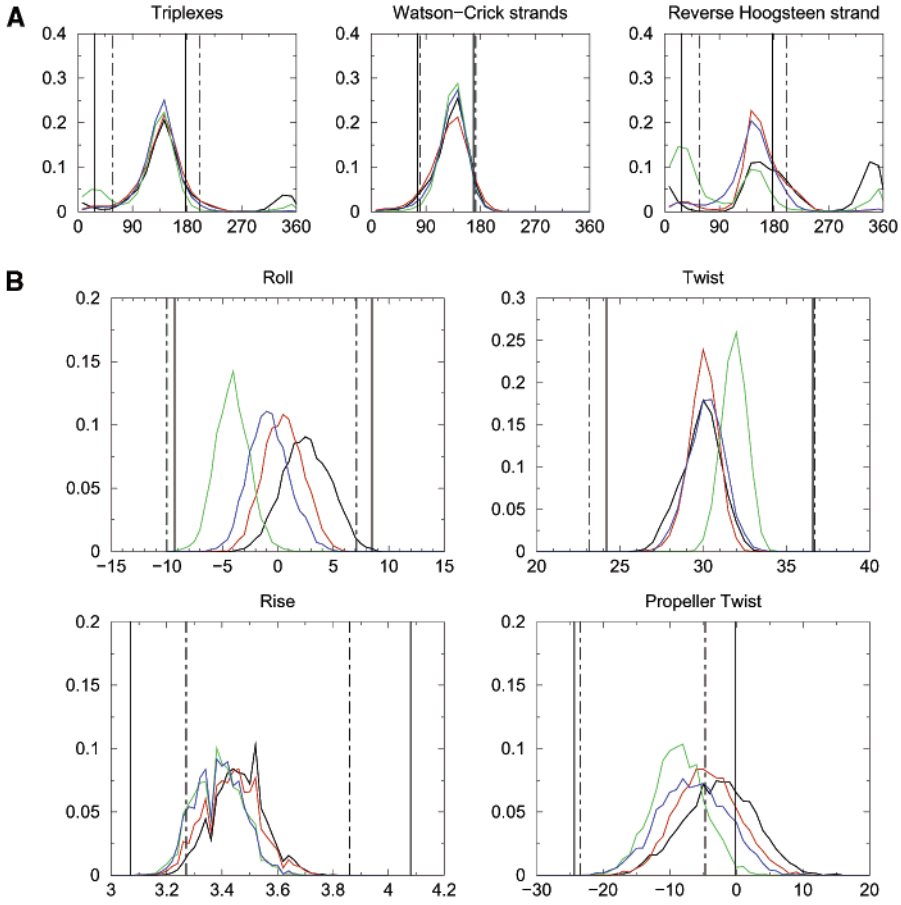


Figure 3. Distribution plots corresponding to (A, left top, left middle, and left bottom) the phase angle of the 2′-deoxyriboses corresponding to all nucleotides, those located in the Watson−Crick strands, and those located in the Hoogsteen strand and (B, middle top, middle bottom, right top, and right bottom) selected helical parameters. Values were obtained by collecting data along the trajectories. The ranges of values found in NMR structures (entries pdb134d and pdb135d) are displayed as lines (solid and dotted−dashed, respectively) in the plots. In all the plots T1 is black, T2 red, T3 green, and T4 blue.

the sugars of the Watson−Crick strands were restricted to the south−southeast region of the pseudorotational circle, but sugars in the third strand were more free to move and sample north regions.

As noted by the standard deviations (Table 4) and the distribution plots (Figure 3), the triplexes displayed a nonnegligible flexibility and sampled wide regions of the helical space during the trajectories. All the values observed during the trajectories fell within the range of variability found in NMR structures (Table 4 and Figure 3), demonstrating that helical parameters obtained from the unrestrained trajectories obtained here are consistent with experimental NMR data. Not surprisingly, the largest deviation in the canonical structure is found in triplexes containing d(T#A·T) triads, which, when present adjacent in the sequence, introduce significant distortions in the third strand.

Interestingly, most helical values mirror those obtained for parallel triplexes,[20a,b] which is not surprising considering the small RMSD between parallel and antiparallel triplexes. The only major differences between the two conformations were in the grooves (Table 4). In parallel triplexes the major groove of the duplex is divided into two asymmetric grooves (see Figure 1 and ref 20a for nomenclature): the minor part (mM) and the major part (MM) of the major groove. For parallel d(T−A·T) triplexes the widths (measured as the shortest P−P distance) of the grooves were around 17 Å (MM), 12 Å (m), and 9 Å (mM). The situation differed completely for the antiparallel triplexes studied here, in which the presence of the reverse-Hoogsteen strand broke the major groove of the duplex more symmetrically. Thus, on average the MM groove was only 1 Å wider than the mM groove, compared to the large difference (8 Å) found in parallel triplexes. The width of the minor groove in antiparallel triplexes was around 12 Å, a value similar to that obtained for parallel triplexes, and for normal B-DNA duplexes,[20a] showing that the minor groove is not dramatically altered by the presence of the third strand. The effect of the sequence in the width of the groove was moderate, implying a reduction in the width of the m groove and a parallel increase in the width of the mM groove for triplexes containing the relatively unstable d(T#A·T) triads.

cMIP analysis shows that the potentiality for interaction of cations with the triplex depends on the sequence (see Figure 4 and methods). For triplex T1 the regions of better interaction were located in the MM and mM grooves, while the m groove was not such a good target probably because of the presence of the 2-amino groups of guanines, as found in DNA duplexes. Triplex T2 showed favorable regions of interaction in the three grooves, the MM groove being the best target for cationic interaction. Finally, triplexes T3 and T4 showed a marked region of favorable interaction in the MM groove and another located in the m groove at steps containing d(T#A·T) triads. The phosphate charge distribution and the presence of a suitable H-bond acceptor at the bottom of the groove (Figure 1) justify the existence of these regions of favorable cMIP(O^+).

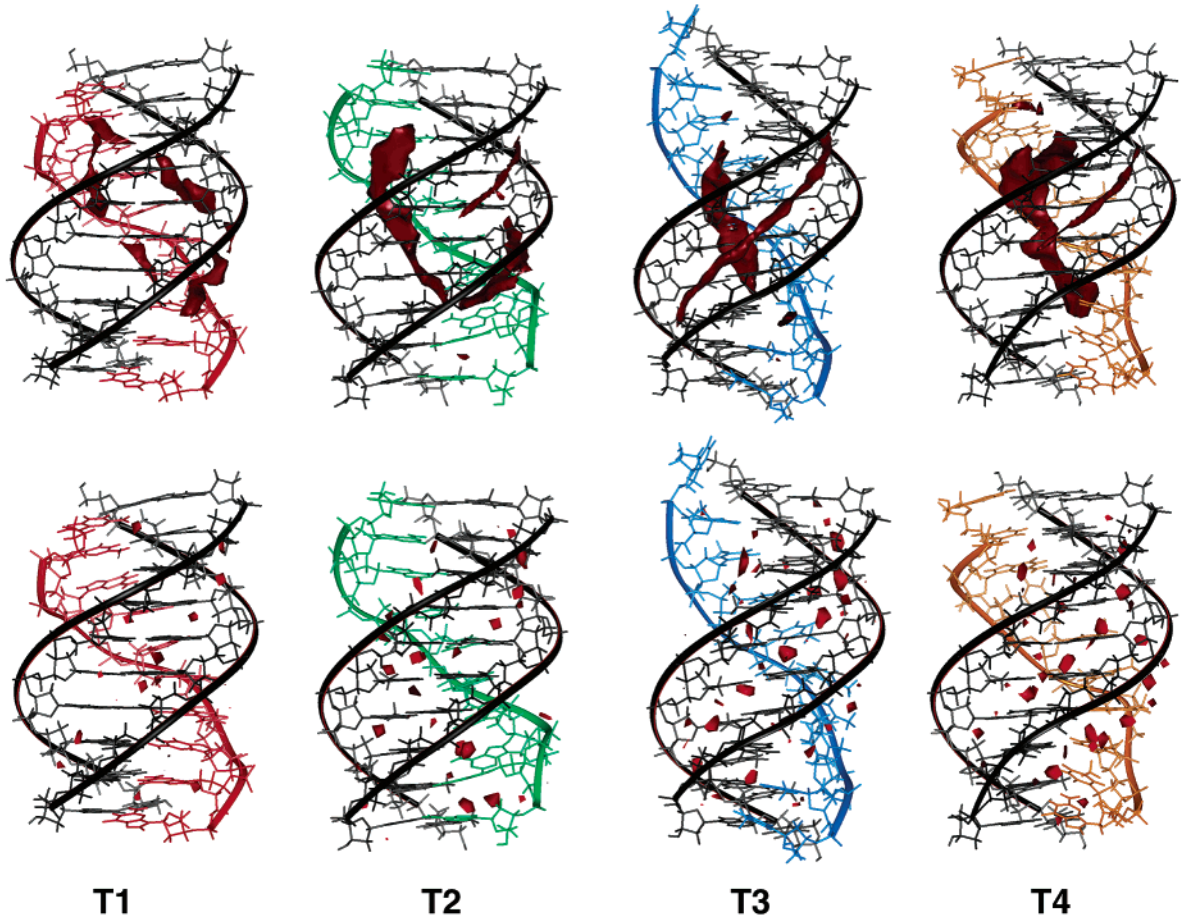


Figure 4. cMIP (top) and solvation (bottom) maps corresponding to triplexes T1−T4 (from left to right). The cMIP contour corresponds to an interaction energy of −7 kcal/mol. Solvation maps correspond to a density of water of 3.5 g/cm^3.

Density maps provide an MD-averaged picture of the ability of the distinct triplexes to interact with water. All the triplexes were well hydrated, with extended regions where the population of water was 3.5-fold higher than the background of the simulation (Figure 4). Clear spines of hydration were found for all the triplexes located in the m groove (Figure 4). These spines were not disrupted in the presence of d(G#G·C) triads, despite the perturbing effect of the 2-amino group of the Watson−Crick guanines. Additional strands of water were found in the mM groove (clearer in triplexes T3 and T4), and small strands of water were also located in the MM groove of the four triplexes. In summary, the water seems to take advantage of its ability to act as H-bond donor and acceptor and interacts very well with all the triplex structures. Overall, the patterns of hydration found here do not differ greatly from those previously obtained for parallel triplexes, despite differences in the groove sizes of parallel and antiparallel conformations.[20a]

Free Energy Calculations. MD calculations allowed us to obtain a qualitative picture of the antiparallel triplex and provided us equilibrated structures which can be used to determine the expected change in stability induced by the presence of modified nucleobases in the structure. Particularly, we wanted to explore the effect of substitution of purines (in the Watson−Crick strand) by 8-aminopurines. These substitutions were able to strongly stabilize parallel triplexes,[21,22,23g,h] and simple inspection of Figure 1 suggested that this could also be the case for antiparallel triplexes. To investigate this hypothesis, models of triplexes containing 8-amino derivatives were generated by manipulation of MD-equilibrated structures, and were then relaxed and reequilibrated as described in the Methods (see Table 1). Comparison of the corresponding trajectories (T1n, T1nn, T4n, and T6n) with the parent trajectories T1, T4, and T6 shows that the structural impact of the presence of 8-aminopurines in the triplexes is very small and localized in the substituted triad, which becomes more planar.

The substitution of the Watson−Crick guanine (d(G#G·C) triad) by 8-aminoguanine and the substitution of adenine (d(T#A·T) triad) by 8-aminoadenine led to an improvement of nearly 6 kcal/mol in the reverse-Hoogsteen hydrogen-bonding energy, with little (0.4−0.1 kcal/mol) impact on the stacking energy and Watson−Crick hydrogen-bonding energy (around −0.2 kcal/mol), which suggests that the introduction of 8-aminopurine derivatives in antiparallel triplexes could be energetically favorable, at the level of nucleobase−nucleobase interaction. To account for other energy terms, solvation, and entropic effects, we performed MD/TI simulations in which the reversible works associated with the mutations A ↔ A^N and G ↔ G^N were computed using the procedure explained in the Methods (see Table 5). The mutation profiles were smooth and did not show clear discontinuities (data available upon request), suggesting the lack of hysteresis effects, and the standard errors in the free energy estimates were very small (0.2−0.4 kcal/mol for the eight independent estimates), indicating a good convergence in the averages. In summary, despite the caution necessary

Table 5. Free Energy Changes in the Triplex → Duplex Transition Induced by Changes from Purine to 8-Aminopurine[a]

mutation	triplex	ΔΔG (kcal/mol)
G → G^N	T1n	2.1(0.2)
G → G^N	T1nn	1.5(0.2)
A → A^N	T4n	2.6(0.2)

[a] A positive sign implies stabilization of the triplex. Standard errors are shown in parentheses.

for subnanosecond MD/TI calculations, we can be confident in the statistical quality of the free energy values reported in Table 5.

Free energy values suggest that the presence of a single 8-aminopurine strongly stabilizes triplexes based on both d(G#G·C) and d(A#A·T) triads. Interestingly, the results for triplexes T1n and T1nn were similar, which suggests that the presence of G^N in the Hoogsteen strands (as in some of the experimental models used in this study) does not alter the triplex-stabilizing effect of the Watson−Crick G^N. Remarkably, the gains in stability of antiparallel triplexes induced by the 8-aminopurines were similar to those induced by the same molecules in parallel triplexes, indicating that the introduction of these molecules is an almost universal mechanism of triplex stabilization. As in antiparallel triplexes, the gain in H-bonding and the entropy gain related to the liberation of waters in the mM groove appeared to be responsible for the gain in stability of the triplexes induced by the presence of G^N or A^N in the Watson−Crick position.

Experimental Studies on the Stability of a Short Intermolecular Antiparallel Triplex Containing 8-Aminoguanines. MD simulations provided us with tridimensional structures of antiparallel triplexes which can be used to obtain quantitative prediction of the effect of substitution of purines by 8-aminopurines in the stability of the triplex. Calculations predict that G → G^N and A → A^N substitutions dramatically stabilize the corresponding antiparallel triplex, which could have large practical importance in antigene- and antisense-oriented methods based on the formation of antiparallel triplexes (see the Introduction), and must be verified experimentally. For this purpose, we prepared oligonucleotides carrying 8-aminopurines and studied the triplex-forming properties of the same.

First, the effect produced by the presence of 8-aminoguanine in antiparallel triplexes was analyzed on a short intermolecular triplex described by Pilch et al.[40] Triplexes formed by d(C3T4C3)·2[d($GG^NG^NA4G^NG^NG$)] were analyzed by melting and gel-shift experiments. The stoichiometry associated with the interaction of these two oligonucleotides was determined by PAGE retardation assay. Mixtures containing 1:1 and 1:2 stoichiometric ratios of d(C3T4C3) and d($GG^NG^NA4G^NG^NG$) were separated by PAGE. As described by Pilch et al.,[40] only a single band, corresponding to the duplex, was present at a 1:1 ratio. This band showed a mobility similar to that of the purine strand. At a 1:2 ratio a new band with reduced mobility was observed, which corresponded to the triplex (see the Supporting Information).

Thermal denaturation curves of the d(C3T4C3)·d($GG^NG^NA4G^NG^NG$) duplex and d(C3T4C3)·2[d($GG^NG^NA4G^NG^NG$)] triplex were determined spectrophotometrically at 260 nm in 50 mM $MgCl_2$, 10 mM sodium cacodylate, and 0.1 mM EDTA at pH 7.2. As previously reported[40] in these conditions a single transition (triplex → random coil and duplex → random coil) is detected. When the dependence of the melting temperature on DNA concentration was studied, a lineal dependence between the concentration and $1/T_m$ was found, which allowed us to obtain ΔH, ΔS, and ΔG from the regression equation (Table 6). The substitution of four guanines for 8-aminoguanines changed the ΔG of the duplex to random coil transition from −12.6 to −10 kcal/mol (a decrease of 2.6 kcal/mol). On the contrary, the same substitution changed the ΔG of the triplex to random coil transition from −26.3 to −28.4 kcal/mol (an increase of 2.1 kcal/mol). Combination of these numbers shows that the presence of four G^N's stabilizes the triplex (with respect to the duplex) by 4.7 kcal/mol. Assuming that the effect of G^N is fully additive, stabilization of around 1.2 kcal/mol per substitution is found. This experimental value compares very well with the MD/TI estimate (1.5 kcal/mol from Table 5). That is, as theoretically suggested, the G → G^N substitution in the Watson−Crick position strongly stabilize antiparallel triplexes containing the d(G#G·C) triad.

Table 6. Thermodynamic Parameters of the Triplex and the Duplex[a]

structure	T_m (°C)	ΔH (kcal/mol)	ΔS [cal/(mol K)]	ΔG_{25} (kcal/mol)
unmodified duplex[b]	52.0[d]	−72[d]	−198[d]	−12.6[d]
8-amino-G duplex[c]	44.7[e]	−28	−63	−10.0
unmodified triplex[b]	54.0[d]	−152[d]	−424[d]	−26.0[d]
8-amino-G triplex[c]	53.5[e]	−133	−350	−28.4

[a] Data obtained in 10 mM sodium cacodylate, 50 mM $MgCl_2$, and 0.1 mM EDTA at pH 7.2. ΔG_{25} refers to the standard free energy change at 25 °C. [b] Unmodified duplex, d($C_3T_4C_3$)·d($G_3A_4G_3$); unmodified triplex, d($C_3T_4C_3$)·2[d($G_3A_4G_3$)]. [c] 8-Amino-G duplex, d($C_3T_4C_3$)·d($GG^NG^NA_4G^N$-G^NG); 8-amino-G triplex, d($C_3T_4C_3$)·2[d($GG^NG^NA_4G^NG^NG$)]. [d] Data from ref 1. T_m at a 15 μM concentration. [e] T_m at a 4 μM concentration.

Thermal Stability of Triplexes Formed from Reverse-Watson−Crick Hairpins. As discussed in previous studies (see the Introduction), an interesting alternative to create triplexes which takes advantage of the stabilizing effect of 8-aminopurines in the Watson−Crick position consists of using duplexes to target single-stranded nucleic acids, instead of the usual approach of targeting duplexes with single-stranded DNA. In this context, we aimed to analyze whether oligonucleotides carrying 8-aminopurines and designed to have the possibility to form antiparallel reverse-Hoogsten hairpins can be templates for antiparallel triplex formation. In this part of the study, and mimicking our previous work with parallel triplexes, the effect of 8-aminoguanine, 8-aminoadenine, and 8-aminohypoxanthine (I^N) in the triplex-forming properties of reverse-Hoogsteen hairpins was studied using the oligonucleotides displayed in Figure 5, which were designed to make antiparallel triplexes with the polypyrimidine sequence WC-11-mer (5′TCTCCTC-CTTC3′). The parent polypurine−polypyrimidine sequences (H26GA and H26GT) were taken from a parallel triplex studied by several authors,[23i,41,42] where the linking between the Watson−Crick polypurine and the reverse-Hoogsteen strands was done by a tetrathymidine loop. 8-Aminopurine derivatives were introduced in a number of positions of the Watson−Crick strand of the putative hairpin. Thus, in oligonucleotides H26GA-($2A^N$) and H26GT($2A^N$) two adenines were replaced by two 8-aminoadenines, in oligonucleotides H26GA($2G^N$) and H26GT-($2G^N$) two guanines were replaced by two 8-aminoguanines, and in H26GA($2I^N$) and H26GT($2I^N$) two guanines were

H26GT : 5'GAAGGAGGAGA-TTTT-TGTGGTGGTTG3'
H26GA: 5'GAAGGAGGAGA-TTTT-AGAGGAGGAAG3'
H26GT(2A^N): 5'GAAGG**A^N**GG**A^N**GA-TTTT-TGTGGTGGTTG3'
H26GA(2A^N: 5'GAAGG**A^N**GG**A^N**GA-TTTT-AGAGGAGGAAG3'
H26GT(2G^N): 5'GAAG**G^N**AG**G^N**AGA-TTTT-TGTGGTGGTTG3'
H26GA(2G^N): 5'GAAG**G^N**AG**G^N**AGA-TTTT-AGAGGAGGAAG3'
H26GT(2I^N): 5'GAAG**I^N**AG**I^N**AGA-TTTT-TGTGGTGGTTG3'
H26GA(2I^N): 5'GAAG**I^N**AG**I^N**AGA-TTTT-AGAGGAGGAAG3'
H26GT(5A^N): 5'G**A^{N}A^N**GG**A^N**GG**A^N**G**A^N**-TTTT-TGTGGTGGTTG3'
H26GA(6G^N): 5'**G^N**AA**G^{N}G^N**A**G^{N}G^N**A**G^N**A-TTTT-AGAGGAGGAAG3'
H26GT(2I): 5'GAAG**I**AG**I**AGA-TTTT-TGTGGTGGTTG3'
H26GA(2I): 5'GAAG**I**AG**I**AGA-TTTT-AGAGGAGGAAG3'

H26contGT: 5'GAAGGAGGAGA-TTTT-GTGTGGTTTGT3'
H26contGA: 5'GAAGGAGGAGA-TTTT-GAGAGGAAAGA3'

S11pur: 5'GAAGGAGGAGA3'
S11pur2A^N: 5'GAAGG**A^N**GG**A^N**GA3'
S11pur2G^N: 5'GAAG**G^N**AG**G^N**AGA3'
S11pur2I: 5'GAAG**I**AG**I**AGA3'

Figure 5. Sequences of oligonucleotides carrying 8-aminopurines as prepared in this study: G^N, 8-aminoguanine; A^N, 8-aminoadenine; I^N, 8-aminohypoxanthine.

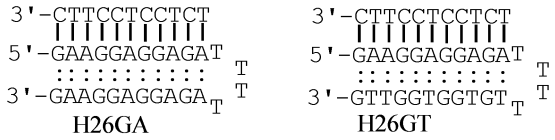
replaced by two 8-aminohypoxanthines. The corresponding oligonucleotides carrying two hypoxanthines, H26GA(2I) and H26GT(2I), were also prepared for comparison. Finally, in H26GT(5A^N) and H26GA(6G^N) all adenines and guanines at Watson−Crick positions were replaced by 8-aminoadenines and 8-aminoguanines, respectively. Control oligonucleotides with a scrambled Hoogsteen strand (H26contGT and H26contGA) and without the reverse-Hoogsteen strand were also prepared (S11pur, S11pur2A^N, S11pur2G^N) as described in the Methods.

The relative stability of triple helices formed by H26GA and H26GT hairpins and the polypyrimidine target sequence (WC-11-mer) was measured spectrophotometrically at 260 nm in 50 mM $MgCl_2$, pH 7.2. In all cases, one single transition characterized as a transition from a triple helix to a random coil was observed with 15% hyperchromicity. Monophasic curves were observed only when H26GA and H26GT hairpins were mixed with the polypyrimidine target sequence (WC-11-mer). This finding strongly suggests that H26GA and H26GT are not fully preorganized before binding to the third strand. This observation clearly contrasts with the behavior for Hoogsteen hairpins,[23h,i] which are fully organized even in the absence of the Watson−Crick polypyrimidine strand.

Triplexes obtained by incubation of hairpins H26GA and H26GT and their derivatives with the polypyrimidine oligonucleotide d(TCTCCTCCTTC) showed melting temperatures in the range 54−77 °C (Table 7). Interestingly, the control duplex formed by the WC-11-mer and the corresponding polypurine strand (without the reverse-Hoogsteen strand, S11 derivatives) melted at lower temperatures (42−50 °C). The addition of a non-sense Hoogsteen strand (H26contGT and H26contGA) gave a similar melting temperatures to control duplexes without a Hoogsteen strand (S11 derivatives). In summary, H26GA and H26GT lead to a very stable and specific triplex at physiological temperatures.

According to our MD and MD/TI calculations, replacement of purines by 8-aminopurines in the Watson−Crick strand should stabilize the triple helix. This is confirmed once again

Table 7. Melting Temperatures (°C) for Triplexes Formed by H26GA and H26GT Derivatives and WC-11-mer[a]



hairpin	T_m[b]	ΔT_m[c]	ΔT_m/substitution	duplex (S11)	duplex (H26con)
H26GT	60.5			50.0[d]	51.0[h]
H26GT(2A^N)	61.7	+1.2	+0.6	45.6[e]	
H26GT(2G^N)	68.0	+7.5	+3.7	41.1[f]	
H26GT(2I^N)	61.5	+1.0	+0.5		
H26GT(2I)	52.8	−7.7	−3.8	38.5[g]	
H26GT(5A^N)	66.6	+6.1	+1.2		
H26(GA)	62.7			50.0[d]	50.0[i]
H26GA(2A^N)	56.5	−6.2	−3.1	45.6[e]	
H26GA(2G^N)	71.0	+8.3	+4.1	41.1[f]	
H26GA(2I^N)	62.4	−0.3	−0.1		
H26GA(2I)	53.1	−9.6	−4.8	38.5[g]	
H26GA(6G^N)	77.2	+14.5	+2.4		

[a] Data obtained in 10 mM sodium cacodylate, 50 mM $MgCl_2$, and 0.1 mM EDTA at pH 7.2. [b] At a concentration of approximately 3 μM. [c] $\Delta T_m = T_m - T_m$ of the corresponding unmodified H26 derivative (H26GT or H26GA). [d] S11pur:WC-11-mer control duplex. [e] S11pur2AA:WC-11-mer control duplex. [f] S11pur2AG:WC-11-mer control duplex. [g] S11purI: WC-11-mer control duplex. [h] H26contGT:WC-11-mer control duplex. [i] H26contGA:WC-11-mer control duplex.

by the observation of an increase in melting temperature of the triplex between 2 and 4 °C per substitution (H26GT(2G^N), ΔT_m/substitution = +3.7 °C; H26GA(2G^N), ΔT_m/substitution = +4.1 °C; H26GA(6G^N), ΔT_m/substitution = +2.4 °C). Replacement of adenine by 8-aminoadenine in the H26GT hairpin stabilized the triplex around 1 °C per substitution (H26GT(2A^N), ΔT_m/substitution = +0.6 °C; H26GT(5A^N), ΔT_m/substitution = +1.2 °C), demonstrating that, as predicted from MD and MD/TI calculations, the A → A^N substitution also stabilizes antiparallel triplexes. It is worth analyzing the results obtained by the introduction of 8-aminoadenines in the H26GA hairpin, since in this case the triads involving adenines should be d(A#A·T), where molecular models suggest that 8-aminoadenines should destabilize the triplex due to strong amino−amino interactions (test MD simulations lead to opened triads). The experimental data in Table 7 confirm this prediction (H26GA(2A^N), ΔT_m/substitution = −3.1 °C). This strongly suggests that the effect of 8-aminopurine in the melting curves is due to specific (d(G#G·C) and d(T#A·T)) triad stabilization, as predicted by simulations, and not by an unspecific aggregation effect.

As a complement to the study, and following previous studies,[22b,47] we analyzed experimentally the effect of the substitution of guanine by hypoxanthine. If the triplex models explained above are correct, this should produce a strong destabilization in the triplex caused by the loss of one hydrogen bond in the Watson−Crick pair, and this was confirmed (Table 7) (H26GT(2I), ΔT_m/substitution = −3.8 °C; H26GA(2I), ΔT_m/substitution = −4.8 °C). Following the same reasoning, the substitution of hypoxanthine by 8-aminohypoxanthine is expected to recover part of the stability, since the loss of one Watson−Crick hydrogen bond is partially compensated by a new hydrogen bond in the reverse-Hoogsteen pair (Figure 6). Once again, the experimental data fully agreed with the

(47) Rueda, M.; Luque, F. J.; Orozco, M. *Biopolymers* **2002**, *61*, 52.

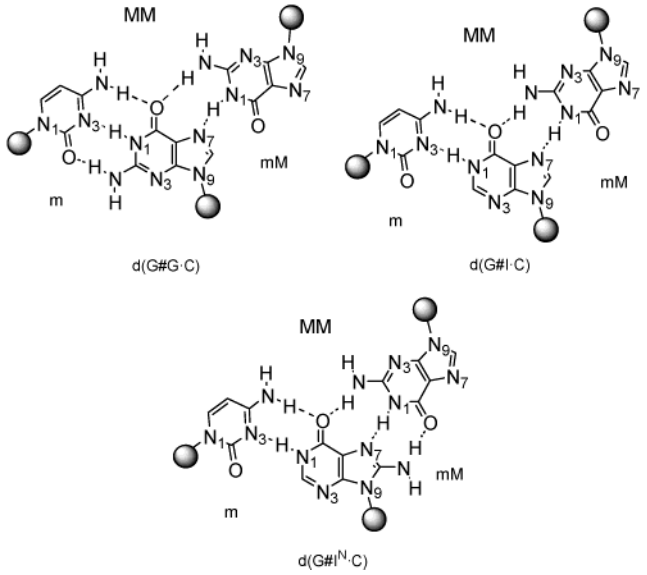


Figure 6. Schematic representation of the reverse-Hoogsteen d(G#G·C) triad and those formed by inosine (I) and 8-aminoinosine (I^N).

Table 8. Thermodynamic Parameters for Triplex to Random Coil Transitions in 10 mM Sodium Cacodylate, 50 mM $MgCl_2$, and 0.1 mM EDTA at pH 7.2 from the Slope of the Plot $1/T_m$ versus ln C[a]

triplex	T_m[b] (°C)	ΔH (kcal/mol)	ΔS [cal/(mol K)]	ΔG (kcal/mol)
H26GT + WC-11-mer	60.7	−83	−222	−17.1
H26GT(2A^N) + WC-11-mer	61.5	−123	−342	−21.7
H26GT(2G^N) + WC-11-mer	67.6	−101	−269	−20.8
H26GT(2I^N) + WC-11-mer	62.0	−96	−260	−18.8
H26GA + WC-11-mer	63.4	−88	−235	−18.2
H26GA(2A^N) + WC-11-mer	57.4	−95	−260	−17.5
H26GA(2G^N) + WC-11-mer	72.5	−122	−325	−24.9

[a] ΔH, ΔS, and ΔG are given as round numbers, and ΔG is calculated at 25 °C, with the assumption that ΔH and ΔS do not depend on temperature; analysis was carried out using melting temperatures obtained from denaturation curves. [b] At 4 μM triplex concentration.

predictions derived from structural models, indirectly supporting their quality.

Thermodynamic Studies. The dependence of the triplex to random coil transition on DNA concentration was studied in several triplexes. In all cases, the melting temperatures of the transitions decreased with concentration, as expected for a bimolecular transition. The plot of $1/T_m$ versus ln(concentration) was linear, giving a slope and a y-intercept from which ΔH, ΔS, and ΔG were obtained. The ΔG values for the triplex dissociation were −17.1 and −18.2 kcal/mol for the unmodified H26GT:WC-11-mer and H26GA:WC11mer triplexes, respectively (Table 8). The substitution of two adenines by two 8-aminoadenines in the H26GT(2A^N):WC-11-mer triplex gave a difference in ΔG of 4.6 kcal/mol. Assuming that the stabilization of A^N is additive, this would imply a stabilization of the triplex by 2.3 kcal/mol per substitution (MD/TI calculations shown in Table 5 suggested a stabilization of 2.6 kcal/mol). As expected, the same substitution in the reverse-Hoogsteen position produced a destabilization of 0.7 kcal/mol (0.35 kcal/mol per substitution). Triplexes carrying G^N gave stabilizations in ΔG of 6.7 kcal/mol (3.3 kcal/mol per substitution) and 3.7 kcal/mol (1.8 kcal/mol per substitution), values which compare very well with the theoretical estimate of 2.1 kcal/mol in Table 5.[48] This and previous quantitative agreements

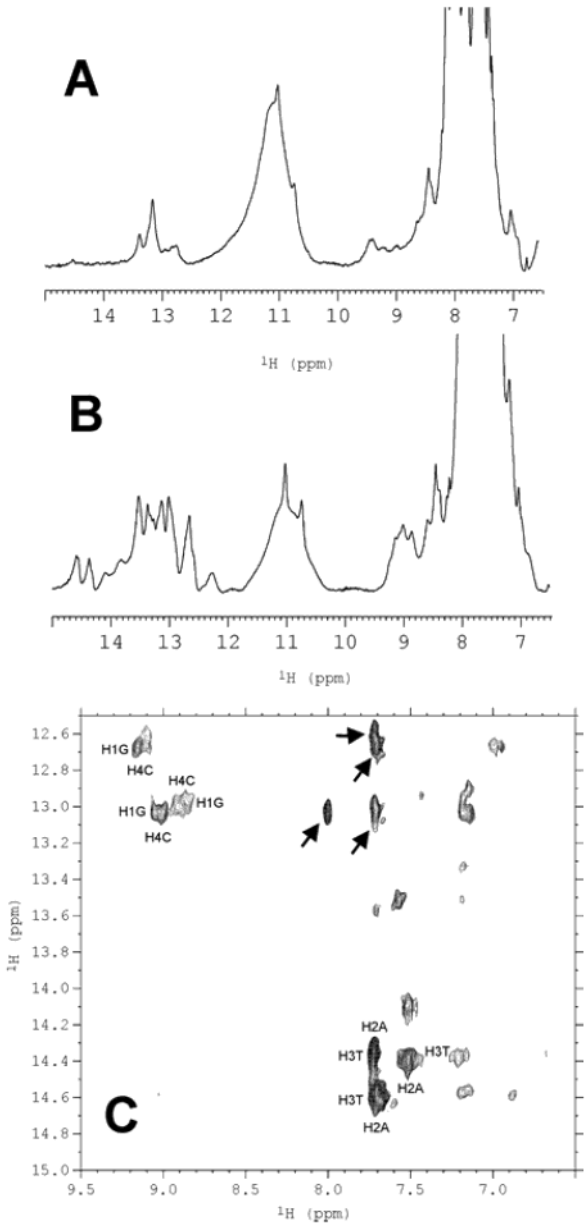


Figure 7. Proton spectra of the triplexes formed by H26GT + WC-11-mer (A) and H26GT(2G^N) + WC-11-mer (B) in H_2O. (C) 2D NOESY spectrum of H26GT(2G^N) + WC-11-mer. NOE cross-peaks characteristic of Watson−Crick base pairs are labeled (amino protons of cytosines with guanine iminos, and iminos of thymines with adenine H2 protons). The arrows indicate cross-peaks between H8 protons of adenines and guanines and Hoogsteen imino protons. All spectra were recorded at $T = 5$ °C and pH 7 with 50 mM Mg_2Cl.

between theory and experiment confirm the utility of "state of the art" molecular simulation as a quantitative tool for predicting the structure and stability of novel nucleic acids.

Structural Data from NMR Experiments. Exchangeable proton NMR spectra of the triplex formed by H26GT and

(48) The transitions followed experimentally here differ slightly from those computed in our simulations, since the experimental values include both Watson−Crick duplex formation and triplex formation from this duplex. The effect of 8-aminopurines in WC duplex stability is always unfavorable (in most cases a few tenths of a kcal/mol; see refs 21, 22a−c, 23g,i, and 47; in other cases larger values can be obtained (see the previous section)). Accordingly, data obtained in this section can be safely used to discuss triplex stability (vs Watson−Crick duplexes or vs single-stranded DNAs), but quantitative comparison with theoretical values must be made with caution.

H26GT(2G^N) with their polypyrimidine target WC-11-mer are shown in Figure 7. In the case of the unmodified triplex, the signals are rather broad, indicating that some conformational or solvent exchange is taking place. This dynamic effect is clearly reduced in the case of the modified triplex, where the signals are much narrower. In both cases the large signal line widths prevent the acquisition of two-dimensional spectra of enough quality for a complete sequential assignment of the proton spectra. However, some key resonances could be identified as shown in Figure 7C.

Chemical shifts of the exchangeable proton signals as well as NOESY cross-peaks clearly show that H26GT(2G^N) hybridizes with WC-11-mer, forming an antiparallel triplex. At least three NOE cross-peaks between the imino protons, with chemical shifts between 14 and 15 ppm, and adenines H2 protons are observed, indicating the formation of A-T Watson−Crick base pairs. Also, cross-peaks between guanine imino protons and some cytosine amino protons were found, showing the occurrence of G-C Watson−Crick base pairs. In addition, some imino protons, with chemical shifts between 12 and 14 ppm, present intense cross-peaks with nonexchangeable base protons. These protons were identified in the D_2O spectra as H8 of adenines or guanines, and their NOEs with iminos are indicative of the formation of reverse-Hoogsteen G#G or T#A base pairs. In summary, NMR data clearly confirm the existence of antiparallel triple helices and the gain in structure obtained by the introduction of 8-aminopurine derivatives.

Conclusions

Molecular dynamics simulations have been used to study a large variety of antiparallel triplexes and to obtain a general view of their structural and dynamic properties. The resulting structural models allow quantitative prediction of the stabilizing effect of the replacement of purines in the Watson−Crick strand by 8-aminopurines. Experimental studies have proven that the structure-based predictions were correct and that 8-aminopurines can be used to obtain stable purine#purine·pyrimidine triplexes. The introduction of 8-aminopurines in purine-rich hairpins allows the targeting of single-stranded polypyrimidine, forming very stable antiparallel triplexes at room temperature. This stability coupled to the lack of an acidic pH requirement strongly suggests that the hairpins described in this work have unique potential use in applications involving oligonucleotide targeting of single-stranded oligonucleotides in vitro and in vivo.

Acknowledgment. M.O. thanks A. López for encouragement and useful discussions. This work was supported by the Spanish Ministry of Science and Technology (Grants BQU2000-0649, CAL01-058-C2-2, and PM99-0046), the Generalitat de Catalunya (Grants 2000-SGR-0018 and 2001-SGR-0049), the Centre de Supercomputació de Catalunya (CESCA), and Cygene Inc.

Supporting Information Available: Figures showing the RMSDs between structures of the 10-mer triplexes sampled during trajectories T1−T7 and reference structures in pdb134d and pdb135d and a photograph of a 15% polyacrylamide gel containg 90 mM TB, 50 mM $MgCl_2$, and differing stoichiometric ratios of d($C_3T_4C_3$) and d($GG^NG^NA_4G^NG^NG$) stained with Stains-All. This material is available free of charge via the Internet at http://pubs.acs.org.

JA035039T