Nucleic Acids Research, 2000, Vol. 28, No. 22 **4531–4539**

DNA-triplex stabilizing properties of 8-aminoguanine

Robert Soliva, Ramón Güimil García[1], J. Ramón Blas, Ramon Eritja[1,*], Juan Luis Asensio[2], Carlos González[3], F. Javier Luque[4] and Modesto Orozco

Departament de Bioquímica i Biologia Molecular, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1, Barcelona 08028, Spain, [1]European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, D-69117 Heidelberg, Germany, [2]Instituto de Estructura de la Materia, C.S.I.C., C/ Serrano 119, Madrid 28006, Spain, [3]Instituto de Química Orgánica, C.S.I.C., C/ Juan de la Cierva 3, Madrid 28006, Spain and [4]Departament de Fisicoquímica, Facultat de Farmàcia, Universitat de Barcelona, Avgda Diagonal sn, Barcelona 08028, Spain

Received July 7, 2000; Revised and Accepted September 25, 2000

ABSTRACT

A DNA-triplex stabilizing purine (8-aminoguanine) is designed based on molecular modeling and synthesized. The substitution of guanine by 8-aminoguanine largely stabilizes the triplex both at neutral and acidic pH, as suggested by molecular dynamics and thermodynamic integration calculations, and demonstrated by melting experiments. NMR experiments confirm the triplex-stabilizing properties of 8-aminoguanine and demonstrate that few changes are found in the structure of the triplex due to the presence of this modified base.

INTRODUCTION

DNA triplexes based on the pyrimidine motif (R·Y-Y) are formed by the interaction of a pyrimidine strand (the triplex-forming oligonucleotide) with the purine strand of a DNA duplex via H-bond interactions in the major groove (1 and references therein). In this pyrimidine motif two triads are found (Fig. 1): (i) d(A·T-T), where a thymine recognizes an adenine (Fig. 1), and (ii) d(G·C-C), where the purine is recognized by cytosine. In the two cases the triplex-forming oligonucleotide binds to the DNA duplex through the major groove and the recognition involves the formation of Hoogsteen H-bonds.

The formation of two H-bonds between guanine and the Hoogsteen cytosine in d(G·C-C) is typically achieved by protonation of the Hoogsteen cytosines, leading to d(G·C-C)$^+$ triads (2–13). The resulting triplexes are then strongly pH-dependent and are often unstable at neutral pH (2–4). It has been suggested that for homoguanine tracks (three or more contiguous guanines) the protonation of all the Hoogsteen cytosines might lead to a strong repulsive interaction, which could unfold the triplex (14,15). Molecular dynamics, calorimetric and NMR data (14–17) suggest that, at least for homoguanine tracks, a minor percentage of neutral cytosines coexists with the protonated cytosines in the triplex. This opens the possibility to have alternative imino and wobble pairings for the d(G·C-C) triad (Fig. 1; 14,15)

The strength of the d(A·T-T) and d(G·C-C) triads is extremely important for the global stability of the triplex. This explains the interest in developing modified bases that enhance the strength of R·Y-Y triads, which might then play a role as anti-gene drugs or diagnostic tools (1,18–37). Molecular dynamics and thermodynamic integration calculations suggested that 8-aminoadenine might be a triplex stabilizing molecule (31–33). Synthesis, incorporation into DNA and melting experiments demonstrated the strong triplex-stabilizing properties of 8-aminoadenine (31–33). It was suggested that the formation of an extra Hoogsteen H-bond (N8→O2), and the placement of the 8-amino group in the narrow and polar minor part of the major groove (mM groove) of the triplex were responsible for the stabilizing effect of 8-aminoadenine (Fig. 1; 31,38).

In this paper we present a theoretical and experimental study on the effect of guanine→8-aminoguanine substitution on the stability of DNA triplexes (from hereon 8-aminoguanine is abbreviated to 8aG except when it is within a nucleotide sequence—in this case the position of it is indicated by **1** to avoid confusion with adenine or guanine). This molecule was previously suggested as a triplex-forming molecule by other authors (34,35). Unfortunately, no clear conclusions were obtained, due to the high stability of the hemiprotonated G·G self-structure, which precluded the formation of stable complexes beyond the level of monophosphate. Here we have explored theoretically the effect of the G→8aG mutation for all possibilities for Hoogsteen-like G-C recognition: d(G·C-C)$^+$, d(G·C-C)$_i$ and d(G·C-C)$_w$, which might appear in different triplex sequences in different ranges of pH (Fig. 1). Calculations suggest that the G→8aG substitution leads to a remarkable stabilization of the triplex at all range of pH values. This is confirmed and quantified experimentally from spectroscopic techniques after synthesis of the molecule and incorporation into DNA, extending preliminary results reported recently by our group (33).

*To whom correspondence should be addressed. Modesto Orozco, 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
Present address:
Ramon Eritja, Instituto de Biología Molecular de Barcelona, C.S.I.C., Jordi Girona 18-26, Barcelona 08034, Spain

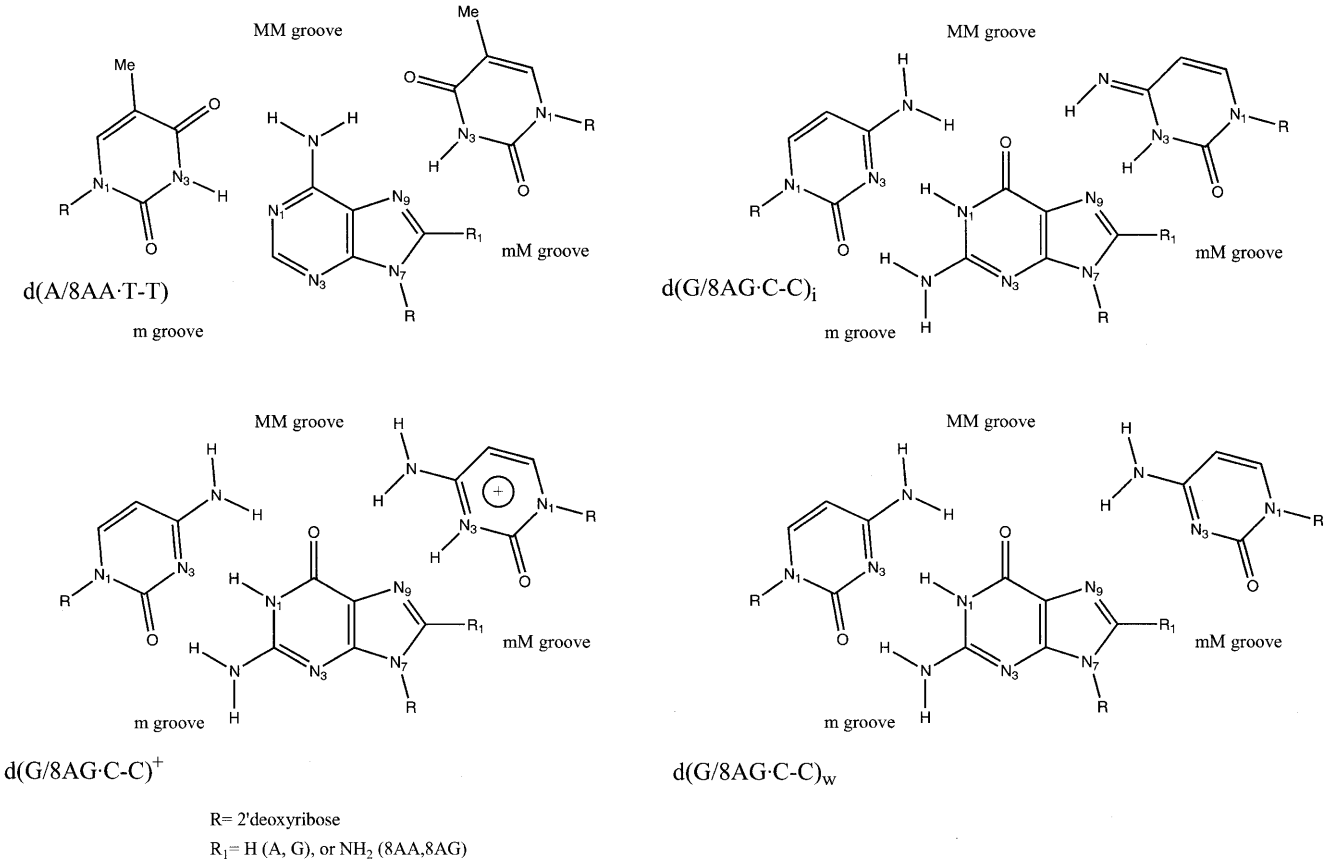


Figure 1. Schematic representation of d(A/8AA·T-T) and d(G/8aG·C-C) triads. The three grooves: major part of the major groove (MM), minor part of the major groove (mM), and minor groove (m) are displayed.

MATERIAL AND METHODS

Theoretical calculations

Triplex structures of sequence d(GAGAGAGAGA) were generated using our previously equilibrated structure (14). Hoogsteen cytosines at positions 1, 3, 7 and 9 were protonated, and guanine at position 5 was replaced by 8aG. Three different models were then generated corresponding to the protonated d(8aG·C-C)$^+$, the wobble d(8aG·C-C)$_i$ and the imino d(8aG·C-C)$_w$ triads. In all cases the sequences flanking the mutation site were kept unaltered. All these models were surrounded by sodium ions to maintain neutrality, and 4141 water molecules. The hydrated systems were optimized and equilibrated using our standard protocol (38,39). The final equilibrated structures were then used in 1.5 ns of molecular dynamics (MD) trajectory at constant pressure (1 atm) and temperature (298K).

Models of DNA duplex d(GAGAGAGAGA) were generated using Arnott's fiber parameters (40). Structures were then surrounded by counterions to maintain neutrality, and hydrated by 4391 water molecules. The duplex was then optimized and equilibrated using our standard protocol (38,39). The final structures were then used in 0.5 ns of MD simulation at constant pressure (1 atm) and temperature (298K).

Periodic boundary conditions and the PME method (41,42) were used to account for long range effects. SHAKE (43) was used to constrain all chemical bonds, which allowed us to use a 2 fs time step for integration of Newton equations. AMBER-99 force-field parameters were used in conjunction with previously optimized parameters for protonated and imino cytosine (14), as well as with HF/6-31G(d)-RESP-derived atomic charges (44) for 8aG (parameters are available from the authors upon request).

The three final structures of the duplexes and triplexes obtained in the MD simulation were then used as starting points for thermodynamic integration (TI) calculations (45), where the 8aG molecule was mutated into guanine both in the duplex and the three triplexes [containing the d(G·C-C)$^+$, d(G·C-C)$_i$, and d(G·C-C)$_w$ trios]. The difference between the reversible work necessary to perform these mutations gives the change in stability of the triplex formation (duplex→triplex process) due to the substitution of G by 8aG (Fig. 2). Mutations were performed using 21 windows consisting of 10 ps of equilibration and 10 ps of collection, leading to MD/TI simulation of 420 ps. All mutations were repeated using the same simulation protocol, but starting from different conformations and velocities (those obtained 100 ps before the end of the MD). All technical details of MD/TI simulations are identical to those of normal MD simulations.

MD-averaged structures were determined using the last 0.4 (duplex) or 1 ns (triplexes) of the MD trajectories. All MD and MD/TI calculations were done using the AMBER-5.1 computer program. Analysis of the trajectories was carried out using

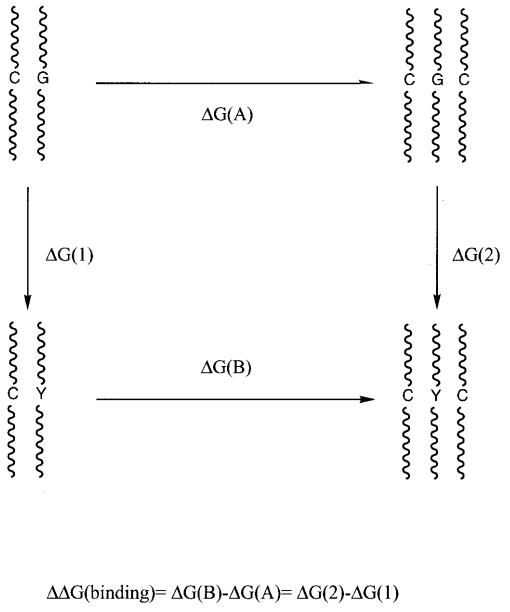


Figure 2. Thermodynamic cycle used to compute the stabilization of triplex structures due to the G→8aG mutation. Y stands for 8aG.

CURVES, analysis modules in AMBER, as well as 'in-house' programs. All calculations were performed in the Centre de Supercomputació de Catalunya (CESCA) as well as at the workstations in our laboratories.

Chemicals

Protected phosphoramidites and DNA synthesis reagents were purchased from Perkin Elmer-Applied Biosystems (USA). Nucleosides were purchased from Pharma-Waldorf (Germany). Anhydrous solvents were obtained from SDS (France). The rest of the chemicals were from Fluka (Switzerland) and Aldrich (USA) and were used without further purification. Starting from 2′-deoxyguanosine, 8-amino-2′-deoxyguanosine was prepared as previously described (46). Essentially the preparation of this nucleoside has three steps: (i) bromination of position 8 of dG with bromide in water, (ii) displacement of bromide with hydrazine and (iii) reduction of the resulting hydrazine derivative with Nickel Raney. The phosphoramidite derivative of 8-amino-2′-deoxyguanosine protected by the (dimethylamino)methylidene (DMF) group was prepared as previously described (47,48). The DMT-protected hexaethyleneglycol phosphoramidite was obtained from Glen Research (USA).

Oligonucleotide synthesis

Oligonucleotide sequences containing 8aG were prepared on an automatic DNA synthesizer using standard 2-cyanoethyl phosphoramidites and the DMF-protected 8aG phosphoramidite (47,48). Sequences were: A, 5′-GAA G**1**A GGA GAT TTT TCT CCT CCT TC-3′; B, 5′-GAA G**1**A **1**GA **1**AT TTT TCT CCT CCT TC-3′; and C, 5′-AGA **1**AG AA$(EG)_6$ TTC TCT CT$(EG)_6$ TCT CTC TT-3′ where **1** represents 8aG and $(EG)_6$ represents hexaethyleneglycol. During solid-phase assembly of the sequences, 0.02 M iodine solution (tetrahydrofuran:water:pyridine) was used to prevent the formation of the N-cyano nucleosides (49). Complementary oligonucleotides containing natural bases were also prepared using commercially available chemicals and following standard protocols. After the assembly of the sequences, oligonucleotide-supports corresponding to the 8aG sequences were treated with 32% aqueous ammonia containing 0.1 M 2-mercaptoethanol (7 µl/ml of ammonia) at 55°C for 24 h (48). Ammonia solutions were concentrated to dryness and the products were desalted with a NAP-10 (Sephadex G-25) column and purified by reverse-phase HPLC. Oligonucleotides were synthesized on 0.2–1 µmol scales and with the last DMT group at the 5′ end (DMT on protocol) to help reverse-phase purification. HPLC solutions are 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). For analytical runs the following conditions were used: column, Nucleosil 120C_{18}, 250 × 4 mm, flow rate of 1 ml/min with a 30 min linear gradient from 0 to 60% B. For preparative runs the following conditions were used: columns, PRP-1 (Hamilton), 250 × 10 mm, flow rate of 3 ml/min with a 30 min linear gradient from 10–80% B (DMT on), or a 30 min linear gradient from 0–50% B (DMT off). Yield (in OD_{260} units after HPLC purification): sequence A (0.2 µmol), 15.8; sequence B (0.2 µmol), 10.4; sequence C (1 µmol), 85.

Melting experiments

Melting experiments described in Table 2 were performed as follows. Solutions of equimolar amounts of the hairpin oligonucleotide (h_{26}) and the 11mer (s_{11}) were mixed in the appropriate 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 with a programmed temperature increase of 0.5°C/min. Melts were run on a duplex concentration of 4 µM at 270 nm.

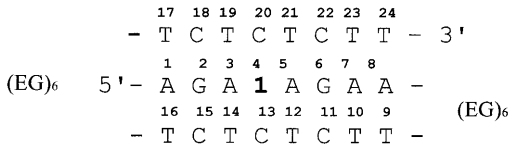
The samples used in the experiments described in Table 3 were prepared in a similar way but melting experiments were recorded at 260 nm and using 0.1, 0.5 and 1 cm path-length cells. The DNA concentration was determined by UV absorbance measurements (260 nm) at 90°C, using for the DNA coil state the following extinction coefficients: 7500, 8500, 12 500, 12 500 and 15 000 $M^{-1}cm^{-1}$ for C, T, G, 8aG and A, respectively (24).

Analysis of the thermodynamic data was performed as described by Xodo *et al.* (24). Melting curves were obtained at concentrations ranging from 0.77 to 32.6 µM of triplex containing three 8aGs. The melting temperature (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 (where C is the concentration) was linear. Linear regression of the data gave a slope of –0.0382 and a y-intercept of 0.00255, from which ΔH_t, and ΔS_t were obtained (24). The free energy was obtained from the standard equation: $\Delta G_t = \Delta H_t - T\Delta S_t$. Melting temperatures of the second transition (hairpin denaturation) were independent of the concentration (76–77°C) as expected for a unimolecular transition.

NMR

NMR experiments were performed on an intramolecular DNA triplex of sequence 5′-AGA **1**AG AA$(EG)_6$ TTC TCT

$CT(EG)_6$ TCT CTC TT-3′ (for numbering see Scheme 1). This particular sequence was chosen for the NMR study because intramolecular triplexes formed by oligonucleotides of similar sequences have been extensively investigated by Prof. Feigon's group (50). The solution structure of this sequence with unmodified oligonucleotide is available for comparison (51). Also, this sequence has been used to study the effect of including chemically modified bases within the structure of the triplex (10,11,52). For this reason, we decided to use the same oligonucleotide to check for the effect of 8aG on triplex formation. Purified samples of this oligonucleotide were dissolved in a phosphate buffer of either D_2O or 9:1 H_2O/D_2O (25 mM sodium phosphate buffer, 100 mM NaCl), giving a final oligonucleotide concentration of ~1 mM.



Scheme 1.

NMR spectra were acquired in a Bruker AMX spectrometer operating at 600 MHz, and processed with the UXNMR software. In the experiments in D_2O, presaturation was used to suppress the residual H_2O signal. NOESY (53) spectra were acquired with mixing times of 50 and 150 ms. TOCSY (54) spectra were recorded with standard MLEV-17 spin-lock sequence, and 80 ms mixing time. In 2D experiments in H_2O, water suppression was achieved by including a WATERGATE (55) module in the pulse sequence prior to acquisition. NOESY spectra in H_2O were acquired with 100 ms mixing time at 5°C. A jump-and-return pulse sequence (56) was employed to observe the rapidly exchanging protons in 1D H_2O experiments.

RESULTS AND DISCUSSION

Theoretical calculations

Molecular dynamics simulations show that no important structural changes occur from the G→8aG mutation in the duplex. The trajectory of the duplex containing 8aG deviates only 1.5 Å from the MD-averaged structure of the reference duplex. This value is just slightly larger than the root mean square deviation (r.m.s.d.) between the trajectory of the reference duplex and its MD-averaged structure (1.4 Å). Helical parameters, grooves and sugar puckering of the duplex containing 8aG are not significantly different from those of the reference duplex. In all cases the expected pattern of H-bond interactions is preserved during the trajectory. In summary, the structural impact of the G→8aG on the duplex is very small.

The effect of the G→8aG mutation in the triplex structure is also small, and located at the mutation site. For instance, the r.m.s.d. between the trajectory of the triplex with the $d(8aG{\cdot}C\text{-}C)^+$ triad and the MD-averaged structure of the reference triplex is just 1.05 Å, which compares with the self r.m.s.d. (the r.m.s.d. deviation between a trajectory and its averaged structure), as shown in Table 1. The three triplex structures containing 8aG were very similar in structure, irrespective to the nature of the Hoogsteen cytosine (protonated, amino or imino). This is shown in Table 1, where the r.m.s.d. between the three trajectories, and the MD-averaged structures of the remaining trajectories are shown. The r.m.s.d. values are very small (~1 Å), similar to those obtained when the trajectories are compared with their MD-averaged structures, and clearly within the thermal noise. All the triplex structures pertain to the B-family (average r.m.s.d. in the range of 1.0–1.2 Å), and are far from a canonical A-type triplex (average r.m.s.d. ~2.4 Å), as was found for the parent triplex (14).

Analysis of grooves, helical parameters, and sugar puckering do not show any clear change in the general structure of the triplex that could be related to the G→8aG mutation, even some local and slight local distortions were found near the substitution site. The expected patterns of H-bonded interactions are found during all the trajectories. The H-bond distances in neutral d(8aG·C-C) triads are in general slightly larger and more flexible than in the protonated one. The helical parameters of the triplex containing a central $d(8aG{\cdot}C\text{-}C)^+$ triad are in general similar to those of the triplexes containing neutral $d(8aG{\cdot}C\text{-}C)_i$, and $d(8aG{\cdot}C\text{-}C)_w$ triads. The most significant differences are found for the minor part of the major groove, which is slightly narrower (0.3 Å), and less electronegative [as seen in the Molecular Interaction Potential (MIP) and solvation maps, Fig. 3] when the central Hoogsteen cytosine is protonated than when it is neutral. This can be easily justified considering that the presence of a positive charge in the Hoogsteen position reduces the negative potential generated by the phosphates. Such a reduction implies a less favorable environment for an NH_2 group, which suggests that the presence of an amino group in the mM groove should be less favorable for protonated triads than for the neutral ones.

Table 1. r.m.s.d. (in Å) between the three triplex trajectories [triplexes with the $d(8aG{\cdot}C\text{-}C)^+$, $d(8aG{\cdot}C\text{-}C)_i$ and $d(8aG{\cdot}C\text{-}C)_w$ triads] and the MD-averaged structures

	Traj. $d(8aG{\cdot}C\text{-}C)^+$	Traj. $d(8aG{\cdot}C\text{-}C)_i$	Traj. $d(8aG{\cdot}C\text{-}C)_w$
Av. $d(8aG{\cdot}C\text{-}C)^+$	1.0	1.2	1.2
Av. $d(8aG{\cdot}C\text{-}C)_i$	1.1	1.0	1.1
Av. $d(8aG{\cdot}C\text{-}C)_w$	1.2	1.2	1.0
B-type triplex[a]	1.1	1.1	1.2
A-type triplex[a]	2.4	2.4	2.4

[a]As defined previously (38,14).

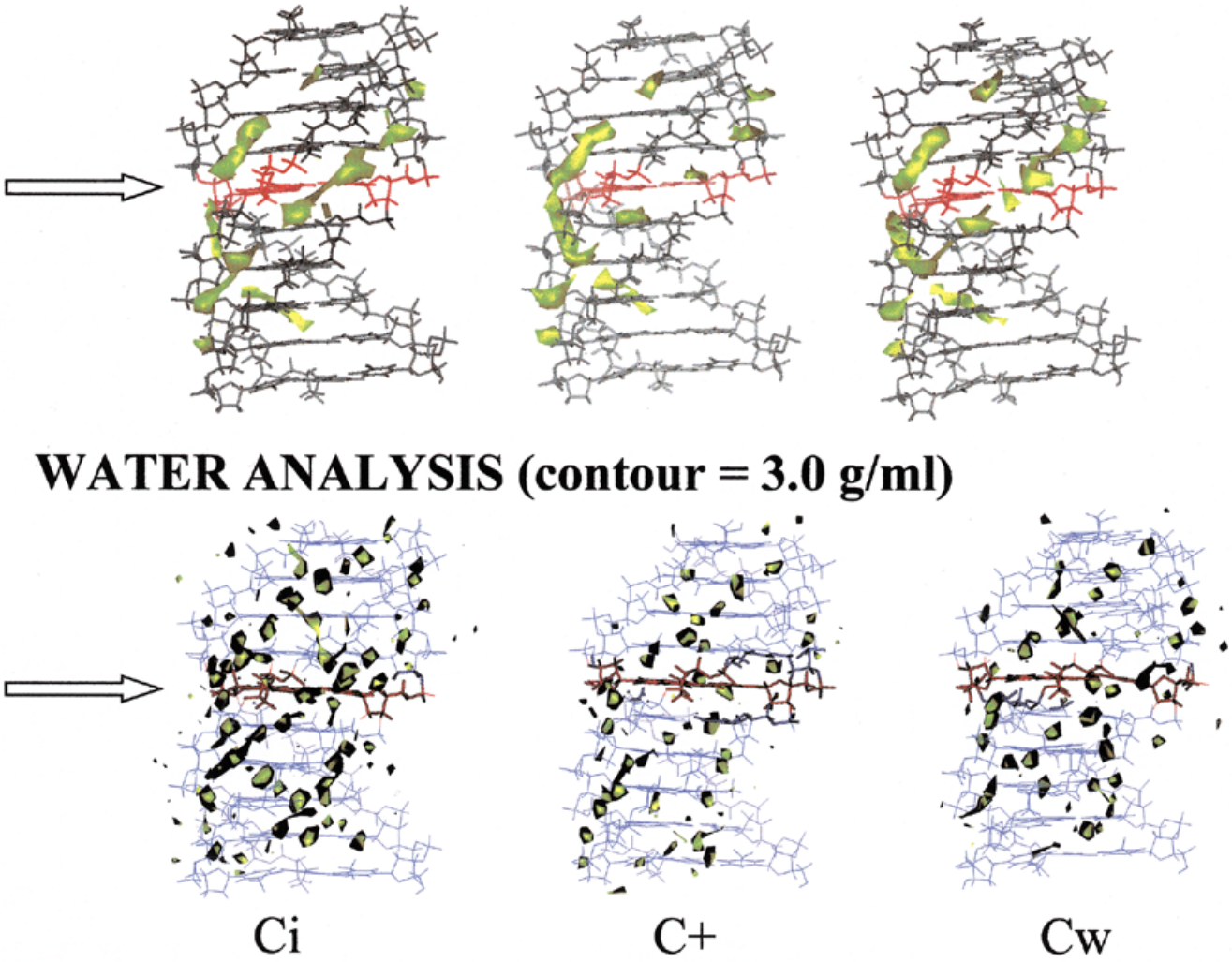


Figure 3. Molecular interaction potential (top) and molecular solvation maps (bottom) for the triplexes containing the d(G·C-C)$_i$ (left) d(G·C-C)$^+$ (middle) and d(G·C-C)$_w$ (right) trios. The –6.0 kcal/mol contour line for the interaction with O^+ is plotted for the MIP, while the 3 g/cm^3 is represented in solvation maps.

In summary, MD trajectories suggest that the G→8aG mutation does not introduce dramatic structural changes in either the duplex or the triplex. These results indicate that 8aG could be incorporated into DNA substituting G, and leading to stable duplexes and triplexes whose structures are not largely distorted with respect to the parent ones.

TI calculations were performed to estimate the changes in stability of triplexes due to the G→8aG mutation (see Materials and Methods and Fig. 2). As noted before, free energy change for the process: duplex + TFO (triplex forming oligonucleotide)→triplex due to the G→8aG mutation was determined for the three possible triads: d(G·C-C)$^+$, d(G·C-C)$_i$ and d(G·C-C)$_w$, even though for the particular sequence studied here the protonated triad is theoretically expected to be the most relevant form at neutral or acidic pH (14).

Differential free energy profiles (Fig. 4) for the 8aG(λ = 0)→G(λ = 1) are smooth. The difference between the free energy estimates obtained with the first (equilibration) and second (collection) halves of each window, and between the two independent mutations are very small. All these findings suggest that no large uncertainties in the results are expected due to errors in the simulation protocol. Results in Figure 4 clearly suggest that the mutation G→8aG stabilizes the triplex by ~1 kcal/mol. The magnitude of this stabilization depends on the nature of the triads, it being slightly larger for the neutral than for the protonated ones, as expected from the less negative values of the electrostatic potential in the mM groove.

Energy analysis of the three trajectories (data not shown) demonstrate that stacking interactions disfavor the G→8aG mutation, which is probably related to unfavorable electrostatic interactions of the exo-amino group with the neighbor triads. However, this effect is compensated by the larger strength of H-bonding interactions in triplexes containing 8aG, due mostly to the formation of an extra H-bond between the guanine and Hoogsteen cytosine (Fig. 1, 31,32).

Melting studies

In order to confirm the triplex-stabilizing properties of 8aG, oligonucleotides carrying 8aG were prepared. Oligonucleotide sequences A, B and C were prepared (see Materials and

Table 2. Melting temperatures[a] (°C) for the triplex h_{26}:s_{11}-containing 8-amino-2′-deoxyguanosine

	pH 4.5	pH 5.0	pH 5.5	pH 6.0	pH 6.5	pH 7.0
gaaggaggaga	56, 81	44, 81	40, 82	20, 82	–, 82	–, 82
gaag1aggaga	66, 81	53, 81	46, 80	36, 79	24, 79	–, 80
gaag1a1ga1a	74[b]	74[b]	72[b]	56, 73	45, 72	35, 72

The first value corresponds to the triplex dissociation (triplex→duplex), the second value is the denaturation of the hairpin (duplex→single strand).
h_{26}: 5′-GAAGXAXGAXATTTTTCTCCTCCTTC-3′; s_{11}: 5′-CTTCCTCCTCT-3′.
[a]1 M NaCl, 100 mM sodium phosphate/citric acid buffer.
[b]Only one transition with a hyperchromicity of 35% is found, suggesting similar stability for triplex and duplex.

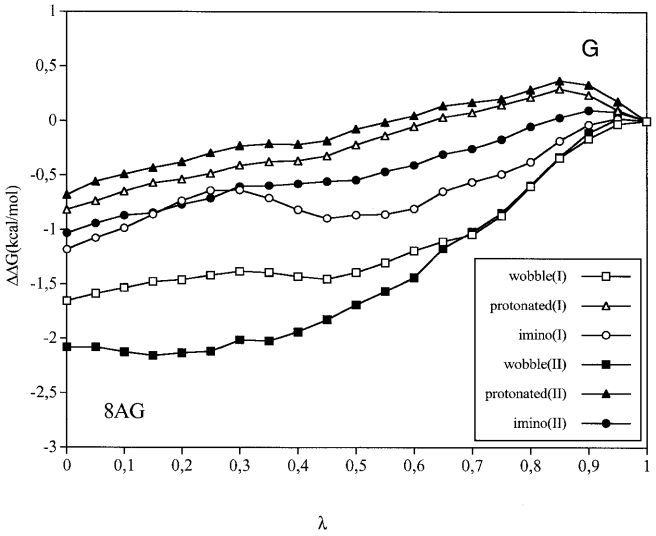


Figure 4. Differential free energy profiles (triplex–duplex) for the G(λ = 0)→8aG(λ = 1) mutation. Negative value means stabilization of the triplex by the mutation

Methods). Deprotection of oligonucleotides was performed with 32% aqueous ammonia containing 0.1 M 2-mercaptoethanol as described previously (48). The resulting products were purified by standard reverse-phase HPLC protocols. Characterization of the products was performed by mass spectrometry on shorter sequences (data not shown) and on sequence C by high field NMR (see below).

The triplex-stabilizing properties of 8aG were studied using a triplex model formed by a self-complementary hairpin of 26 bases (h_{26}) and an all pyrimidine single stranded oligonucleotide (s_{11}) described previously (24). In most cases two transitions were observed. As described previously (25), the first transition corresponds to the dissociation of the Hoogsteen strand (s_{11}) from the hairpin (triplex→duplex) and the second transition is the melting of the hairpin (duplex→single-strand). Melting temperatures at different pH are shown in Table 2. Substitution of G by 8aG in a triplex results in a high stabilization of the triplex (between 6 and 15°C per substitution), as predicted by MD and TI calculations.

Moreover, the dependence of the triplex→duplex transition from DNA concentration was studied on the triplex with three 8aGs. Table 3 shows the melting temperatures of the triplex→duplex transition at triplex concentrations ranging from 0.77 to 32.6 µM. The melting temperature at the higher concentration could not be measured due to the proximity of the transition corresponding to the hairpin denaturation. This second transition remained independent of the concentration, as expected for a unimolecular transition. Conversely, melting temperatures of the triplex→duplex transition were decreasing with concentration from 63°C at 17 µM to 50°C at 0.77 µM. The plot of $1/T_m$ versus lnC was linear giving a slope of –0.0382 and y-intercept of 0.00255, from which the following parameters were obtained: ΔH_t = –217 kJ/mol; ΔS_t = –544 J/mol·K; ΔG_t = –55 kJ/mol. In the same conditions, the thermodynamic parameters for the dissociation of an s_{11} derivative having 5-methylcytosine instead of cytosine from unmodified h_{26} were ΔH_t = –274 kJ/mol; ΔS_t = –784 J/mol·K and ΔG_t = –40 kJ/mol. Although the differences in the chemical nature of s_{11} should be taken into account, the comparison between these values gives a ΔG_t difference of –15 kJ/mol for three G→8aG mutations which is 5 kJ/mol per substitution (–1.2 kcal/mol) in good agreement with theoretical calculations (around –1 kcal/mol).

Table 3. Thermodynamic parameters for transition h_{26}:s_{11} = h_{26} + s_{11} in 100 mM sodium acetate (pH 6.0), 20 mM NaCl, $MgCl_2$ from the slope of the plot $1/T_m$ versus lnC[a]

Concentration	T_m	ΔH_t	ΔS_t	ΔG_t
µM/triplex	°C	kJ/mol	J/mol K	kJ/mol
0.77	49.7	–217	–544	–55
1.3	53.6			
2.1	54.4			
5.2	57.2			
8.0	60.9			
17.5	63.2			
32.6	nd[b]			

h_{26}: 5′-GAAG**1**A**1**GA**1**ATTTTTCTCCTCCTTC-3′; s_{11}: 5′-CTTCCTCCTCT-3′.
[a] ΔH_t, ΔS_t and ΔG_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; error on T_m is 0.7°C.
[b]not determined; the transition triplex to duplex was closed to the transition duplex to coil and therefore it was not possible to measure the melting temperature.

NMR spectroscopy

^{1}H NMR spectra were assigned using established techniques for right-handed helical nucleic acid structures using DQF-COSY, TOCSY and 2D NOESY spectra (57,58). Cytosine H5 and H6 protons were easily identified by their strong cross-peak in the TOCSY spectra in D_2O, and they were connected to their sugar spin system by the cross-peaks between H6 and H1′ and H2′/2″ in the NOESY spectra in D_2O. A similar procedure, starting from the Met-H6 TOCSY cross-peaks, was used to assign the thymines. NOE connectivities between sugar protons with their own base and their 5′-neighbor allowed for the sequential assignment of all deoxyribose spin systems. The deoxyribose spin system of the 8aG was identified by the sequential NOEs between H1′, H2′ and H2″ with H8 of the adjacent adenine 5.

Cytosine amino protons were identified by their strong cross-peaks with H5 in the NOESY spectra in H_2O. Connections between H5C→HN4C→H1G or H5C→HN4C→H3C$^+$ could be followed for the Watson–Crick and protonated Hoogsteen cytosines, respectively. Imino protons of the Watson–Crick or Hoogsteen thymines were assigned by their strong cross-peaks with adenine H2 or H8. In most d(A·T-T) triads, the complete path H8A→H3T$_h$→H6A→H3T$_{wc}$→H2A could be observed (Fig. 5; Scheme 1). Most of the exchangeable protons of adenines, cytosines and thymines were assigned, with the only exception of the imino and amino protons of the terminal trios. None of the amino protons of guanines was assigned. The 8-amino protons of the modified guanine were not detected in an ample range of pH and temperature, due probably to a very rapid intrinsic exchange with the solvent.

NMR-derived structure calculations are necessary to get a detailed description of the effect of 8aG in the structure of the triplex. However, some features of the structure of this molecule can be readily derived from the NMR spectra. Most importantly, the integrity of the triplex along the whole sequence can be assessed from the NOE pattern between exchangeable protons in canonical d(A·T-T) and d(G·C-C)$^+$ triplets. Also, the three Hoogsteen imino protons chemical shifts are typical of protonated cytosines (15.75, 14.97 and 14.96 p.p.m. for C$^+$18, C$^+$20 and C$^+$22, respectively). It is worth noting that a full protonation state for cytosines (specially at acidic pH) was expected for this sequence since the d(G·C-C) triad is surrounded by neutral triads (14,15).

Although the key cross-peak between H3C$^+$20 and the 8-amino of G4 was not observed, the chemical shift of this imino is consistent with the formation of a Hoogsteen base-pair with the modified guanine. Chemical shifts of the amino protons of cytosines 11, 13 and 15 are typical of Watson–Crick base-pairs (Fig. 5). In the case of C$^+$18 and C$^+$22, these protons are shifted to lower field, as usual in the Hoogsteen strand of triplexes, but this is not the case for the amino protons of C$^+$20 (8.74 and 7.70 p.p.m.). Although less pronounced, significant changes with respect to the unmodified triplex are also observed for the chemical shifts of the imino of the G4 and the amino protons of C13 (0.4 and 0.3 p.p.m., respectively). These observations may indicate some kind of slight distortion in the modified triad, which might be related to local rearrangements necessary to accommodate an additional H-bond.

It is interesting to notice that the differences in chemical shifts of the non-exchangeable protons between the modified and the unmodified triplex are localized in the neighborhood of the 8-amino group (mainly residues 4, 5, 20 and 21). Since the chemical shifts are very sensitive to any structural change, it can be concluded that the slight structural alterations at the mutation site are not transferred to the rest of the triplex, as suggested by MD simulations on a related triplex (see above).

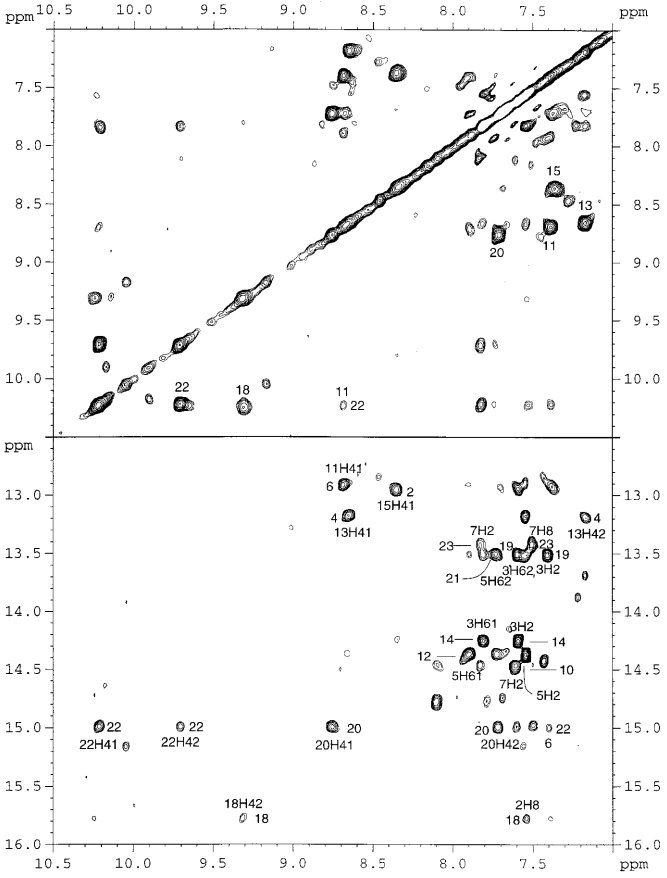


Figure 5. Two regions of the NOESY spectrum of 5′-AGA **1**AGAA(EG)$_6$ TTCTCT CT(EG)$_6$ TCTCTCTT-3′ in H_2O (150 ms mixing time, 1.0 mM oligonucleotide concentration, T = 5°C, pH 5). In the upper part, cross-peaks between vecinal cytosine amino protons are indicated. In the lower part, intraresidual imino–amino cross-peaks for the protonated cytosines are shown (18, 20 and 22), together with some inter-strand NOEs.

The stability of the triplex containing 8aGs, observed in the UV melting curves, is confirmed by the NMR spectra in H_2O. Exchangeable protons spectra at pH 7 at different temperatures are shown in Figure 6. The imino signals of the protonated cytosines are clearly observed at low temperature. Some signals, including the imino resonances of two Hoogsteen thymines, can be observed up to 45°C. These data are consistent with an stable triplex at neutral pH, and confirm that the low temperature transition observed in the UV melting experiments is, in fact, a triplex denaturation.

CONCLUSIONS

Theoretical calculations predict that the substitution of G by 8aG induces only small changes in the general structure of DNA triplexes. UV melting experiments, and NMR data confirm these suggestions.

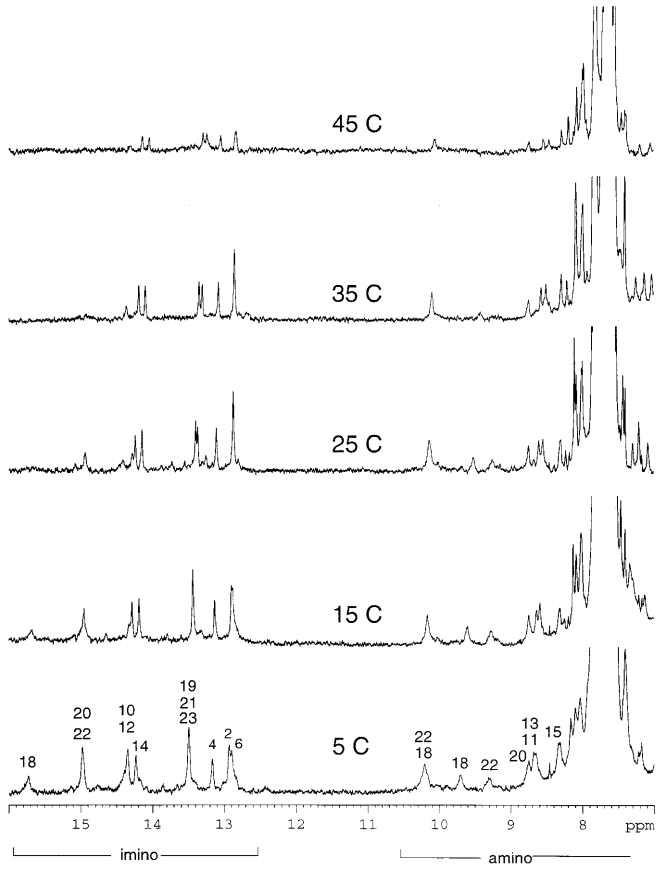


Figure 6. 1D exchangeable proton spectra of 5′-AGA **1**AGAA$(EG)_6$ TTCTCT CT$(EG)_6$ TCT CTC TT-3′ at pH 7 at different temperatures. Imino and some cytosine amino protons are labeled.

MD/TI calculations strongly suggest that the substitution G→8aG stabilizes the triplex in a wide range of pH. NMR and UV experiments demonstrate experimentally the large magnitude of the stabilization effect due to the G→8aG mutation. MD/TI and melting experiments suggest ~1 kcal/mol of stabilization for each mutation.

It is suggested that the stabilizing effect of the G→8aG mutation is related to: (i) the extra Hoogsten H-bond formed, and (ii) the favorable electrostatic interactions between the 8-NH_2 group and the negative mM groove of the triplex.

ACKNOWLEDGEMENTS

This work has been supported by the Centre de Supercomputació de Catalunya (CESCA, Mol. Recog. Project) and by the Spanish DGICYT (Projects PB96-1005, PM99-0046, PB98-1222, PB97-0941-C02-02). R.S. and J.R.B thank the CIRIT for predoctoral fellowships. This is a contribution of the Centre Especial de Recerca en Química Teòrica.

REFERENCES

1. Soyfer,V.N. and Potaman,V.N. (1996) *Triple-Helical Nucleic Acids.* Springer, New York.
2. Lee,J.S., Johnson,D.A. and Morgan,A.R. (1979) *Nucleic Acids Res.*, **6**, 3073–3091.
3. Völker,J. and Klump,H.H. (1994) *Biochemistry*, **33**, 13502–13508.
4. Wittung,P., Nielsen,P. and Norden,B. (1997) *Biochemistry*, **26**, 7973–7979.
5. Rajagopal,P. and Feigon,J. (1989) *Biochemistry*, **28**, 7859–7870.
6. Rajagopal,P. and Feigon,J. (1989) *Nature*, **339**, 637–640.
7. Radhakrishnan,I. and Patel,D.J. (1994) *Biochemistry*, **33**, 11405–11416.
8. Radhakrishnan,I. and Patel,D.J. (1994) *Structure*, **2**, 17–32.
9. Wang,E., Koshlap,K.M., Gilllespie,P., Dervan,P.B. and Feigon,J. (1996) *J. Mol. Biol.*, **257**, 1052–1069.
10. Koshlap,K.M., Schultze,P., Brunar,H., Dervan,P.B. and Feigon,J. (1997) *Biochemistry*, **36**, 2659–2668.
11. Asensio,J.L., Dhesai,J., Bergquist,S., Brown,T. and Lane,A.N. (1998) *J. Mol. Biol.*, **275**, 811–822.
12. Asensio,J.L., Brown,T. and Lane,A.N. (1999) *Structure*, **7**, 1–11.
13. Asensio,J.L., Brown,T. and Lane,A.N. (1998) *Nucleic Acids Res.*, **26**, 3677–3686.
14. Soliva,R., Laughton,C.A., Luque,F.J. and Orozco,M. (1998) *J. Am. Chem. Soc.*, **120**, 11226–11233.
15. Soliva,R., Luque,F.J. and Orozco,M. (1999) *Nucleic Acids Res.*, **27**, 2248–2255.
16. Plum,G.E. and Breslauer,K.J. (1995) *J. Mol. Biol.*, **248**, 679–695.
17. Leitner,D., Schröder,W. and Weisz,K. (1998) *J. Am. Chem. Soc.*, **120**, 7123–7129.
18. Cooney,M., Czernuszewicz,G., Postel,E.H., Flint,S.J. and Hogan,M.E. (1988) *Science*, **241**, 456–459.
19. Hélène,C. and Toulme,J.J. (1990) *Biochim. Biophys. Acta*, **1049**, 99–125.
20. Strobel,S.A. and Dervan,P. (1992) *Methods Enzymol.*, **216**, 309–321.
21. Grigoriev,M., Praseuth,D., Guieysee,A.L., Robin,P., Thuong,N.T., Hélène,C. and Harel-Bellan,A. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 3501–3505.
22. Sun,J.S. and Hélène,C. (1993) *Curr. Opin. Struct. Biol.*, **3**, 345–356.
23. Thuong,N.T. and Hélène,C. (1993) *Angew. Chem. Int. Ed. Engl.*, **32**, 666.
24. Xodo,L.E., Manzini,G., Quadrifolio,F., van der Marel,G.A. and van Boom,J.H. (1991) *Nucleic Acids Res.*, **19**, 5625–5631.
25. Sun,S., François,J.C., Montenay-Garesier,T., Saison-Behmoaras,T., Roig,V., Thuong,N.T. and Hélène,C. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 9198–9202.
26. Povsic,T.J. and Dervan,P. (1989) *J. Am. Chem. Soc.*, **111**, 3059–3061.
27. Ferrer,E., Fàbrega,C., Güimil García,R., Azorín,F. and Eritja,R. (1996) *Nucl. Nucl.*, **15**, 907–921.
28. Rana,V.S., Barawkar,D.A. and Ganesh,K.N. (1996) *J. Org. Chem.*, **61**, 3578–3579.
29. Barawkar,D.A., Rajeev,K.G., Kumar,V.A. and Ganesh,K.N. (1996) *Nucleic Acids Res.*, **24**, 1229–1237.
30. Xiang,G., Soussou,W. and McLaughlin,L.W. (1994) *J. Am. Chem. Soc.*, **116**, 11155–11156.
31. Güimil García,R., Ferrer,E., Macías,M.J., Eritja,R. and Orozco,M. (1999) *Nucleic Acids Res.*, **27**, 1991–1998.
32. Eritja,R., Ferrer,E., Güimil García,R. and Orozco,M. (1999) *Nucl. Nucl.*, **18**, 1619–1621.
33. Güimil García,R., Bachi,A., Eritja,R., Luque,F.J. and Orozco,M. (1998) *Bioorg. Med. Chem. Lett.*, **8**, 3011–3016.
34. Hattori,M., Frazier,J. and Miles,H.T. (1975) *Biochemistry*, **14**, 5033–5045.
35. Hattori,M., Frazier,J. and Miles,H.T. (1976) *Biopolymers*, **15**, 523–531.
36. Kumar,R.K., Gunjal,A.D. and Ganesh,K.N. (1994) *Biochem. Biophys. Res. Commun.*, **204**, 788–793.
37. Kawai,K., Saito,J. and Sugiyama,H. (1998) *Tetrahedron Lett.*, **39**, 5221–5224.
38. Shields,G., Laughton,C.A. and Orozco,M. (1997) *J. Am. Chem. Soc.*, **119**, 7463–7469.
39. Shields,G., Laughton,C.A. and Orozco,M. (1998) *J. Am. Chem. Soc.*, **120**, 5895–5904.
40. Arnott,S. and Hukins,D.W.L. (1972) *Biochem. Biophys. Res. Commun.*, **47**, 1504–1510.
41. Essmann,U., Perera,L., Berkowitz,M.L., Darden,T., Lee,H. and Pedersen,L.G. (1995) *J. Chem. Phys.*, **103**, 8577–8593.
42. Darden,T.A., York,D. and Pedersen,L. (1993) *J. Chem. Phys.*, **103**, 8577–8584.
43. Ryckaert,J.P., Ciccote,G. and Berendsen,J.C. (1977) *J. Comput. Phys.*, **23**, 327–342.
44. Bayly,C.I., Cieplak,P., Cornell,W.D. and Kollman,P.A. (1983) *J. Phys. Chem.*, **97**, 10269–10273.
45. Kirkwood,J.J. (1935) *J. Chem. Soc.*, **3**, 300–308.
46. Long,R.A., Robins,R.K. and Townsend,L.B. (1967) *J. Org. Chem.*, **32**, 2751–2756.
47. Rao,T.S., Durland,R.H. and Revankar,G.R. (1994) *J. Heterocyclic Chem.*, **31**, 935–940.

48. Rieger,R.A., Iden,C.R., Gonikberg,E. and Jonson,F. (1999) *Nucl. Nucl.*, **18**, 73–88.
49. Mullah,B., Andrus,A., Zhao,H. and Jones,R.A. (1995) *Tetrahedron Lett.*, **36**, 4373–4376.
50. Macaya,R., Wang,E., Schultze,P., Slenar,W. and Feigon,J. (1992) *J. Mol. Biol.*, **225**, 755–773.
51. Tarkoy,M., Phipps,A.K., Schultze,P. and Feigon,J. (1998) *Biochemistry*, **37**, 5810–5819.
52. Phipps,A.K., Tarkoy,M., Schultze,P. and Feigon,J. (1998) *Biochemistry*, **37**, 5820–5830.
53. Kumar,A., Ernst,R.R. and Wüthrich,K. (1990) *Biochem. Biophys. Res. Commun.*, **95**, 1–6.
54. Bax,A. and Davis,D.G. (1985) *J. Magn. Reson.*, **65**, 355–360.
55. Piotto,M., Saudek,V. and Sklenar,V. (1992) *J. Biomol. NMR*, **2**, 661–665.
56. Plateau,P. and Güeron,M. (1982) *J. Am. Chem. Soc.*, **104**, 7310–7311.
57. Wüthrich,K. (1986) *NMR of Proteins and Nucleic Acids.* John Wiley and Sons, New York.
58. Feigon,J., Koshlap,K.M. and Smith,F.W. (1995) *Methods Enzymol.*, **261**, 225–255.