2522–2534 *Nucleic Acids Research, 2001, Vol. 29, No. 12*

The effect of amino groups on the stability of DNA duplexes and triplexes based on purines derived from inosine

Elena Cubero, Ramon Güimil-García[1], F. Javier Luque[2], Ramon Eritja[1] and Modesto Orozco*

Departament de Bioquímica i Biologia Molecular, Facultat de Química, Universitat de Barcelona, Martí i Franques 1, Barcelona 08028, Spain, [1]Instituto de Biología Molecular de Barcelona, C.S.I.C., Jordi Girona 18-26, Barcelona 08034, Spain and [2]Departament de Fisicoquímica, Facultat de Farmàcia, Universitat de Barcelona, Avgda Diagonal s/n, Barcelona 08028, Spain

Received February 22, 2001; Revised and Accepted May 1, 2001

ABSTRACT

The effect of amino groups attached at positions 2 and 8 of the hypoxanthine moiety in the structure, reactivity and stability of DNA duplexes and triplexes is studied by means of quantum mechanical calculations, as well as extended molecular dynamics (MD) and thermodynamic integration (MD/TI) simulations. Theoretical estimates of the change in stability related to 2′-deoxyguanosine (G) → 2′-deoxyinosine (I) → 8-amino-2′-deoxyinosine (8AI) mutations have been experimentally verified, after synthesis of the corresponding compounds. An amino group placed at position 2 stabilizes the duplex, as expected, and surprisingly also the triplex. The presence of an amino group at position 8 of the hypoxanthine moiety stabilizes the triplex but, surprisingly, destabilizes the duplex. The subtle electronic redistribution occurring upon the introduction of an amino group on the purine seems to be responsible for this surprising behavior. Interesting 'universal base' properties are found for 8AI.

INTRODUCTION

Hydrogen bonds (H-bonds) have a decisive influence on DNA structure and function. Watson–Crick H-bonds are the physical basis of specific A·T and G·C recognition in DNA duplexes (1). Hoogsteen H-bonds play a key role in the formation of pyrimidine-based triplexes of DNA (2,3), and G.G H-bonded pairs are known (4,5) to be essential for the stability of G-DNA (Fig. 1). The stability of different DNA duplexes and triplexes is frequently explained on the basis of the number of H-bond interactions between complementary strands. Thus, duplexes rich in G·C pairs are expected to be more stable than duplexes rich in A·T pairs due to the existence of 3/2 H-bonds in the G·C/A·T pairs (1). Furthermore, triplexes containing d(G·C-C) triads are suggested to be more stable at acidic pH than at neutral pH due to the gain of an additional H-bond contact in the protonated d(G·C-C) triad at acidic pHs (2,6–8).

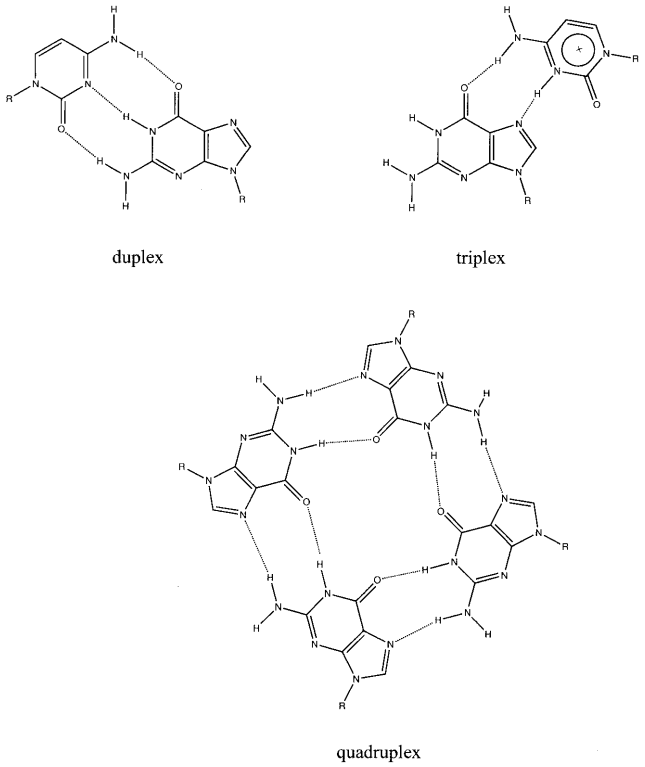

Figure 1. Schematic representations of duplex d(G·C), triplex d(G·C-C)$^+$ and quadruplex d(G)$_4$ pairings.

Increasing the number of H-bond interactions is one of the main strategies to design pseudobases for the stabilization of nucleic acid structures. For instance, the substitution adenine→2,6-diaminopurine was designed to stabilize duplexes by gaining an extra H-bond in the Watson–Crick pairing with thymine (9). Very recently, we studied the ability of 8-aminoadenine and 8-aminoguanine to stabilize DNA triplexes (10–13). One of the reasons for the design of these

*To whom correspondence should be addressed. Tel: +34 93 402 1719; Fax: +34 93 402 1219; Email: modesto@luz.bq.ub.es
Correspondence may also be addressed to Ramon Eritja. Tel: +34 93 400 6145; Fax: +34 93 204 5904; Email: recgma@cid.csic.es

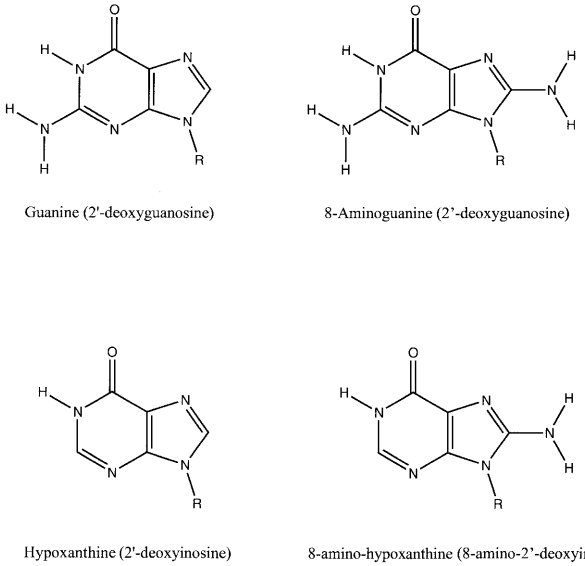

Figure 2. Schematic representation of the nucleobases and 2′-deoxynucleotides (names in parentheses) considered in this study.

two molecules was the expected gain of one extra Hoogsteen-like H-bond with the pyrimidine strand in the triplex structure.

Experimental (14–16) and theoretical (17,18) studies on the structure and stability of apolar bases have raised doubts about the suitability of criteria for predicting the stability of nucleic acid structures based only on counting the number of H-bond contacts. These pseudobases are found in DNA helices that are stable at room temperature, despite the fact that they cannot form H-bonds, even in apolar environments (14–18). It is clear that H-bond interactions might contribute to the stability of DNA helices, but the assumption of a direct relationship between the number of H-bonds and the stability of the helix might not be correct.

In this study we try to gain new insight into the importance of amino H-bond interactions in the stability of duplexes and triplexes based on the hypoxanthine motif. Hypoxanthine (Fig. 2) has a more relaxed recognition pattern than guanine (19,20). The corresponding 2′-deoxy nucleotide (2′-deoxyinosine; I) has been used as a 'universal' base, one that could base pair equally well with any of the four bases, in duplex hybridization (21). It is also known that I can be incorporated into triple-stranded complexes (22–26), specially substituting a 2′-deoxyguanosine (G) in a G·C-C+ triad, but the impact of this substitution on the stability of the triplex remains to be elucidated. Therefore, it is important to analyze the structural and reactive characteristics of duplexes and triplexes containing I, as well as to determine how hypoxanthine can be altered to enhance the stability of the triplex without changing its Watson–Crick recognition pattern, i.e. preserving its 'ambiguity'.

Here we examine, both theoretically and experimentally, the changes in structure, stability and reactivity of duplexes and triplexes when I is mutated to G, and to 8-amino-2′-deoxyinosine (8AI) (Fig. 2). In particular, this is the first time that this latter compound has incorporated into DNA. The combination of experimental and theoretical calculations allows us to obtain a clear picture of the impact of amino groups on duplex and triplex recognition, as well as to determine the triplex and duplex forming characteristics of a new interesting molecule.

MATERIALS AND METHODS

Theoretical calculations

Molecular dynamics (MD) and thermodynamic integration (TI) techniques were used to analyze the changes in structure, reactivity and stability of duplexes and triplexes containing G, I or 8AI (the abbreviation I^N is used to represent 8AI in oligonucleotide sequences to avoid confusion with either A or I). Two duplex families of sequences d(GAGGXTCCAG) and d(GAAGXAGGAG), and one triplex family of sequence d(GAAGXAGGAG), where X is G, I or 8AI, were studied. Starting structures for the MD simulations of duplexes were taken from canonical B-conformation (27), while the starting conformation for triplexes was taken from previously equilibrated MD structures (28,29). The triplexes were defined considering that all the Hoogsteen cytosines were protonated. All systems were surrounded by sodium counterions to obtain a neutral system, and 2140–2170 (duplexes) or 2870 (triplexes) water molecules were added. The systems were then optimized and equilibrated using our standard protocol (28–30). Finally, the equilibrated structures were used in 3 ns of unrestrained MD simulations in the isothermic (T = 298 K), isobaric (P = 1 atm) ensemble. Periodic boundary conditions and the PME technique (31,32) were used to account for long range effects. All bond lengths were kept constant using SHAKE (33), which allowed us the use of a 2 fs time step for integration of Newton equations. AMBER-99 (34,35) and TIP3P (36) force-field parameters were used for all the standard residues. For I and derivatives electrostatic charges were derived from HF/6-31G(d) wavefunctions using the standard RESP procedure (37).

MD/TI calculations (38) were performed to determine the change in stability of duplexes and triplexes related to the mutations G→I→8AI following well-known thermodynamic cycles (Fig. 3). Mutations were performed in single-stranded DNAs of sequences d(GGXTC) and d(AGXAGG), and the duplex and triplex structures mentioned above. The starting structures for these simulations were those obtained after 3 ns of MD trajectories of duplexes and triplexes, and after 1 ns of MD simulation for the corresponding single-stranded oligonucleotides. All TI simulations were performed four times, corresponding to two different simulation lengths (420 and 820 ps), and two directions (forward and reverse) for each mutation (starting from independently equilibrated structures in all cases). A total of 21 or 41 windows were used for each mutation, and free energy values obtained using the first and second halves of each window were recorded. This lengthy and time-consuming simulation protocol allows us to obtain statistical errors in average free energy values (obtained from eight independent estimates), which are necessary to evaluate the predictive value of our simulations. All the technical details of MD/TI simulations are identical to those of the MD simulations mentioned above. Note that MD simulations reported in this paper correspond to >45 ns of unrestrained MD trajectories. To our knowledge, this is one of the largest sets of MD simulations reported on DNA structures.

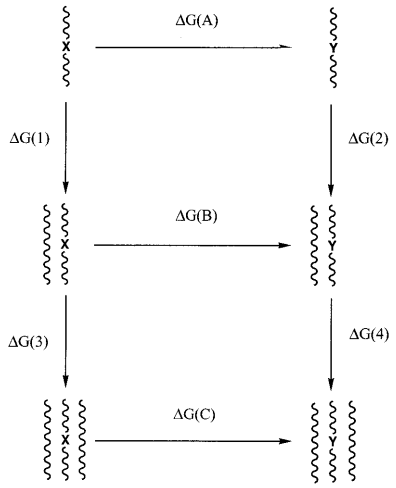

Figure 3. Thermodynamic cycle used to compute differences in the stability of duplex and triplex structures induced by the substitutions G↔I and I↔8AI.

The basic nature of the X·C and X-C H-bonding was analyzed by means of classical energy calculations, as well as QM calculations at the DFT level using the B3LYP/6-31G(d,p) functional (39). Geometries of the G·C, I·C, 8AI·C, G-C$^+$, I-C$^+$ and 8AI-C$^+$ H-bond dimers (where the symbols · and - stand for Watson–Crick and Hoogsteen pairs, respectively) were fully optimized at the DFT level. The nature of minimum energy of the optimized geometries was verified by frequency analysis. Basis set superposition error was corrected using the counterpoise method (40).

Self consistent reaction field (SCRF) calculations were performed to check the ability of MD/TI calculations to properly represent changes in hydration free energy related to the I→G and I→8AI mutations. SCRF calculations were carried out using our HF/6-31G(d) optimized version (41–43) of the MST method (44–46), and HF/6-31G(d) optimized geometries of guanine, hypoxanthine and 8-aminohypoxanthine. Finally, QM-MIPp (47–50) calculations were performed to determine the change in Watson–Crick and Hoogsteen H-bonding properties of inosine due to the introduction of amino groups at positions 2 (guanine) and 8 (8-aminohypoxanthine). QM-MIPp calculations were determined using O$^+$ and O$^-$ as probes, and the HF/6-31G(d) wavefunctions.

Classical calculations were carried out using AMBER5.1 (51) suite of programs. AMBER5.1, CURVES (52), as well as 'in house' developed programs were used to analyze the trajectories. DFT calculations were performed using Gaussian-94 (53). SCRF calculations were done using a locally modified version of the MonsterGauss computer program (54). QM-MIPp calculations were carried out using the MOPETE computer program (55).

Chemical synthesis

8AI (**1**). In a round-bottom flask 2 g (7.5 mmol) of 8-amino-2′-deoxyadenosine (54) was dissolved in 100 ml of 0.1 M aqueous sodium phosphate buffer pH 7.5. Adenosine deaminase (100 mg, 149 U, of crude powder from calf intestinal mucosa) was dissolved with 100 ml of 0.1 M aqueous sodium phosphate buffer pH 7.5 and the solution was added. The resulting mixture was stirred at 37°C for 72 h and concentrated to dryness. The solid was treated with hot methanol and filtered immediately. The methanol solution was concentrated to dryness yielding 2 g (7.48 mmol, 99% yield) of 8AI. TLC (50% ethanol in dichloromethane) R_f = 0.32. Elemental analysis: ($C_{10}H_{13}N_5O_4H_2O$) expected C, 42.11%; H, 5.30%; N, 24.55%; found C, 41.51%; H, 5.32%; N, 24.01%. UV (methanol) max. 263 and 290 (shoulder) nm. ^{1}H-NMR δ (DMSO-d_6) p.p.m.: 8.29 (1H, s, NH), 7.77 (1H, s, H-2), 6.50 (2H, broad s, NH_2), 6.28 (1H, dd, H-1′), 5.43 (1H, t, OH-5′), 5.27 (1H, d, OH-3′), 4.39 (1H, m, H-3′), 3.85 (1H, m, H-4′), 3.63 (2H, m, H-5′), 2.1–1.9 (2H, m, H-2′). ^{13}C-NMR δ (DMSO-d_6) p.p.m.: 155.7 (C-6), 151.2 (C-4), 142.4 (C-2), 146.8 (C-8), 121.5 (C-5), 87.6 (C-1′), 83.2 (C-4′), 71.4 (C-3′), 61.8 (C-5′), 37.7 (C-2′). EM (electrospray, positive mode): 268 (M+H$^+$), 152 (nucleobase +H$^+$) expected for $C_{10}H_{13}N_5O_4$ 267.2.

8-Amino-2′-deoxy-5′-O-dimethoxytrityl-N^8-isobutyryl-inosine (**2**). 8AI (1.72 g, 6.4 mmol) was dissolved in DMF and 2.7 ml of hexamethyldisilazane (12.9 mmol) were added. After 1 h of magnetic stirring at room temperature the solution was concentrated to dryness giving a white foam with the following elemental analysis: ($C_{19}H_{29}N_5O_4Si_2$) expected C, 46.69%; H, 7.10%; N, 17.01%; found C, 46.51%; H, 7.23%; N, 16.92%. The solid was dried by evaporating dry pyridine (10 ml) twice. The resulting residue was dissolved in 20 ml of dry pyridine and treated with phenoxyacetic anhydride (5.6 g, 19.3 mmol) for 16 h at room temperature followed by the addition of 1.3 ml of isobutyryl chloride (12.9 mmol). After 2 h of magnetic stirring at room temperature a new product was observed by TLC (5% ethanol in dichloromethane, R_f = 0.87). The reaction mixture was concentrated to dryness and the oily residue was treated with toluene and evaporated (3 × 10 ml). The same operation was performed with tetrahydrofuran (3 × 10 ml). The residue was dissolved in tetrahydrofuran and 24 ml of 1 M tetrabutylammonium fluoride solution in tetrahydrofuran was added. After 30 min of stirring the mixture was concentrated to dryness. The residue was dissolved in 20 ml of a pyridine/methanol/water (3:1:1) and 30 ml of Dowex 50Wx4 (pyridinium form) were added. The mixture was stirred for 20 min and filtered. The resulting product was purified by column chromatography on silica gel eluted with a 7–15% ethanol gradient in dichloromethane. NMR spectra showed the presence of one phenoxyacetyl group and one isobutyryl group so the product obtained was assigned to 8-amino-2′-deoxy-*N*^{8}-isobutyryl-*N*^{8}-phenoxyacetyl-inosine. Yield 2.19 g (4.64 mmol, 72%). TLC (10% ethanol in dichloromethane) R_f = 0.29. UV (methanol) max.: 269 nm. ^{1}H-NMR (DMSO-d_6) p.p.m.: 8.28 (1H, s, NH), 8.03 (1H, s, H-2), 7.3–6.7 (5H, m, phenyl), 5.92 (1H, dd, H-1′), 5.25 (1H, m, OH), 4.75 (1H, broad s, OH), 4.40 (1H, m, H-3′), 4.25 (2H, s, CH_2 phenoxyacetyl), 3.82 (1H, m, H-4′), 3.64 (2H, m, H-5′), 3.22 (1H, q, CH isobutyryl), 2.18 (2H, m, H-2′), 1.12 (6H, dd, CH_3 isobutyryl).

^{13}C-NMR (DMSO-d_6) p.p.m.: 180.4 (C=O, isobutyryl), 176.6 (C=O, phenoxyacetyl), 160.4 (C-1, phenyl), 159.3 (C-6), 149.3 (C-4), 147.1 (C-2), 144.1 (C-8), 131.2 (C-3 and C-5 phenyl), 124.1 (C-5), 122.8 (C-4 phenyl), 116.4 (C-2 and C-6 phenyl), 90.4 (C-1′), 87.4 (C-4′), 73.9 (C-3′), 68.6 (CH_2 phenoxyacetyl), 64.2 (C-5′), 41.2 (C-2′), 37.0 (CH isobutyryl), 20.7 y 20.4 (CH_3 isobutyryl). EM (electrospray, positive mode): 472.2 (M+H), 376.2 (M-phenoxy), 338.2 (M-phenoxyacetyl), 222.2 (*N*-phenoxyacetyl-*N*-isobutyrylamide) expected for $C_{22}H_{25}N_5O_7$ 471.4.

8-Amino-2′-deoxy-N^8-isobutyryl-N^8-phenoxyacetyl-inosine (2.19 g, 4.6 mmol) was dissolved into dry pyridine (20 ml) and dimethoxytrityl chloride (1.73 g, 5.1 mmol) was added to the solution. After 2 h of magnetic stirring at room temperature, methanol (2 ml) was added and the solution was concentrated to dryness. The residue was dissolved in dichloromethane and the solution was washed with aqueous sodium bicarbonate solution and brine. The organic phase was dried over Na_2SO_4 and evaporated. Silica gel chromatography (2–10% ethanol gradient in dichloromethane) yielded 2.1 g (3.3 mmol, 72% yield) of 5′-dimethoxytrityl-N^8-isobutyryl-8-amino-2′-deoxy-inosine. Unexpectedly, the phenoxyacetyl group was eliminated during the work-up. TLC (10% ethanol in dichloromethane) R_f = 0.4. Elemental analysis: ($C_{35}H_{37}N_5O_7$) expected C, 65.72%; H, 5.83%; N, 10.95%; found C, 64.62%; H, 5.87%; N, 10.60%. UV (methanol) max.: 236, 274 (shoulder) and 283 (shoulder) nm. ^{1}H-NMR (DMSO-d_6) p.p.m.: 8.29 (1H, s, NH), 7.63 (1H, s, H-2), 7.4–6.7 (13H, m, DMT), 5.92 (1H, t, H 1′), 4.82 (1H, m, 3′-OH), 4.45 (1H, m, H-3′), 3.95 (1H, m, H-4′), 3.35 (2H, m, H-5′), 3.18 (1H, q, CH isobutyryl), 2.2 (2H, m, H-2′), 1.16 (6H, CH_3 isobutyryl). ^{13}C-RMN (Cl_3CD): 179.2 (C=O, isobutyryl), 158.4 (C-4 anisoyl DMT), 157.7 (C-6), 147.9 (C-4), 144.7 (C-8), 144.7 (C-1 phenyl DMT), 143.9 (C-2), 136.1 (C-1 anisoyl DMT), 130.0 (C-2 and C-6 anisoyl DMT), 128.2 (C-2 and C-6 phenyl DMT), 127.7 (C-3 and C-5 phenyl DMT), 126.8 (C-4 phenyl DMT), 121.6 (C-5), 113.0 (C-3 and C-5 anisoyl DMT), 86.3 (C-1′), 86.3 (Cquat. DMT), 85.9 (C-4′), 72.6 (C-3′), 64.2 (C-5′), 55.1 (CH_3O DMT), 38.1 (C-2'), 35.4 (CH isobutyryl), 19.4 and 18.9 (CH_3 isobutyryl). EM (electrospray, positive mode): 640.1 (M+H), 303.3 (DMT+) expected for $C_{35}H_{37}N_5O_7$ 639.7.

8-Amino-2′-deoxy-5′-O-dimethoxytrityl-N^8-isobutyryl-inosine-3′-O-(2-cyanoethyl)-N,N-diisopropylphosphoramidite (**3**). The protected nucleoside (1 g, 1.56 mmol) described above was dissolved in dry dichloromethane (30 ml) and diisopropylethylamine was added (0.85 ml, 4.7 mmol). Chloro 2-cyanoethoxy diisopropylamino phosphine (0.6 ml, 0.34 mmol) was added to the solution dropwise with a syringe. After 1 h of magnetic stirring at room temperature, methanol (2 ml) was added and the solution was concentrated to dryness. The resulting residue was dissolved in dichloromethane and washed with aqueous sodium bicarbonate solution and brine. The organic phase was dried over Na_2SO_4 and evaporated. Silica gel chromatography (ethyl acetate/hexane 1:2) yielded 1.1 g (1.31 mmol, 84% yield) of the desired phosphoramidite. TLC (ethyl acetate/hexane 1:2) R_f = 0.34. ^{31}P-NMR (Cl_3CD): 143.07 and 143.37 (two diastereoisomers).

Oligonucleotide synthesis

Oligonucleotides were prepared on an automatic DNA synthesizer using standard and the modified phosphoramidite of the 8-aminohypoxanthine described above. The phosphoramidite of protected 8AI (**3**) was dissolved in dry dichloromethane to make a 0.1 M solution. The rest of the phosphoramidites were dissolved in dry acetonitrile (0.1 M solution). Sequences of oligonucleotides containing 8-aminohypoxanthine were: A, 5′-CTAI^{N}G-3′; B, 5′-GCA ATG GAI^N CCT CTA-3′; C, 5′-GAA GI^{N}A GGA GAT TTT TCT CCT CCT TC-3′; D, 5′-GAA GI^{N}A I^{N}GA I^{N}AT TTT TCT CCT CCT TC-3′, I^N being = 8AI. Complementary oligonucleotides containing the natural bases and I were also 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. Ammonia solutions were concentrated to dryness and the products were purified by reverse-phase HPLC. Oligonucleotides were synthesized on 0.2 µmol scale and with the last DMT group at the 5′-end (DMT on protocol) to help reverse-phase purification. All purified products presented a major peak, which was collected. Characterization of pentanucleotide A, 5′-CTAI^{N}G-3′, was perfomed by mass spectrometry and by snake venom phosphodiesterase and alkaline phosphatase digestion followed by HPLC analysis of the nucleosides (HPLC conditions B). 8AI eluted together with dG. EM (MALDI): 1501.5 expected for $C_{49}H_{62}N_{20}O_{28}P_4$ 1502.8. Nucleoside composition: dC 1.1 (**1**), dG + 8-amino-dI 1.8 (**2**), T 1.1 (**1**), dA 1.1 (**1**). Yield (OD units at 260 nm after HPLC purification, 0.2 µmol) were 15mer B 10.3 OD, 26mer C 18.4 OD, 26mer D 15.4 OD. HPLC conditions: HPLC solutions were as follows. Solvent A, 5% ACN in 100 mM triethylammonium acetate pH 6.5 and solvent B, 70% ACN in 100 mM triethylammonium acetate pH 6.5. For analytical runs the following conditions were used. Column, Nucleosil 120C_{18}, 250 × 4 mm; flow rate: 1 ml/min. (i) Conditions, a 40 min linear gradient from 0 to 75% B; (ii) a 20 min linear gradient from 0 to 20% B. For preparative runs the following conditions were used: columns, 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).

Melting experiments

Solutions of equimolar amounts of the pentadecamer carrying 8AI or I at the central position and its complementary sequences carrying each of the four natural bases opposite the modified base were mixed in 0.15 M NaCl, 0.05 M Tris–HCl buffer pH 7.5. Melting experiments with triple helix were performed by mixing equimolar amounts of the modified oligonucleotide (26mer) and the Hoogsteen pyrimidine strand (11mer) in 1 M NaCl, 0.1 mM sodium phosphate/citric acid buffer. The solutions were heated to 90°C, allowed to cool slowly to room temperature and then samples were kept in the refrigerator overnight. UV absorption spectra and melting experiments (absorbance versus temperature) were recorded in 1 cm path length cells using a spectrophotometer, which has a temperature controller with a programmed temperature increase of 0.5°C/min. Melting experiments were performed using a concentration of 4 µM monitoring the absorbance at 260 nm.

Thermodynamic analysis

Analysis of the melting curves was carried out as described by Loakes and Brown (56). The enthalpy of the melting transition (in kcal/mol) was obtained using the equation $\Delta H = -4.38/(1/T_{1/2} - 1/T_{3/4})$, where $T_{1/2}$ is the temperature at the maximum of the first derivative of the melting curve, and $T_{3/4}$ is the temperature at which the differential curve is half of $T_{1/2}$ (both in degrees Kelvin). A similar analysis on the triplex was not carried out due to several factors such as the absence of transitions at certain pHs, the lower melting temperature (T_m) of hypoxanthine triplexes and the proximity of the triplex transition to the duplex transition of the 8AI triplexes.

Analysis of the thermodynamic data of triplex→duplex transition was performed as previously described (13). Melting curves were obtained at concentrations ranging from 0.5 to 40 μM of triplex. The T_m of the first transition (corresponding to the dissociation of s_{11} from the triplex) was measured at the maximum of the first derivative of the melting curve. The plot of $1/T_m$ versus lnC was linear. Linear regression of the data gave a slope, and a *y*-intercept, from which ΔH_t and ΔS_t were obtained (13). The free energy was obtained from the standard equation: $\Delta G_t = \Delta H_t - T\Delta S_t$.

The samples used in the concentration-dependent experiments were prepared in a similar way to that described in the previous section, but melting experiments were obtained using 0.1, 0.5 and 1 cm path length cells. The DNA concentration was determined by UV absorbance measurements (260 nm) at 90°C, using for the DNA coil state the following extinction coefficients: 7500, 8500, 12500, 12500, 12500 and 15000 M^{-1} cm^{-1} for C, T, G, 8AI, I and A, respectively.

RESULTS AND DISCUSSION

Synthesis

To our knowledge the synthesis of 8AI has not been described, although the synthesis of the corresponding ribonucleoside, 8-aminoinosine, was reported in 1967 (57). First, a synthetic route similar to the route described for the preparation of 8-aminoinosine was tried. This route had the following steps: (i) bromination of position –8 of I, (ii) displacement of bromine with azide or hydrazine and (iii) catalytic hydrogenation. Unfortunately bromination of I under the conditions described for the bromination of inosine was not successful and the conditions described by Long *et al.* (57) for the next steps were considered too extreme for a 2′-deoxynucleoside. For these reasons an alternative route was considered (Fig. 4). In this route, 8-amino-2′-deoxyadenosine was first prepared from dA as described (57). 8AI was obtained by enzymatic deamination of 8-amino-2′-deoxyadenosine by adenosine deaminase. The formation of a more polar compound was followed by TLC and HPLC showing complete reaction after 72 h. Spectral data (mass spectrometry, UV, ^{1}H- and ^{13}C-NMR) of the new compound was consistent with the desired compound (**1**).

Next, protection of the amino group was studied (Fig. 4). Reaction with dimethylaminoformamidine dimethylacetal in methanol and in *N*,*N*-dimethylformamide did not yield the desired dmf-protected nucleoside as described for 8-aminoguanine (12) and 8-aminoadenine (10) derivatives. Instead the

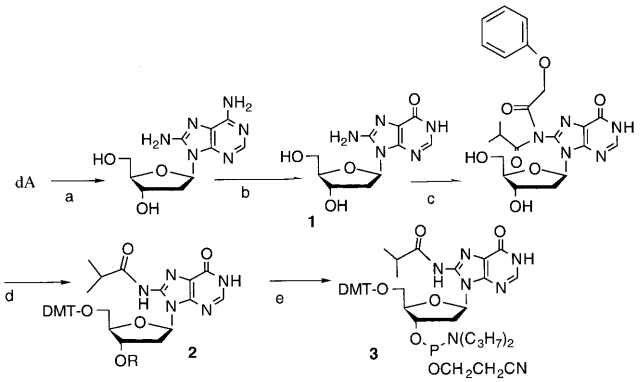

Figure 4. Preparation of the phosphoramidite of 8AI. (a) Reference (56), (b) adenosine deaminase, (c) hexamethyldisilazane followed by phenoxyacetic anhydride in pyridine followed by isobutyryl chloride, (d) dimethoxytrityl chloride in pyridine, (e) chloro 2-cyanoethoxy diisopropylamino phosphine.

starting 8AI was recovered. Other groups such as phenoxyacetyl (PAC) (58) and isobutyryl (ibu) were considered. The introduction of these groups was assayed with 5′,3′-*O*-bis(trimethylsilyl) derivative of 8AI (transient protection). Treatment of the silylated derivative of 8AI with phenoxyacetyl anhydride and isobutyryl chloride in pyridine yielded the expected protected monomers but they were isolated in very low yields (data not shown). We believe that the low recovery yields are due to solubility problems during silica gel purification. We realized that if the reaction of the silylated derivative of 8AI with phenoxyacetyl anhydride was followed by reaction with isobutyryl chloride, a new product was formed that was isolated in good yield (74%). NMR showed the presence of one phenoxyacetyl group and an isobutyryl group on the nucleoside. We believed the PAC group was first incorporated and the PAC-protected nucleoside was further reacted with isobutyryl chloride. The addition of the isobutyryl group may happen at positions N-1 or N-8 and, at this point, no data was conclusive to assign the position of the groups. Most probably the mono-protected derivative has lower solubility on organic solvent than the di-protected derivative and for this reason it was isolated in lower yield.

Reaction of the PAC, ibu-protected nucleoside with dimethoxytrityl chloride in pyridine gave the 5′-*O*-DMT-protected nucleoside (**2**) in 84% yield but during the characterization of the product by NMR the signals corresponding to the PAC group were not present anymore. The removal of the PAC group during the mild conditions used for the preparation of the DMT-derivative are indicative that the PAC and ibu groups were linked to the 8-amino position. In this way, during the work-up following the introduction of the DMT group, the PAC group of a disubstituted (PAC, ibu) amide was hydrolyzed due to its extreme lability. Although during the route described for the preparation of the 5′-*O*-DMT-N^8-ibu derivative of 8AI, the use of the PAC groups seems to be unnecessary, we would like to stress the solubility problems found during the isolation of the ibu-protected derivative of 8AI which are avoided by the preparation of the PAC, ibu-protected nucleoside. We believe that improvements to this route may come by changing the transient protection by the

Table 1. r.m.s.d. (in Å) between the three different trajectories (X = G, I or 8AI) for the two duplexes families [d(GAGGXTCCAG) and d(GAAGXAGGAG)] and different reference structures: canonical A and B forms, and the MD-averaged structures obtained in the three trajectories

Trajectory	B-form	A-form	Av. (X = G)	Av. (X = I)	Av. (X = 8AI)
Duplexes d(GAGGXTCCAG)					
X = G	2.4 (0.4)	4.6 (0.6)	1.7 (0.4)	1.6 (0.4)	1.6 (0.4)
X = I	2.5 (0.4)	4.4 (0.5)	1.8 (0.4)	1.3 (0.3)	1.4 (0.3)
X = 8AI	2.5 (0.5)	4.3 (0.5)	1.8 (0.4)	1.3 (0.3)	1.3 (0.3)
Duplexes d(GAAGXAGGAG)					
X = G	2.7 (0.4)	4.0 (0.5)	1.4 (0.3)	1.7 (0.4)	1.4 (0.4)
X = I	2.4 (0.4)	4.4 (0.5)	1.5 (0.2)	1.5 (0.3)	1.6 (0.3)
X = 8AI	2.7 (0.4)	4.1 (0.5)	1.4 (0.3)	1.5 (0.3)	1.3 (0.3)

The standard deviations (Å) in the r.m.s.d. are given in parentheses.

classical per-acylation method although this route has not been tried in this work (59).

Finally, the phosphitylation of the DMT, ibu-protected nucleoside with the corresponding chlorophosphine gave the desired phosphoramidite (**3**) in 84% yield.

The stability of the isobutyryl group to ammonia was analyzed. First, DMT, ibu-protected nucleoside was dissolved in dioxane and it was treated with concentrated ammonia at 55°C overnight. Complete removal of the ibu group was observed in <6 h as seen by TLC (data not shown). Afterwards, the pentanucleotide 5′-CTAI^{N}G-3′ (I^{N} = 8AI) was prepared and the support was treated with concentrated ammonia at 55°C overnight. The resulting pentanucleotide was analyzed by analytical HPLC giving a major product, which had the expected mass and nucleoside composition.

Theoretical calculations

MD analysis of DNA duplexes. MD simulations of the two families of duplexes studied here [d(GAGGXTCCAG) and d(GAAGXAGGAG)] provide stable trajectories, irrespective of the nature of X (G, I or 8AI). This is clearly noted in the root mean square deviations (r.m.s.d.), which amount to 1.3–1.7

Table 2. Curves-helical parameters (angles in degrees, distances in Å) for the MD averaged structures (the last 1.5 ns were used for averaging) of the duplexes (X = G, I or 8AI) of the two families: d(GAGGXTCCAG) (in roman) and d(GAAGXAGGAG) (in italics)

Structure	X = G	X = I	X = 8AI	B-form[a] (fiber)	B-form[b] (crystal)	A-form[a] (fiber)
Rise	3.4	3.4	3.3	3.3	[3.3,3.4]	3.4
	3.3	*3.4*	*3.4*			
Twist	33.6	32.4	33.0	36	[34,36]	31
	32.5	*33.4*	*33.1*			
Inclination	0.7	–2.6	3.7	–6	[–0.4,1.7]	19
	2.6	*0.4*	*–0.2*			
Roll	2.9	3.1	1.8	–4	[1.7,3.9]	11
	1.8	*2.1*	*2.5*			
X-disp	–0.3	–1.1	–1.7	–1	[–0.8,1.2]	–5
	–1.8	*–1.1*	*–1.5*			
P	141	136	134	191	[129,160]	13
	136	*136*	*134*			
Minor groove width[c]	11.7	11.5	12.3	12	[10,13]	17
	12.3	*11.4*	*12.0*			
Major groove width[c]	18.5	19.4	18.3	17	[17,19]	8
	19.1	*19.0*	*19.7*			

When local and global parameters are available the local values are shown.
[a]Experimental helical values from structures in Arnott and Hukins (27).
[b]Average values reported in Hernández *et al.* (61).
[c]Measured as the shortest P-P distance along the minor/Major grooves.

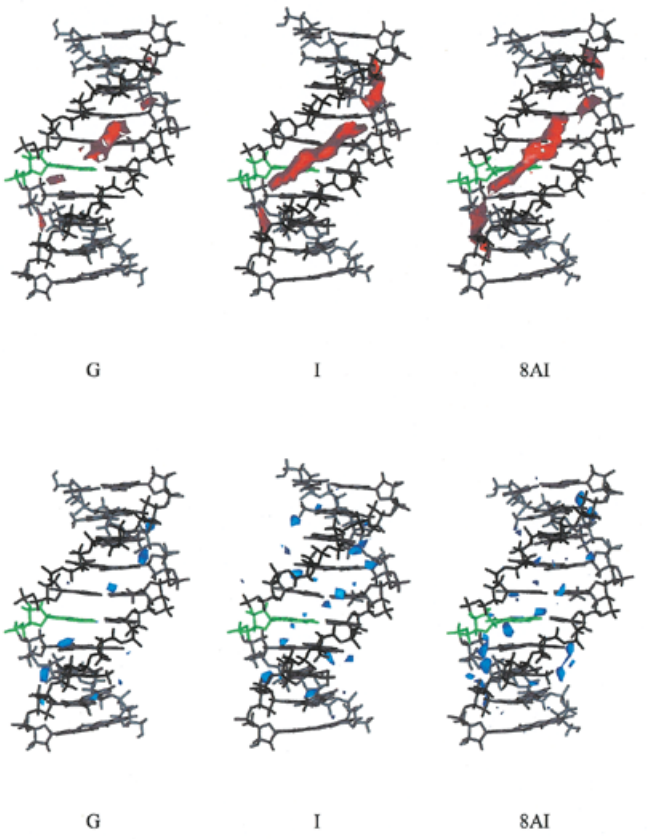

Figure 5. Classical molecular interaction potential (cMIP) (top, in red) and solvation maps (bottom, in blue) for the three duplexes of the family d(GAAG**X**AGGAG) considered in the study (from left to right duplexes X = G, I and 8AI). Contour levels are –5.0 kcal/mol (cMIP), and apparent density of 2.5 g/ml for solvation maps. The base **X** is shown in green.

Table 3. r.m.s.d. (in Å) between the three different trajectories (X = G, I or 8AI) for the triplex family d(GAAGXAGGAG) and different reference structures: canonical A and B forms (*28*), and the MD-averaged structures obtained in the three trajectories

Trajectory	B-form	A-form	Av. (X = G)	Av. (X = I)	Av. (X = 8AI)
X = G	1.4 (0.2)	2.3 (0.3)	1.0 (0.2)	1.1 (0.2)	1.0 (0.2)
X = I	1.4 (0.3)	2.3 (0.2)	1.0 (0.2)	1.0 (0.1)	1.0 (0.2)
X = 8AI	1.4 (0.2)	2.3 (0.2)	0.9 (0.1)	1.0 (0.2)	0.9 (0.2)

The standard deviations (Å) in the r.m.s.d. are given in parentheses.

Table 4. Curves-helical parameters (angles in degrees, distances in Å) for the MD averaged structures (the last 1.5 ns were used for averaging) of the triplexes (X = G, I or 8AI) of the family d(GAAGXAGGAG)

Structure	X = G	X = I	X = 8AI	B-form[a]	A-form[a]
Rise	3.3	3.3	3.3	3.4	3.4
Twist	30.7	30.6	30.7	30	29
Inclination	0.7	2.8	1.3	–2	5
Roll	0.7	0.5	1.9	0	3
X-disp	–3.3	–3.4	–3.4	–3	–3
P	132	131	132.6	117	13
Minor groove width[b]	11.6	12.0	11.8	11.9	16.2
Minor–Major groove width	8.5	8.7	8.5	9.2	7.5
Major–Major groove width	14.0	14.4	14.4	17.1	8.9

When local and global parameters are available the local values are shown.
[a]See Shields *et al.* (28).
[b]Measured as the shortest P-P distance along the different grooves.

(±0.3) Å when the different trajectories are compared with their corresponding MD-averaged structures (Table 1). These r.m.s.d. values are typical of converged MD simulations of DNA duplexes (60,61). Disruptions of the helix or breathing movements are not observed, thus supporting the stability of the helical structure along the trajectories. The structures found for each family of duplexes are in general identical, irrespective of the nature of X. This is clearly seen in cross-r.m.s.d. values (those obtained when one trajectory is compared with the averaged structure of another trajectory) of 1.3–1.8 Å. In summary, our results suggest that the DNA is flexible enough to accommodate possible local distortions arising from the presence of the 8-amino group, without any important structural alteration.

The structures sampled for the six trajectories are those expected for a B-type structure, as noted in r.m.s.d. values of 2.4–2.5 (±0.4) Å for the entire duplexes. This shows that MD-trajectories sample typical regions of the B-type configurational space. This is confirmed by helical analysis of the six trajectories (Table 2), which shows helical values in good agreement to those expected for B-type duplexes, and far from those of an A-type duplex (27).

The classical molecular interaction potential (MIP; 28,30) and solvation (28,30) maps are those expected for B-type duplexes (60). Reactive regions, corresponding to areas of large preferential solvation, are located in the minor groove of the DNA (Fig. 5), defining clearly the shape of the 'spine of hydration' (60,62–64). As expected, the region of large electronegative potential and large solvation is partially disrupted in the vicinity of d(G·C) pairs. Interestingly, when G is substituted by I or 8AI, the MIP in the region of the mutation is more negative, and the solvation increases. This finding supports the hypothesis that the amino group at position 2 of guanine is responsible for the weakness of the spine of hydration in the vicinity of d(G·C) steps in DNA duplexes.

MD analysis of DNA triplexes. MD simulations of the triplexes studied here provide stable trajectories, irrespective of the nature of X (G, I or 8AI). The small r.m.s.d. values [~1.0 (±0.2) Å; Table 3] found between the different trajectories and their corresponding MD-averaged structures point out the stability of the three trajectories, and the larger rigidity of triplexes compared to duplexes. Helical structures are well preserved along the 3 ns trajectories, despite the fact that the selected triplex contains two contiguous d(G·C-C)$^+$ steps. We should note here that previous studies suggested that three contiguous d(G·C-C)$^+$ steps destabilize the triplex (29), while present simulations suggest that two of them do not affect the integrity of the triplex, in agreement with NMR findings (65–72).

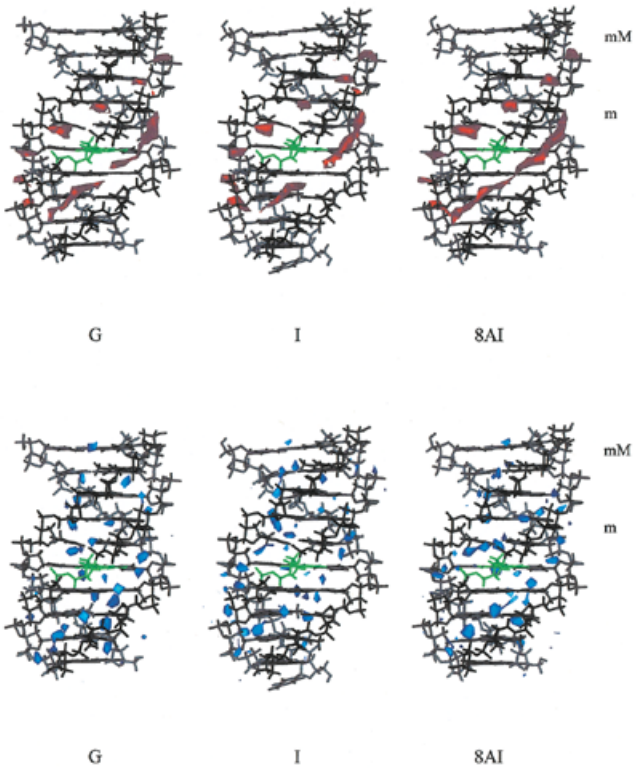

Figure 6. cMIP (top, in red) and solvation maps (bottom, in blue) for the three triplexes considered in the study (from left to right triplexes X = G, I and 8AI). Contour levels are –5.0 kcal/mol (cMIP), and apparent density of 3.0 g/ml for solvation maps. The base X is shown in green.

The three trajectories sample the same region of the configurational space, as noted in the cross r.m.s.d. values (~1 Å; Table 3). As expected (28,29), the structures sampled in the trajectories pertain to the B-family, as noted in the r.m.s.d. of 1.4 Å from the canonical B-form, and in the helical parameters shown in Table 4.

MIP and solvation analyses show clearly two regions of the triplex able to interact with small polar molecules, which are the minor (-m) groove, and the minor part of the Major groove (-mM; see 28 for nomenclature). The regions of large negative MIP are mostly located in the plane of the purine for the -m groove, and between triad planes for the -mM groove (Fig. 6). The presence of an amino group in the -m groove of the triplex slightly decreases the ability of this groove to interact with small polar solutes (Fig. 6). However, an amino group in the -mM groove does not lead to detectable alterations in the MIP and solvation maps, suggesting that the water environment in the -mM groove is not largely disrupted. This can be understood considering that the regions of most negative MIP and preferential solvation in the -mM groove are not located in the same plane as the purines (Fig. 6).

Free energy calculations. MD calculations suggest that duplex and triplex structures containing I or 8AI are stable and have similar structures. Based on the lack of large structural alterations upon replacement of G by I or 8AI, MD/TI calculations were performed to determine the changes in stability induced by the G→I→8AI substitutions

Free energy profiles for the G→I and 8AI→I mutations were smooth and without hysteresis (Fig. 7). The excellent convergence of the results is clearly noted in the small standard error

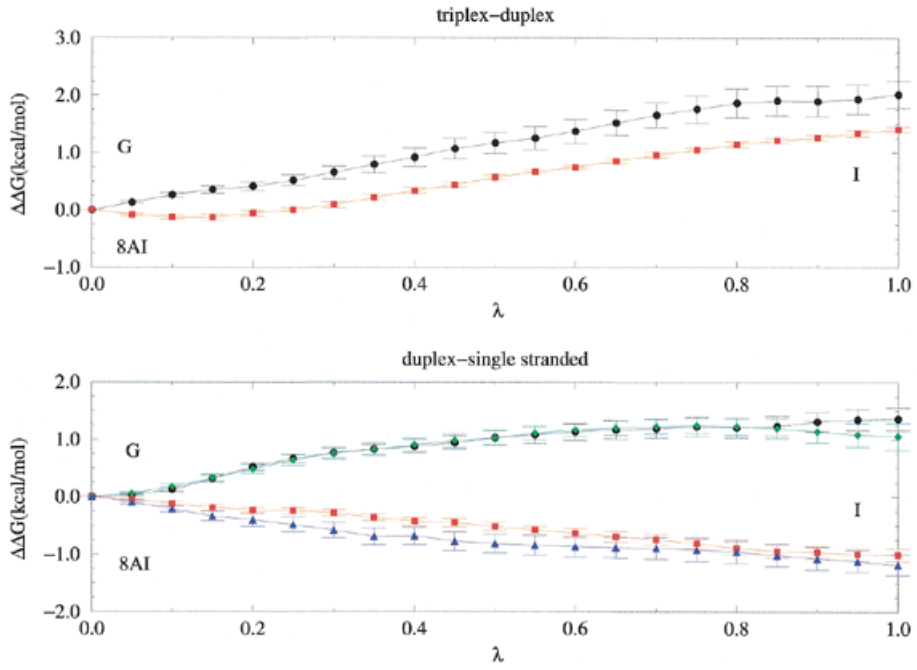

Figure 7. Difference (triplex–duplex and duplex–single-stranded) free energy profiles (in kcal/mol) obtained from the different mutations. Black, mutation G→I in the d(GAAG**X**AGGAG) duplex and triplex; green, mutation G→I in the d(GAGG**X**TCCAG) duplex; red, mutation 8AI→I in the d(GAAG**X**AGGAG) duplex and triplex; blue, mutation 8AI→I in the d(GAGG**X**TCCAG) duplex. Error bars are displayed.

Table 5. Differences in free energy of solvation (in kcal/mol) associated with the mutations G→I, and 8AI→I determined from MD/TI simulations in two single-stranded pentamers, and from SCRF calculations for the isolated bases

Mutation	SCRF	MD/TI[a]	MD/TI[b]
G→I	4.9	4.0	3.8
8AI→I	4.0	3.4	3.3

[a]Using the GGXTC sequence.
[b]Using the AGXAG sequence.

in the averages (0.1–0.2 kcal/mol). Small statistical uncertainties are then expected in our averaged free energy estimates. However, to further check the quality of MD/TI calculations, we computed the difference in free energy of solvation between pentamers containing a central G, I and 8AI. The results are compared (Table 5) with SCRF estimates obtained using our *ab initio* HF/6-31G(d)-optimized version of the MST method (see Materials and Methods). Considering that MD/TI estimates are obtained for a DNA pentamer, while the SCRF values are derived for the isolated bases, the agreement between the results calculated from the two techniques is remarkable, thus supporting the quality of the MD/TI calculations.

Changes in duplex and triplex stability due to the G→I and I→8AI mutations are given in Table 6. The G→I mutation destabilizes the duplex by ~1.2 kcal/mol, which can be explained by the loss of H-bond between the 2-amino group and the carbonyl oxygen of cytosine (Fig. 1). More intriguing is the loss of stability (~1 kcal/mol) of the duplex when I is mutated to 8AI since there is no obvious change in the H-bonding pattern (Fig. 1).

The mutation I→8AI stabilizes the triplex (with respect to the duplex) by ~1.4 kcal/mol (Table 6), which is explained by the gain of an extra Hoogsteen-like H-bond between the purine and the Hoogsteen cytosine (Fig. 1). More surprising is the strong (~2 kcal/mol) destabilization of the triplex (with respect to duplex) as a consequence of the G→I mutation (Table 6), since it does not affect the number of Hoogsteen H-bonds (Fig. 1).

Table 6. MD/TI predicted changes in stability (in kcal/mol) of duplexes and triplexes related to G→I and 8AI→I mutations

Sequence	Mutation	Folding process	$\Delta\Delta G$ (SE)
d(GAGGXTCCAG)	G→I	single→duplex	1.1 (0.2)
d(GAGGXTGGAG)	8AI→I	single→duplex	–1.2 (0.2)
d(GAAGXAGGAG)	G→I	single→duplex	1.3 (0.2)
d(GAAGXAGGAG)	8AI→I	single→duplex	–1.0 (0.1)
d(GAAGXAGGAG)	G→I	duplex→triplex	2.0 (0.2)
d(GAAGXAGGAG)	8AI→I	duplex→triplex	1.4 (0.1)

Positive values mean destabilization of the folding process.

Energy calculations. Energy calculations were performed to better understand the changes in stability due to the G→I and 8AI→I mutations in duplexes and triplexes. Energy calculations performed using the central three steps of the duplex (Table 7) suggest the following order of stability: G > I > 8AI. The presence of an amino group at position 2 of guanine slightly increases the stacking stabilization, while the same group at position 8 does not. The attachment of an amino group at position 2 of hypoxanthine largely increases (~8–9 kcal/mol) the Watson–Crick H-bonding. However, its attachment at position 8, which should be neutral from the H-bonding point of view, leads to a small, but not negligible (~1 kcal/mol) reduction of this latter intramolecular interaction. This finding, which agrees with MD/TI results, is confirmed by B3LYP/6-31G(d,p) calculations in the gas phase, which shows that the Watson–Crick dimer 8AI·C is 0.5 kcal/mol less stable than the I·C one, the canonical G·C pair being the most stable dimer (~6 kcal/mol more stable than the I·C pair).

Energy calculations for the triplexes (Table 7) show that the I→8AI mutation leads to a small destabilization of stacking interactions, and to an important (near 4 kcal/mol) gain in H-bonding interactions in the Hoogsteen side. These energy changes agree well with B3LYP calculations (see Materials and Methods), which suggest that the Hoogsteen 8AI-C$^+$ pair is

Table 7. Average interaction energies for the central triplets (underlined) obtained during the different trajectories for (top) duplexes and (bottom) triplexes

Interaction	Sequence	Interaction energies (kcal/mol)		
		d(G·C)	d(I·C)	d(8AI·C)
Stacking (total)	d(GAG<u>GXT</u>CCAG)	–26.6	–25.3	–25.5
Stacking (total)	d(GAA<u>GXA</u>GGAG)	–24.7	–23.9	–23.3
H-bond	d(GAG<u>GXT</u>CCAG)	–64.7	–56.8	–56.0
H-bond	d(GAA<u>GXA</u>GGAG)	–64.8	–57.4	–56.5
Total	d(GAG<u>GXT</u>CCAG)	–91.3	–82.1	–81.5
Total	d(GAA<u>GXA</u>GGAG)	–89.5	–81.3	–79.8
		d(G·C-C)$^+$	d(I·C-C)$^+$	d(8AI·C-C)$^+$
Stacking (total)	d(GAG<u>GXT</u>CCAG)	21.9	22.3	23.0
H-bond	d(GAG<u>GXT</u>CCAG)	–79.0	–75.2	–78.8
Total	d(GAG<u>GXT</u>CCAG)	–57.1	–52.9	–55.8

For the triplex only the interactions involving the Hoogsteen strand are represented.

Table 8. MIPp minima values (in kcal/mol) for the interaction of guanine, hypoxanthine and 8-amino-hypoxanthine with O^+ (H-bond donors) and O^- (H-bond acceptors) along H-bond directions, i.e. the straight line defined by the acceptor-donor atoms in the MD-averaged structures)

Atom	Guanine	Hypoxanthine	8-Amino-hypoxanthine
N1H	–30	–28	–26
O6 (N1)	–30	–29	–30
O6 (N7)	–52	–49	–51
N7	–55	–51	–51

For O6 both the N1 and N7 H-bond directions were considered.

Table 9. Melting temperatures (°C) of 8AI and I duplexes[a]

	Y = 8AI	Y = I	Y = G
X = C	56	58	61
X = A	53	55	54
X = G	54	53	57
X = T	53	52	52

0.15 M NaCl, 50 mM Tris–HCl buffer pH 7.5.
[a]5′-d(TAGAGG**X**TCCATTGC)-3′:3′-d(ATCTCC**Y**AGGTAACG)-5′

4 kcal/mol more stable than the I-C^+ one. It should be noted that the stabilization obtained by the extra Hoogsteen H-bond is then smaller than that expected for a normal H-bond, which is easily explained considering that the presence of the 8-amino group is not favored by the positive charge of the Hoogsteen cytosine.

Classical energy calculations (Table 7) suggest that the Hoogsteen G-C^+ pair is very stable due to the large strength of the two Hoogsteen H-bonds. In fact, classical calculations suggest similar stability for the G-C^+ and 8AG-C^+ pairs, not fully supported by B3LYP calculations, which suggest that the 8AI-C^+ dimer is slightly more stable (~2 kcal/mol) than the G-C^+ one. This discrepancy might overestimate the stability of triplexes containing d(8AI·C-C)$^+$ triads with respect to those containing d(G·C-C)$^+$ triads. However, there is general agreement between classical and QM calculations. It is then clear that (i) the Hoogsteen H-bond involving the 8-amino group is not very strong and (ii) the presence of the 2-amino group enhances the stability of Hoogsteen H-bonds.

Both MD/TI and energy calculations show that the presence of an amino group at position 2 of I strongly stabilizes duplexes relative to single-stranded oligonucleotides, and triplexes relative to duplexes. On the contrary, the presence of an amino group at position 8 stabilizes the triplex relative to duplex, but destabilizes the duplex relative to single-stranded oligonucleotides. As noted above, simple counting of H-bond interactions explains the effect of the 2-amino group in the Watson–Crick pairing, and that of the 8-amino group in the Hoogsteen side. However, they do not explain the effects of the 2-amino in the Hoogsteen pairing, and of the 8-amino in the Watson–Crick interaction.

To investigate the effect of distant amino groups on H-bond interactions we performed QM-MIPp calculations in the Watson–Crick and Hoogsteen regions for G, I and 8AI. The results in Table 8 show that the presence of an amino group at position 2 of guanine increases the negative potential at N7 and O6 (N7-side) lone pairs (3–4 kcal/mol from the results in Table 8), which explains the larger stability of Hoogsteen H-bonds in guanine compared to hypoxanthine. The amino group at position 8 of hypoxanthine increases the electron density in the Watson–Crick region. As a result, the O6 (N1-side) becomes a slightly better H-bond acceptor, but the N1-H group becomes a poorer H-bond donor. The overall effect is a small loss (~1 kcal/mol) of H-bonding properties in the Watson–Crick side (MIPp minima located without restrains show a loss of 1.5 kcal/mol in binding due to the 8-NH_2 group).

Melting experiments

A homologous series of pentadecanucleotide duplexes with sequence 5′-d(TAGAGG**X**TCCATTGC)-3′/3′-d(ATCTCC**Y**-AGGTAACG)-5′ containing 8AI, I and G in the center, were prepared by annealing equimolar mixtures of the corresponding strands. Thermal melting curves were determined spectrophotometrically by measuring UV absorbance at 260 nm as function of temperature. Melting temperatures are shown in Table 9. The most stable base pair in the three series was the Y·C base pair, G·C being the most stable (61°C), followed by I·C (58°C) and 8AI·C (56°C). These results are in good agreement with the MD/TI predictions (Tables 6 and 7) and TPIp values (Table 9). The decrease in T_m by the mutation of G→I is due to the loss of one hydrogen bond, and the decrease in T_m by the mutation of I→8AI is due (see above) to the change on the electronic distribution induced by the addition of the 8-amino group.

Table 10. Melting temperatures (°C) for the triplex[a] h_{26}: s_{11} containing 8AI (I^N) and I (1M NaCl, 100 mM sodium phosphate/citric acid buffer)

Modifications	Sequence purine strand (5′→3′)	pH 5.5	pH 6.0	pH 6.5	pH 7.0
No	GAAGGAGGAGA	40, 82	20, 82	–, 82	–, 82
1 I	GAAGIAGGAGA	32, 75	18, 74	–, 74	–, 75
3 I	GAAGIAIGAIA	26, 64	–, 64	–, 64	–, 64
1 I^N	GAAGI^{N}AGGAGA	43, 73	29, 73	18, 73	–, 73
3 I^N	GAAGI^{N}AI^{N}GAI^{N}A	55[b]	40, 55	27, 55	18, 55

[a]h_{26} 5′-d(GAAGGAGGAGATTTTTCTCCTCCTTC)-3′:s_{11} 5′-d(CTTCCTCCTCT)-3′.
[b]Only one transition was observed with 25% hyperchromicity suggesting a triplex to random coil transition.

Table 11. Melting enthalpies (kcal/mol) data of pentadecamer duplex[a] formation determined from the UV melting curves

	Y = 8AI	Y = I	Y = G
X = C	–98	–96	–105
X = A	–99	–94	–86
X = G	–88	–90	–103
X = T	–87	–84	–92

Buffer conditions 50 mM Tris–HCl, 0.15 M NaCl pH 7.5
[a]5′-d(TAGAGG**X**TCCATTGC)-3′:3′-d(ATCTCC**Y**AGGTAACG)-5′.

The stability of the duplexes containing 8AI mismatches with A, G and T are similar to the duplexes containing I mismatches. Because of the 2°C decrease in the melting temperature of the duplex carrying an 8AI·C base pair, the differences in stability between the duplexes containing 8AI is only 3°C compared with the 6°C of difference between duplexes containing I (Table 9). This tendency to similar melting temperatures for the different mismatches may be beneficial in the design of degenerate primers or probes, taking into account that I is already being used for this purpose. The spread on melting temperatures when 8AI was paired with natural bases (3°C) compares well with the spread on T_m reported for other candidates for universal base [3°C for 3-nitropyrrole (73,56) and 5-nitroindole (56); 2°C for acyclic 5-nitroindazole (74) and 8-aza-7-deaza-dA (75)]. Also the differences in melting temperatures of 8AI mismatches relative to fully matched duplexes are similar to the differences reported for 3-nitropyrrole (56), 5-nitroindole (56) and acyclic 5-nitroindazole (74). Only 8-aza-7-deaza-dA seems to have better binding properties (75) than 8AI. The 'universal-base' characteristics of 8AI are of special interest considering the ability of this molecule to form stable triplexes (see below). This point will deserve detailed study in the future.

The effect of 8AI and I on triple helices was investigated using a triple helix model formed by a self-complementary hairpin of 26 bases (h_{26}, 5′-GAA GGA GGA GAT TTT TCT CCT CCT TC-3′) and an all-pyrimidine single-strand oligonucleotide (s_{11}, 5′-CTT CCT CCT CT-3′) described previously (76). For this purpose h_{26} derivatives carrying one or three I^N or I residues were synthesized and triplexes were prepared by annealing equimolar mixtures of h_{26} and s_{11} oligonucleotides. Thermal melting curves were determined spectrophotometrically by measuring UV absorbance at 260 nm as function of temperature at pH 5.5–7.0. In most cases two transitions were observed that were assigned to triplex to duplex transition (first melt) and duplex to random coil transition (second melt) (76). Melting temperatures are shown in Table 10. As described previously, the stability of the triplex increases at acidic pH due to protonation of C while the stability of the duplex remains stable in the pH range 5.5–7.0 (76) Sequential replacement of G by I results in a destabilization of both duplex and triplex. The decrease in duplex stability is easily explained by the loss of one H-bond, but the decrease in triplex stability is not evident. As described above, the decrease in triplex stability may be explained by electronic redistribution after the removal of the amino group at position 2 that affects the N-7 and O-6 responsible of the Hoogsteen pairing (Table 7).

Sequential replacement of I by 8AI results in a destabilization of the duplex and in a stabilization of the triplex, as predicted by theoretical calculations. Again the presence of the amino group at position 8 results in an electronic distribution that does not favor Watson–Crick pairing and stabilizes the triplex by the gain of an H-bond and by a new electronic distribution that favors Hoogsteen pairing (Table 7).

Thermodynamic analysis

The shape of the melting curve defined by the differential curves was used for the determination of melting enthalpies in pentadecamer duplexes (these enthalpies contain both solute and solvent terms and cannot be directly compared with energy calculations reported in the theoretical part of the manuscript). As seen in Table 11, the substitution of G by I leads to a decrease (in absolute value) of 9 kcal/mol in $\Delta H°$, while the substitution of G by 8AI leads to a decrease (in absolute values) of 7 kcal/mol in melting enthalpy. Hypoxanthine (I and 8AI) base pairs with A have an increased melting enthalpy than G.A base pair, while the melting enthalpies of the G.T and G.G base pairs are bigger than the corresponding I and 8AI base pairs.

In the case of the triplexes we measured the thermodynamic parameters from concentration-dependent melting curves as described for 8-amino-G triplexes (13). Table 12 shows the thermodynamic parameters (ΔH_t, ΔS_t and ΔG_t) for triplexes having G, I and 8AI at the central position. The substitution of G by I provoked a decrease of 1.4 kcal/mol in the stability of the triplex, in good agreement with theoretical calculations. As expected from theoretical calculations, the presence of the 8-amino group stabilises of the triplex, as noted in an increase (in absolute terms) of 2.6 kcal/mol in the stability of the triplex upon I→8AI mutation. These data confirm the importance of the amino groups at the 2 and 8 positions of inosine derivatives on the stability of triplexes.

Table 12. Thermodynamic parameters for transition h_{26}: $s_{11} = h_{26} + s_{11}$ in 100 mM sodium phosphate/citric acid, 1 M NaCl pH 5.5 from the slope of the plot $1/T_m$ versus lnC[a]

Modifications	Sequence purine strand (5′→3′)	ΔH_t (kcal/mol)	ΔS_t (cal/mol K)	ΔG_t (kcal/mol)
None	GAAGGAGGAGA	–67	–189	–10.8
1 I	GAAGIAGGAGA	–88	–263	–9.4
1 I^N	GAAGI^{N}AGGAGA	–74	–207	–12.0

h_{26}: 5′-GAAGGAGGAGATTTTTCTCCTCCTTC-3′, s_{11}: 5′-CTTCCTCCTCT-3′.
[a]ΔH_t, ΔS_t are given as round numbers, ΔG_t is calculated at 25°C, with the assumption that ΔH_t and ΔS_t do not depend on temperature; analysis has been carried out using melting temperatures obtained from denaturation curves.

CONCLUSIONS

A combination of state of the art theoretical and experimental techniques have allowed us to investigate the effect that amino groups added on positions 2 and 8 of the hypoxanthine moiety has on the structure, flexibility and stability of duplexes and triplexes. A new pseudobase 2-aminohypoxanthine is theoretically designed, synthesized and experimentally tested, showing interesting properties as a unique 'universal base' and triplex-stabilizing pseudobase. Good agreement is found between theoretical prediction and experimental results, demonstrating the predictive power of high level theoretical calculations. Overall, our results demonstrate that the stability of nucleic acid structures is a subtle balance of different terms, and that simple counting of elemental interactions can lead to erroneous conclusions.

ACKNOWLEDGEMENTS

We are grateful to Prof. J.Tomasi for providing us with his original code of the PCM model, which was modified by us to carry out the MST calculations. We also acknowledge the Dirección General de Investigación Científica y Técnica (grants PB98-1222, PM99-0046 and BQU2000-0649) and the Generalitat de Catalunya (2000-SGR-0018) for financial support, and the Centre de Supercomputació de Catalunya for computational facilities.

REFERENCES

1. Saenger,W. (1984) *Principles of Nucleic Acid Structure*. Springer-Verlag. New York, NY.
2. Soyfer,V.N. and Potaman,V.N. (1996) *Triple Helical Nucleic Acids*. Springer-Verlag. New York, NY.
3. Sun,J.S. and Hélène,C. (1993) Oligonucleotide-directed triple-helix formation. *Curr. Opin. Struct. Biol.*, **6**, 327–333.
4. Laughlan,G., Murchie,A.I., Norman,M.H., Moody,P.C., Lilley,D.M. and Luisi,B. (1994) The high-resolution crystal structure of a parallel-stranded guanine tetraplex. *Science*, **265**, 520–524.
5. Williamson,J.R. (1994) G-Quartet structures in telomeric DNA. *Annu. Rev. Biophys. Biomol. Struct.*, **23**, 703–730.
6. Lee,J.S., Johnson,D.A. and Moogan,A.R. (1979) Complexes formed by (pyrimidine)n. (purine)n DNAs on lowering the pH are three-stranded. *Nucleic Acids Res.*, **6**, 3073–3091.
7. Völker,J. and Klump,H.H. (1994) Electrostatic effects in DNA triple helices. *Biochemistry*, **33**, 13502–13508.
8. Wittung,P., Nielsen,P. and Norden,B. (1997) Extended DNA-recognition repertoire of peptide nucleic acid (PNA): PNA–dsDNA triplex formed with cytosine-rich homopyrimidine PNA. *Biochemistry*, **36**, 7973–7979.
9. Chollet,A. and Kawashima,E. (1988) DNA containing the base analogue 2-aminoadenine: preparation, use as hybridization probes and cleavage by restriction endonucleases. *Nucleic Acids Res.*, **16**, 305–317.
10. Güimil-García,R., Ferrer,E., Macías,M.J., Eritja,R. and Orozco,M. (1999) Theoretical calculations, synthesis and base pairing properties of oligonucleotides containing 8-amino-2′-deoxyadenosine. *Nucleic Acids Res.*, **27**, 1991–1998.
11. Eritja,R., Ferrer,E., Güimil-García,R. and Orozco,M. (1999) Modified oligonucleotides with triple-helix stabilization properties. *Nucl. Nucl.*, **18**, 1619–1621.
12. Güimil-García,R., Bachi,A., Eritja,R., Luque,F.J. and Orozco,M. (1998) Triple helix stabilization properties of oligonucleotides containing 8-amino-2′-deoxyguanosine. *Bioorg. Med. Chem. Lett.*, **8**, 3011–3016.
13. Soliva,R., Güimil-García,R., Blas,J.R., Eritja,R., Asensio,J.L., González,C., Luque,F.J. and Orozco,M. (2000) DNA-triplex stabilizing properties of 8-aminoguanine. *Nucleic Acids Res.*, **28**, 4531–4539.
14. Schweitzer,B.A. and Kool,E.T. (1995) Hydrophobic, non-hydrogen-bonding bases and base pairs in DNA. *J. Am. Chem. Soc.*, **117**, 1863–1972.
15. Moran,S., Ren,R. and Kool,E.T. (1997) A thymidine triphosphate shape analog lacking Watson–Crick pairing ability is replicated with high sequence selectivity. *Proc. Natl Acad. Sci. USA*, **94**, 10506–10511.
16. Matray,T.J. and Kool,E.T. (1998) Selective and stable DNA base pairing without hydrogen. *J. Am. Chem. Soc.*, **120**, 6191–6192.
17. Cubero,E., Sherer,E.C., Luque,F.J., Orozco,M. and Laughton,C.A. (1999) Observation of spontaneous base pair breathing events in the molecular dynamics simulation of a difluorotoluene-containing DNA oligonucleotide. *J. Am. Chem. Soc.*, **121**, 8653–8654.
18. Cubero,E., Laughton,C.A., Luque,F.J. and Orozco,M. (2000) Molecular dynamics study of oligonucleotides containing difluorotoluene. *J. Am. Chem. Soc.*, **122**, 6891–6899.
19. Ohtsuka,E., Matsuki,S., Ikehara,M., Takahashi,T. and Matsubara,K.J. (1985) An alternative approach to deoxyoligonucleotides as hybridization probes by insertion of deoxyinosine at ambiguous codon positions. *J. Biol. Chem.*, **260**, 2605–2608.
20. Martin,F.H., Castro,M.M., Aboul-ela,F. and Tinoco,I. (1985) Base pairing involving deoxyinosine: implications for probe design. *Nucleic Acids Res.*, **13**, 8927–8938.
21. Loakes,D., Brown,D.M., Linde,S. and Hill,F. (1995) 3-Nitropyrrole and 5-nitroindole as universal bases in primers for DNA sequencing and PCR. *Nucleic Acids Res.*, **23**, 2361–2366.
22. Thiele,D. and Guschlbauer,W. (1969) Protonated polynucleotides. VII. Thermal transitions between different complexes of polyinosinic acid and polycytidylic acid in an acid medium. *Biopolymers*, **8**, 361–378.
23. Gargallo,R., Tauler,R. and Izquierdo-Ridorsa,A. (1997) Acid-base and copper (II) complexation equilibria of poly(inosinic)-poly(cytidylic). *Biopolymers*, **42**, 271–283.
24. Letai,A.G., Palladino,M.A., Fromm,E., Rizzo,V. and Fresco,J.R. (1988) Specificity in formation of triple-stranded nucleic acid helical complexes: studies with agarose-linked polyribonucleotide affinity columns. *Biochemistry*, **27**, 9108–9112.
25. Mills,M., Völker,J. and Klump,H.H. (1996) Triple helical structures involving inosine: there is a penalty for promiscuity. *Biochemistry*, **35**, 13338–13344.
26. Hogeland,J.S. and Weller,D.D. (1993) Investigations of oligodeoxyinosine for triple helix formation. *Antisense Res. Dev.*, **3**, 285–290.
27. Arnott,S. and Hukins,D.W.L. (1972) Optimized parameters for A-DNA and B-DNA. *Biochem. Biophys. Res. Commun.*, **47**, 1504–1509.
28. Shields,G., Laughton,C.A. and Orozco,M. (1997) Molecular dynamics simulations of the d(T·A·T) triple helix. *J. Am. Chem. Soc.*, **119**, 7463–7469.
29. Soliva,R., Laughton,C.A., Luque,F.J. and Orozco,M. (1998) Molecular dynamics simulations in aqueous solution of triple helices containing d(G·C·C) trios. *J. Am. Chem. Soc.*, **120**, 11226–11233.
30. Shields,G., Laughton,C.A. and Orozco,M. (1998) Molecular dynamics simulation of a PNA·DNA·PNA triple helix in aqueous solution. *J. Am. Chem. Soc.*, **120**, 5895–5904.
31. Essmann,U., Perera,L., Berkowitz,M.L., Darden,T., Lee,H. and Pedersen,L.G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.*, **103**, 8577–8593.
32. Darden,T.A., York,D. and Pedersen,L. (1993) Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.*, **98**, 10089–10092.
33. Ryckaert,J.P., Ciccote,G. and Berendsen,J.C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys.*, **23**, 327–341.
34. 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. and Kollman,P.A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, **117**, 5179–5197.
35. Cheatham,T.E., Cieplak,P. and Kollman,P.A. (1999) A modified version of the Cornell *et al.* force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn.*, **16**, 845–862.
36. Jorgensen,W.L., Chandrasekhar,J., Madura,J.D., Impey,R. and Klein,M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926–935.
37. Bayly,C.I., Cieplak,P., Cornell,W.D. and Kollman,P.A. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.*, **97**, 10269–10280.

38. McQuarri,D.A. (1976) *Statistical Mechanics.* Harper and Row, New York, NY.
39. Lee,C., Yang,W. and Parr,R.G. (1998) Development of the Colle–Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, **37**, 785–789.
40. Boys,S.F. and Bernardi,F. (1970) The calculation of small molecules interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.*, **19**, 553–559.
41. Luque,F.J., Bachs,M. and Orozco,M. (1994) An optimized AM1/MST method for the MST-SCRF representation of solvated systems. *J. Comp. Chem.*, **15**, 847–857.
42. Orozco,M., Bachs,M. and Luque,F.J. (1995) Development of optimized MST/SCRF methods for semiempirical calculations: the MNDO and PM3 Hamiltonians. *J. Comp. Chem.*, **16**, 563–575.
43. Bachs,M., Luque,F.J. and Orozco,M. (1994) Optimization of solute cavities and van der Waals parameters in ab initio MST-SCRF calculations of neutral molecules. *J. Comp. Chem.*, **15**, 446–454.
44. Miertus,S., Scrocco,E. and Tomasi,J. (1981) Electrostatic interaction of a solute with a continuum. A direct utilization of ab initio molecular potentials for the prevision of solvent effects. *Chem. Phys.*, **55**, 117–129.
45. Miertus,S. and Tomasi,J. (1982) Approximate evaluations of the electrostatic free energy and internal energy changes in solution processes. *Chem. Phys.*, **65**, 239–245.
46. Tomasi,J. and Persico,M. (1994) Molecular interactions in solution: an overview of methods based on continuous distribution of the solvent. *Chem. Rev.*, **94**, 2027–2094.
47. Orozco,M. and Luque,F.J. (1993) Molecular interaction potential: a new tool for the theoretical study of molecular reactivity. *J. Comp. Chem.*, **14**, 587–602.
48. Luque,F.J. and Orozco,M. (1998) Polarization effects in generalized molecular interaction potential: New Hamiltonian for reactivity studies and mixed QM/MM calculations. *J. Comp. Chem.*, **19**, 866–881.
49. Hernández,B., Luque,F.J. and Orozco,M. (1999) Parametrization of the GMIPp for the study of stacking interactions. *J. Comp. Chem.*, **20**, 937–946.
50. Cubero,E., Luque,F.J. and Orozco,M. (1998) Is polarization important in cation-pi interactions? *Proc. Natl Acad. Sci. USA*, **95**, 5976–5980.
51. Case,D.A., Pearlman,D.A., Caldwell,J.W., Cheatham,T.E., Ross,W.S., Simmerling,C.L., Darden,T.A., Merz,K.M., Stanton,R.V., Cheng,A.L. *et al.* (1997) *AMBER 5*. University of California, San Francisco, CA.
52. Lavery,R. and Sklena,J. (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.*, **6**, 63–91.
53. Frisch,M.J., Trucks,G.W., Schlegel,H.B., Gill,P.M.W., Johnson,B.G., Robb,M.A., Cheeseman,J.R., Keith,T.A., Petersson,G.A., Montgomery,J.A. *et al.* (1995) *GAUSSIAN 94* (Rev. A.1). Gaussian Inc., Pittsburgh, PA.
54. Peterson,M. and Poirier,R. (1980) *MonsterGauss*. Department of Chemistry. University of Toronto, Toronto, Canada [modified (1987) by Cammi,R., Bonaccorsi,R. and Tomasi,J., University of Pisa, Pisa, Italy; further modified (1995) by Luque,F.J. and Orozco,M., University of Barcelona, Barcelona, Spain].
55. Luque,F.J. and Orozco,M. (2000) *MOPETE Computer Program*. University of Barcelona, Barcelona, Spain.
56. Loakes,D. and Brown,D.M. (1994) 5-Nitroindole as a universal base analogue. *Nucleic Acids Res.*, **22**, 4039–4043.
57. Long,R.A., Robins,R.K. and Townsend,L.B. (1967) Purine nucleosides XV. The synthesis of 8-amino- and 8-substituted amino purine nucleosides. *J. Org. Chem.*, **32**, 2751–2756.
58. Schulhof,J.C., Molko,D. and Teoule,R. (1987) The final deprotection step in oligonucleotide synthesis is reduced to a mild and rapid ammonia treatment by using labile base-protecting groups. *Nucleic Acids Res.*, **15**, 397–416.
59. Jones,R.A. (1984) Preparation of protected deoxyribonucleosides. In Gait,M.J. (ed.), *Oligonucleotide Synthesis, A Practical Approach.* IRL Press, Oxford, UK, pp. 23–34.
60. Soliva,R., Luque,F.J., Alhambra,C. and Orozco,M. (1999) Role of sugar re-puckering in the transition of A and B forms of DNA in solution. A molecular dynamics study. *J. Biomol. Struct. Dyn.*, **17**, 89–99.
61. Hernández,B., Soliva,R., Luque,F.J. and Orozco,M. (2000) Misincorporation of 2′-deoxyoxanosine into DNA: a molecular basis for NO-induced mutagenesis derived from theoretical calculations. *Nucleic Acids Res.*, **28**, 4873–4883.
62. Dickerson,R.E. (1992) DNA structure from A to Z. *Methods Enzymol.*, **211**, 67–127.
63. Drew,H.R., Wing,R.M., Takana,S., Broka,C., Tanaka,S., Itakura,K. and Dickerson,R.E. (1981) Structure of a B-DNA dodecamer: Conformation and dynamics. *Proc. Natl Acad. Sci. USA*, **78**, 2179–2183.
64. Shui,X., McFail-Isom,L., Hu,G.G. and Williams,L.D. (1998) The B-DNA dodecamer at high resolution reveals a spine of water on sodium. *Biochemistry*, **37**, 8341–8355.
65. Rajagopal,P. and Feigon,J. (1989) NMR studies of triple-strand formation from the homopurine–homopyridine deoxyribonucleotides $d(GA)_4$ and $d(TC)_4$. *Biochemistry*, **28**, 7859–7870.
66. Rajagopal,P. and Feigon,J. (1989) Measurement of γ-rays from cold fusion. *Nature*, **339**, 667–669.
67. Radhakrishnan,I. and Patel,D.J. (1994) DNA triplexes: solution structures, hydration sites, energetics, interactions, and function. *Biochemistry*, **33**, 11405–11416.
68. Radhakrishnan,I. and Patel,D.J. (1994) Hydration sites in purine.purine.pyrimidine and pyrimidine.purine.pyrimidine DNA triplexes in aqueous solution. *Structure*, **2**, 395–405.
69. Wang,E., Koshlap,K.M., Gilllespie,P., Dervan,P.B. and Feigon,J. (1996) Solution structure of a pyrimidine·purine·pyrimidine triplex containing the sequence-specific intercalating non-natural base D3. *J. Mol. Biol.*, **257**, 1052–1069.
70. Asensio,J.L., Dhesai,J., Bergquist,S., Brown,T. and Lane,A.N. (1998) The contribution of cytosine protonation to the stability of parallel DNA triple helices. *J. Mol. Biol.*, **275**, 811–822.
71. Asensio,J.L., Brown,T. and Lane,A.N. (1999) Solution conformation of a parallel DNA triple helix with 5‘ and 3‘ triplex-duplex junctions. *Struct. Fold Des.*, **7**, 1–11.
72. Asensio,J.L., Brown,T. and Lane,A.N. (1998) Comparison of the solution structures of intramolecular DNA triple helices containing adjacent and non-adjacent CG.C+ triplets. *Nucleic Acids Res.*, **26**, 3677–3686.
73. Nichols,R., Andrews,P.C., Zhang,P. and Bergstrom,D.E. (1994) A universal nucleoside for use at ambiguous sites in DNA primers. *Nature*, **369**, 492–493.
74. Van Aerschot,A., Rozenski,J., Loakes,D., Pillet,N., Schepers,G. and Herdewijn,P. (1995) An acyclic 5-nitroindazole nucleoside analogue as ambiguous nucleoside. *Nucleic Acids Res.*, **23**, 4363–4370.
75. Seela,F. and Debelak,H. (2000) The N^8-(2′-deoxyribofuranoside) of 8-aza-7-deazaadenine: a universal nucleoside forming specific hydrogen bonds with the four canonical DNA constituents. *Nucleic Acids Res.*, **28**, 3224–3232.
76. Xodo,L.E., Manzini,G., Quadrifolio,F., van der Marel,G.A. and van Boom,J.H. (1991) Effect of 5-methylcytosine on the stability of triple-stranded DNA—a thermodynamic study. *Nucleic Acids Res.*, **19**, 5625–5631.