T·C·G Triplet in an Antiparallel Purine·Purine·Pyrimidine DNA Triplex. Conformational Studies by NMR†

Karen Dittrich,‡ Juan Gu,‡ Robert Tinder,‡ Michael Hogan,‡ and Xiaolian Gao*,§

Center for Biotechnology, Baylor College of Medicine, 4000 Research Forest Drive, The Woodlands, Texas 77381, and Department of Chemistry, University of Houston, Houston, Texas 77204-5641

Received August 9, 1993; Revised Manuscript Received January 4, 1994®

ABSTRACT: The antiparallel *purine*·purine·pyrimidine DNA triplex, RRY6, which contains a *T*·C·G inverted triplet in the center of the sequence, was examined by proton and phosphoros two-dimensional NMR spectroscopy. The local conformation of the *T*·C·G triplet (*T4*·C11·G18) and the effect of this triplet on the global helical structure were analyzed in detail. The formation of the *T*·C·G triplet is confirmed by a set of cross-strand NOEs, including unusual cross-strand NOEs between the third strand and the pyrimidine strand as opposed to the purine strand of the duplex. NMR data suggest that the *T*·C·G triplet may be present in an equilibrium between a non-hydrogen-bonded form and a T(O4)–C(NH_2) hydrogen-bonded form and that there is a distortion of the in-plane alignment of the three bases. The flanking *G*·G·C base triplets are well-defined on the 5′-side of T4, but somewhat interrupted on the 3′-side of T4. The effect of the third strand binding on the Watson–Crick duplex was probed by an NMR study of the free duplex RY6. NMR parameters are affected mostly around the *T*·C·G inversion site. The perturbations extend to at least two adjacent base triplets on either side. The binding of the third purine strand and the accommodation of a central *T*·C·G inversion in RRY6 does not require a readjustment in sugar pucker, which remains in the range of C2′-endo. ^{31}P resonances of RRY6 distribute over a range of 2.2 ppm. The H–P coupling patterns of the third strand differ from those of the duplex. General spectral patterns defined by the marker protons of the RRY and YRY triplexes are compared.

Significant progress has been made toward understanding the structure and thermodynamics of triple-helix formation (Le Doan *et al.*, 1987; Moser & Dervan, 1987; Cooney *et al.*, 1988; Wells *et al.*, 1988; de los Santos *et al.*, 1989; Rajagopal & Feigon, 1989; Plum *et al.*, 1990; Cheng & Pettitt, 1992; Chubb & Hogan, 1992; Krawczyk *et al.*, 1992; Radhakrishnan *et al.*, 1992a; Roberts & Crothers, 1992; T'so *et al.*, 1992; Radkrishnan & Patel, 1993). This effort has allowed the assessment of the role of triplexes in cellular function and enhanced the prospect of using oligonucleotides and their analogs as a new generation of gene-specific agents. From the biochemical and pharmaceutical perspective, the triple-helix approach would have maximum utility if triple-helix-forming oligonucleotides could be synthesized with the ability to bind to a wide range of potential double- and single-stranded target sites. However, thus far, triplex-forming oligonucleotides comprising the standard bases display a target range limited to contiguous homopurine (R designates purine) and homopyrimidine (Y designates pyrimidine) segments. Base-pair inversions, such as T·A and C·G in homo-RY duplexes, substantially destabilize third strand binding. For instance, C·G inversion was found to form marginally stable *T*·C·G triplets with a third strand, T, in both RRY (purine·purine·pyrimidine) and YRY (pyrimidine·purine·pyrimidine) triplexes, while the *G*·T·A base triplet was found in YRY triplexes (Griffin & Dervan, 1989; Mergny *et al.*, 1991; Beal & Dervan, 1992a; Yoon *et al.*, 1992; Durland *et al.*, 1993). To overcome such limitations, extensive research activities have focused on the expansion of the coding schemes of natural base triplets (Ono *et al.*, 1991; Beal & Dervan, 1992b; Krawczyk *et al.*, 1992; Asseline & Thuong, 1993; Mohan *et al.*, 1993; Stilz & Dervan, 1993).

High-resolution NMR[1] spectroscopy is a powerful tool for understanding the structural basis of triplex formation (Sklenar & Feigon, 1990; Radhakrishnan *et al.*, 1992a; Radhakrishnan & Patel, 1993). Both the parallel YRY and the antiparallel RRY sequences have been investigated to obtain specific structural information, such as the alignment of base triplets, sugar geometry, and local conformation (de los Santos *et al.*, 1989; Rajagopal & Feigon, 1989; Pilch *et al.*, 1990; Mooren *et al.*, 1990; Radhakrishnan *et al.*, 1991, 1992b; Macaya *et al.*, 1992). Structures of a YRY triplex containing a *G*·T·A inverted triplet and an unperturbed RRY triplex have been determined to high resolution using experimental restraints derived from two- and three-dimensional (2D and 3D) NMR spectral analyses (Radhakrishnan *et al.*, 1992; Radhakrishnan & Patel, 1993). These structures provide detailed information about the global helical conformation and interproton contacts specifically related to nonconventional *G*·T·A triplet. These results shed light on specific molecular forces that are critical for stabilizing triplex formation.

This NMR study focuses on an RRY triplex (denoted as RRY6) containing a *T*·C·G triplet in the center of the sequence, as shown in Figure 1. Because preliminary thermodynamic

† This work was funded in part by grants to M.H. (NIHR01 AI3280) and a grant from the Triplex Pharmaceutical Corporation. The 600-MHz NMR spectrometer at the University of Houston was funded in part by the W. M. Keck Foundation. The instrument time on a 750-MHz spectrometer was generously provided by Bruker Instruments, Inc. Acknowledgement is made to the donors of the Petroleum Research Fund, administered by the American Chemical Society, for partial support of this research (26519-G4, to X.G.).

‡ Baylor College of Medicine.

§ University of Houston.

® Abstract published in *Advance ACS Abstracts,* March 15, 1994.

[1] Abbreviations: 1D, one-dimensional; 2D, two-dimensional; 3D, three-dimensional; COSY, correlation spectroscopy; DQF, double quantum filter; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlated spectroscopy.

0006-2960/94/0433-4111$04.50/0

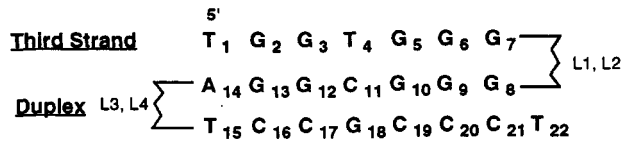

FIGURE 1: Sequence of the RRY6 triplex and RY6 duplex. Linker: O3′–P[$(OCH_2CH_2)_3OP]_2$–O5′.

analysis suggests that a third strand, T, is the least destabilizing of the four natural bases (Durland *et al.*, 1993), we have chosen it as a prototype for subsequent molecular design of triplexes of mixed sequences. In this article, we report a detailed analysis of the structural features of the RRY6 triplex and the local conformation of the *T*·C·G triplet. Additionally, a comparative NMR study of the RRY6 triplex and the related duplex RY6 is described.

MATERIALS AND METHODS

Synthesis of Poly(ethylene glycol)-Linked Oligonucleotide Sequences. RRY6 and RY6 were synthesized on a Milligen 7500 DNA synthesizer using β-cyanoethyl phosphoramidites. The hairpin RY6 duplex and the intramolecular RRY6 triplex contain hexaethylene glycol phosphate linkages (Glen Research), which were incorporated into the sequence using standard phosphoramidite chemistry (McBride & Caruthers, 1983). The oligonucleotides were detritylated on the synthesizer and deprotected by an overnight incubation in aqueous ammonia at 55 °C. Crude oligonucleotides were purified in a single run by ion-exchange chromatography on a Waters 600 HPLC system using a 2 × 10 cm Q-Sepharose column with a NaCl gradient in 10 mM NaOH [eluent A, 10 mM NaOH; eluent B, 1.5 M NaCl and 10 mM NaOH; 0–80 min, 100–98% A; 80–100 min, 98–60% A; 100–120 min, 60–0% A). These high-pH conditions effectively denature the secondary structure of the G-rich oligonucleotides and, thus, greatly improve the separation. The HPLC-purified oligonucleotides were desalted by washing extensively with water using ultrafiltration (Filtron). To suppress the formation of G4 tetraplexes (Sen & Gilbert, 1988), phosphate-bound Na^+ ions were exchanged for Li^+ ions by filtration with 10 mM LiOH. The purity of the oligonucleotides was confirmed by sequencing gel analysis of 5′-^{32}P-labeled sequences.

Sample Preparation. The concentration of RRY6 triplex in different samples ranged from 1–6 mM and that of the RY6 samples was 2–4 mM. The samples were dissolved in solutions containing 20–60 mM LiCl and one Mg^{2+} per DNA phosphate at pH 5.7–7.5. NMR spectra are comparable under these conditions.

NMR Experiments. (a) The RRY6 Triplex. NMR spectra were acquired on Bruker 600-MHz (University of Houston), 750-MHz (Bruker Instruments), and 500-MHz (Rice University) NMR spectrometers. 2D NOESY, TOCSY, COSY-35, and DQF-COSY (Derome, 1987) spectra were acquired at temperatures ranging from 5 to 20 °C, and the indirect-detected ^{1}H–^{31}P COSY spectra (Sklenar *et al.*, 1986) were collected at 20 and 32 °C, respectively. Higher temperatures usually were used for COSY-type experiments to achieve higher spectral resolution. Mixing times were 80, 100, 200, 250, and 500 ms for the NOESY experiments and 44 and 88 ms for the TOCSY experiments. Digital resolution of 2–4 Hz was used for various data sets. Sweep widths of 2000 Hz (3.33 ppm) in the proton dimension and 1200 Hz (4.55 ppm) in the ^{31}P dimension were used to acquire ^{1}H–^{31}P COSY spectra. A NOESY data set (80-ms mixing time) was recorded on a Bruker 750-MHz spectrometer at 5 °C using gradient-tailored water suppression (Piotto *et al.*, 1992). The two gradient pulses were equal in strength and lasted 50 ms. To optimize excitation at the 11- and 14-ppm regions, the corresponding delays used for the 3–9–19 pulse sequence were 60 and 100 μs, respectively. A total of 32 scans was collected with a repetition delay of 1 s. ^{1}H chemical shifts (uncorrected for solvent D_2O or H_2O) were referenced to the HOD resonance of 4.70 ppm at 25 °C. ^{31}P resonances were referenced to external trimethyl phosphate in an aqueous solution containing 0.1 M NaCl (pH 6.5). NMR data were processed using the FELIX program (Biosym/Hare Research) on an INDIGO/ELAN workstation.

The H1′–H2′ and H1′–H2″ coupling constants ($J_{1'2'}$ and $J_{1'2''}$) were measured from the antiphase multiplets of the COSY-35 spectra, and $J_{3'2'}$, $J_{3'2''}$, and $J_{3'4'}$ were derived from analyses of various COSY and TOCSY data sets. The assignments of ^{31}P resonances from ^{1}H–^{31}P correlation spectra were assisted by the assignments of the H3′, H4′, H5′, and H5″ protons. H–P coupling constants were analyzed by 1D simulation of multiple couplings in the proton dimension on the basis of homo- and heteronuclear Karplus equations (Majumdar & Hosur, 1992). For instance, $J_{2'3'}$, $J_{2''3'}$, and $J_{3'4'}$ were considered for the observed H3′–P coupling cross peaks.

(b) The RY6 Duplex. A set of NMR spectra was recorded for the RY6 duplex at 10–20 °C using spectral acquisition and data processing procedures similar to those described above. The comparison of NOE intensities between RY6 and RRY6 was made mainly using 100-ms NOESY data at pH 6.2, except for some weak NOEs, which were assigned from NOESY of longer mixing times.

Model Structure of the T·C·G Triplet. A starting structure for the central three triplets of RRY6 was obtained by modifying a triplex generated using the INSIGHT program (BIOSYM). NOEs related to the central three triplets of RRY6, including about 15 NOEs related to the *T*·C·G site, were used for calculation of the model structure using the XPLOR program (Molecular Simulations). A simulated annealing approach was used to gradually increase the forces of NOE restraints during dynamics simulations. Sugar C3′–C4′ torsion angles were restricted to 130 ± 10°. The final model structure generally satisfies the local NOEs used in calculation.

RESULTS AND DISCUSSION

This study addressed two major questions: (a) the local conformation of the *T*·C·G triplet in the antiparallel RRY triplex and the effect of the triplet on global helical structures, and (b) the induced structural perturbation of the Watson–Crick duplex upon binding of the third strand. The poly(ethylene glycol) linkers were used to connect single strands in RRY6 and RY6 (Figure 1). The spectra of the triplexes containing oligothymine or poly(ethylene glycol) linkers are comparable (data not shown). However, the spectra of the latter are much more simplified, with the majority of the linker protons resonating at 3.0–4.0 ppm and the linker phosphates resonating at a downfield region (greater than −3.8 ppm). Most of these positions are resolved from those of the DNA resonances. We have analyzed NMR spectra of the RRY6 triplex and the RY6 duplex and assigned nearly all proton and ^{31}P resonances (Table 1).

Exchangeable Proton Assignments Confirm Overall RRY Triplex Formation. The formation of triplex was observed in 1D imino proton spectra at 10 °C as a function of the gradual addition of Mg^{2+} in the presence of 20 mM LiCl at neutral pH (Figure S1, supplementary material). During this

Table 1: Proton and ^{31}P Assignments of RRY6 and RY6 and Chemical Shift Differences between the Two Sequences

(a) Proton and ^{31}P Assignments of the RRY6 Triplex[a] (ppm)

	H1/H3	H2/H4(b)	H2/H4(nb)	H8/H6	H2/H5	H1′	H2′	H2″	H3′	H4′	H5′&H5″	^{31}P
T1	11.46			7.14	1.74	5.81	1.49	2.05	4.52	3.86	3.55/3.58	−4.34
G2	13.65			8.00		5.08	2.71	2.58	4.86	4.21	3.86/3.92	−3.35
G3	13.63	6.42		8.05		5.86	2.80	2.53	5.04	4.28	4.00/4.08	−3.11
T4	11.05			7.18	1.49	6.11	1.61	2.11	4.82	4.23	3.93	−4.88
G5	13.30			8.06		5.24	2.61	2.51	4.91	4.28	3.97	−3.81
G6	13.93	6.55		8.02		5.23	2.74	2.61	4.97	4.41	4.04	−2.89
G7	14.17	7.12		8.08		6.37	2.70	2.72	5.06	4.50	4.14/4.19	−3.57
G8	13.46			7.66		5.99	2.71	2.92	4.86	4.26	3.69/3.77	−4.28
G9	13.35			7.45		5.97	2.66	2.91	5.06	4.34	4.12/4.17	−5.12
G10	13.23			7.12		5.66	1.30	2.55	4.41	4.32	3.43/4.22	−4.71
C11		8.53	7.40	6.43	4.62	5.62	1.72	2.14	4.74	4.03	3.62/3.83	−4.06
G12	13.30			7.51		5.82	2.58	2.81	5.00	4.23	3.96/4.03	−4.45
G13	13.31			7.38		5.49	2.30	2.52	4.85	4.29	4.14	−4.42
A14				8.13	8.13	6.29	2.54	2.63	4.92	4.36	4.13	−3.73
T15	14.03			7.57	1.78	6.12	2.36	2.53	4.85	4.18	3.77/3.90	−4.31
C16		8.81	7.30	7.64	5.70	5.94	2.26	2.47	4.79	4.19	4.10	−3.97
C17		8.94	7.23	7.32	5.62	5.89	1.89	2.46	4.85	4.11	4.04/4.11	−3.90
G18	13.18			7.88		5.96	2.62	2.79	4.95	4.33	4.00/4.14	−4.44
C19		8.49	6.97	7.20	5.37	5.70	2.06	2.42	4.58	4.15	4.10/4.28	−4.17
C20		8.77	7.17	7.47	5.42	5.85	2.01	2.35	4.72	4.06	4.06	−4.22
C21		8.71	7.17	7.57	5.59	6.16	2.20	2.38	4.75	4.16	4.00	−4.04
T22	10.81			7.40	1.65	6.06	2.14	2.14	4.43	3.96	3.91/4.07	

Linker Proton and ^{31}P Assignments (25 °C)

	linker proton	^{31}P		linker proton	^{31}P
L1	3.26, 3.46	3.61/3.68	L3	3.88, 3.93	3.34/3.38/3.43	−3.29
L2	3.61/3.68	−2.92	L4	3.34/3.83/3.43	3.87	−2.76

(b) Proton Assignments of RY6 and Chemical Shift Differences between RRY6 Triplex and RY6 Duplex[b] (ppm)

	imino H1/H3		amino H4(b)		amino H4(nb)		H8/H6		H2/H5		H1′		H2′		H2″		H3′		^{31}P	
	RY6	Δ(ppm)	RY6	Δ(ppm)	RY6	Δ(ppm)	RY6	Δ(ppm)	RY6	Δ(ppm)	RY6	Δ(ppm)	RY6	Δ(ppm)	RY6	Δ(ppm)	RY6	Δ(ppm)	RY6	Δ(ppm)
G8							7.80	0.14			5.59	−0.40	2.46	−0.25	2.62	−0.30	4.74	−0.12	−4.00	0.28
G9	12.93	−0.42					7.76	0.31			5.65	−0.32	2.62	−0.04	2.72	−0.19	4.94	−0.12	−4.09	1.03
G10	12.79	−0.44					7.72	0.60			5.83	0.17	2.52	1.22	2.64	0.09	4.93	0.52	−4.21	0.50
C11			8.21	−0.32	6.30	−1.10	7.22	0.79	5.27	0.65	5.61	−0.01	1.80	0.08	2.22	0.08	4.78	0.04	−4.26	−0.20
G12	12.94	−0.36					7.78	0.27			5.42	−0.40	2.55	−0.03	2.61	−0.20	4.91	−0.09	−3.99	0.46
G13	12.76	−0.55					7.69	0.31			5.49	0.00	2.46	0.16	2.57	0.05	4.92	0.07	−3.90	0.52
A14							8.23	0.10	8.08	−0.05	6.30	0.01	2.63	0.09	2.71	0.08	4.96	0.04	−3.74	−0.01
T15							7.64	0.07	1.83	0.05	6.11	−0.01	2.33	−0.03	2.53	0.00	4.83	−0.02	−4.27	0.04
C16			8.34	−0.47	6.88	−0.42	7.58	−0.06	5.73	0.03	5.91	−0.03	2.16	−0.10	2.37	−0.10	4.79	0.00	−3.90	0.07
C17			8.49	−0.45	6.79	−0.44	7.44	0.12	5.61	−0.01	5.53	−0.36	2.03	0.14	2.35	−0.11	4.80	−0.05	−3.88	0.02
G18	12.83	−0.35					7.89	0.01			5.84	−0.12	2.64	0.02	2.66	−0.13	4.94	−0.01	*c*	
C19			8.09	−0.40	6.34	−0.63	7.36	0.16	5.36	−0.01	5.90	0.20	2.11	0.04	2.41	−0.01	4.73	0.15	−4.06	0.11
C20			8.37	−0.40	6.81	−0.36	7.49	0.02	5.59	0.17	5.91	0.06	2.11	0.10	2.38	0.03	4.74	0.02	−3.84	0.38
C21							7.54	−0.03	5.62	0.03	6.15	−0.01	2.20	0.00	2.20	−0.18	4.48	−0.28		

Linker Proton and ^{31}P Resonances

	proton	^{31}P		proton		^{31}P
L3	3.87/3.94	3.24/3.33/3.46	−3.08	L4	3.34/3.83/3.43	3.83	−2.70

[a] Exchangeable proton assignments were from a spectrum recorded in 90% H_2O in the presence of 20 mM LiCl and 44 mM $MgCl_2$ at 10 °C, pH 6.3. Nonexchangeable proton assignments were from a spectrum of the same sample in D_2O. The assignments of the H5′ and H5″ protons are arbitrary. [b] Proton assignments of RY6 were obtained at 20 °C from a sample containing 2 mM duplex in the presence of 20 mM LiCl and 28 mM $MgCl_2$, pH 6.2. Δ(ppm) = ppm(RY6) − ppm(RRY6); a positive number indicates that the chemical shift of the RY6 duplex is downfield shifted compared to that of the RRY6 triplex. [c] Overlaps with G9 or is missing.

process, new peaks gradually appear at the 13.6–14-ppm region and sharpen up as [$MgCl_2$] passes 1 Mg^{2+} per triplex molecule.[2] The spectra remain unchanged when $MgCl_2$ titration exceeds 1.5 Mg^{2+} per triplex molecule. The imino protons of RRY6 do not melt until 32 °C. Beyond that temperature, the signals of the third strand imino protons begin to broaden.

An expanded 2D NOESY spectrum containing imino protons of RRY6 in F_2 and amino/base protons in F_1 dimension is shown in Figure 2. The assignments of the exchangeable protons of RRY6 are given in Table 1a. Imino protons of G8, G9, G10, G12, G13, and G18 at 13.0–13.5 ppm manifest spectral patterns typical of a helical duplex, that is, each base pair gives a pair of intense NOEs linking G imino to C amino protons of the same base pair. There are also weak sequential NOEs from G imino to adjacent C amino protons. Notably, four imino protons of the third strand are well-resolved from their duplex counterparts, with H1 of G3, G6, and G7 shifted downfield (13.6–14.1 ppm) and T4(H3) shifted upfield significantly (11.05 ppm) to a terminal imino proton region. We cannot unambiguously assign the G2(H1) resonance. The only cross peak that is assignable to the G2 imino is the NOE linking G2(H1) with G12(H8).

There are a number of characteristic 2D cross peaks defining the orientation of the third strand (Table 2). (a) Each of the third strand imino protons interact with two base protons: one is from the purine strand of the same base triplet and the other is from the adjacent residue (imino H_k–base H_i and imino H_k–base H_{i-1}, where k and i are the residue numbers of the third strand and the purine strand, respectively), such as G3(H1)–G12(H8) and G3(H1)–C11(H5). These inter-

[2] J. Gu and X. Gao, unpublished results. We have consistently observed such a Mg^{2+} dependence in the RRY triplex formation of several different sequences. The function and quantitative aspect of Mg2+ in triplexes are under investigation.

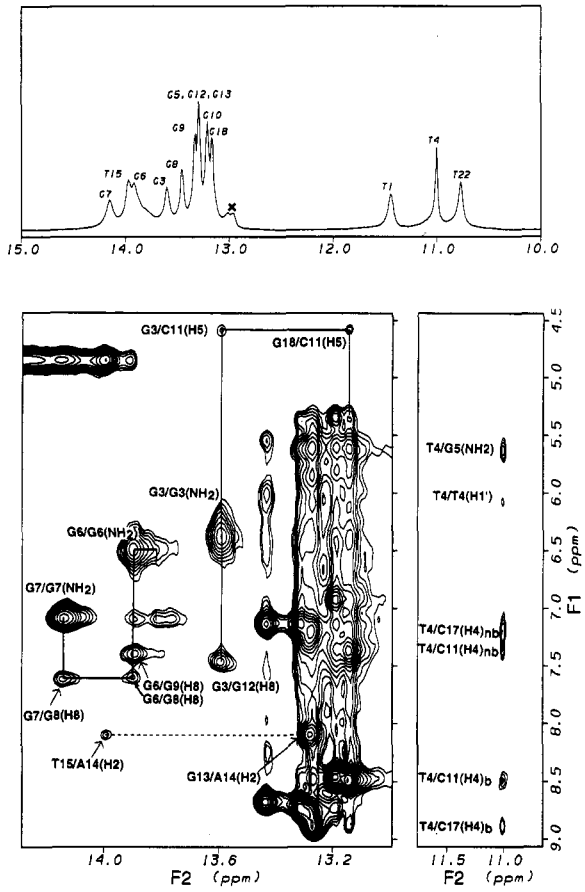

FIGURE 2: NOESY spectra of RRY6 (2 mM) recorded in H_2O in the presence of 20 mM LiCl and 44 mM $MgCl_2$ at 10 °C, pH 6.3. Drawn on top is the 1D projection and assignments of the imino protons. Minor peaks labeled with × are not from the RRY6 triplex. Parts of the NOE assignments of intra- or cross-strand interproton contacts linking third strand imino protons (F_2 axis) and all other base protons (F_1 axis) are indicated.

strand NOEs arise from the in-plane alignment of base triplets, as shown in Figure 3, and the right-handed helical stacking between adjacent base triplets (Figure 4). (b) Third strand imino protons interact with H2′ and H2″ of the purine strand (Table 2a and Figure 4). These NOEs correlate protons of the same triplet and the adjacent residue [imino H_k–(H2′,2″)$_i$ and imino H_k–(H2′,2″)$_{i-1}$], such as G3(H1)–G12(H2′) and G3(H1)–C11(H2′,2″). These characteristic spectral patterns delineate the interface between the base residues of the purine third strand and the major groove of the duplex. In contrast, YRY-type triplexes typically show interactions between the "minor groove phase" (the side that would form the minor groove in a duplex) of a pyrimidine third strand and the major groove of the duplex (Radhakrishnan *et al.*, 1992b). Furthermore, unlike YRY triplexes, cross-strand NOEs involving the sugar protons of the third strand are not present in RRY6. This is consistent with a general model of the RRY triplex, with sugar moieties of the third strand directed away from the duplex purine strand. (c) We observed NOEs connecting third strand G amino protons of residue *k* to duplex C amino protons of residues *j* and *j* – 1 in the pyrimidine strand (Table 2a). These interproton contacts confirm the in-plane alignment of the *G*·G·C triplets in RRY6 in the presence of a *T*·C·G triplet in the center of the triplex.

A noteworthy spectral feature related to the third strand is a set of strong and well-defined NOE cross peaks linking G imino to its own amino protons (Figure 2). Such NOEs are not detected in the duplex motif of RRY6 and usually are absent in Watson–Crick base-paired G residues due to the

Table 2: RRY6 NOEs

(a) NOEs Related to the Third Strand of the RRY6 Triplex[a]

third strand	purine strand	pyrimidine strand	NOE s > m > w
Third Strand Imino to Duplex Base Protons			
T1(H5)	G13(H8)		m
G3(H1)	G12(H8)		s
G3(H1)	C11(H5)		w
G5(H1)	G10(H8)		w
G6(H1)	G9(H8)		m
G6(H1)	G8(H8)		w
G7(H1)	G8(H8)		m
Third Strand Imino to Duplex Sugar Protons			
G2(H1)	G12(H2′)		w–
G3(H1)	C11(H2′)		w
G3(H1)	C11(H2″)		w
G6(H1)	G9(H2′)		w
G6(H1)	G9(H2″)		w
G7(H1)	G8(H2′)		w
Third Strand Amino to Duplex Amino			
G3(H2)	C11[H4(b)]		w–
G3(H2)		C16[H4(b)]	w
G3(H2)		C17[H4(b)]	w–
G6(H2)		C19[H4(b)]	m
G6(H2)		C20[H4(b)]	w

(b) Unusual NOEs Related to the T·C·G Triplet[b]

third strand	duplex purine	duplex pyrimidine	NOE
T4–G3			
T4(H5)	G3(H1′)		s
T4(H5)	G3(H2′)		w–
T4(H5)	G3(H2″)		w+
T4(H5)	G3(H3′)		w
T4(H5)	G3(H4′)		w–
T4–C17			
T4(H3)		C17[H4(nb)]	m
T4(H3)		C17[H4(b)]	w
T4–G10			
T4(H5)	G10(H8)		w+

Predicted and Observed NOEs of the T·C·G Base Triple (Figure 3)[c]

T4–C11		form I	form II	form III	NOE
T4(H3)	C11[H4(b)]	w	w+		w–
T4(H3)	C11[H4(nb)]	s	m	m	m
T4(H3)	C11(H5)	m	w	s	w
T4(H3)	C11(H6)			w+	
T4(H5)	C11[H4(nb)]		w		
T4(H5)	C11(H5)	w	w+		w
T4(H5)	C11(H6)				

[a] NOEs were assigned from the 200-ms NOESY. Experimental conditions were the same as described for Table 1. [b] NOEs were assigned from the 80- and 200-ms NOESY spectra. Experimental conditions were the same as described for Table 1. [c] As shown in Figure 3, form I contains no H-bond but T(H3) and C(H4) are aligned; forms II and III contain potential H-bonds between T(O4) or T(O2) to C[H4(nb)], respectively. m: medium; s: strong; w: weak; +: more, –: less.

broad line widths of G amino protons. The detection of an averaged amino resonance from the residues of the third strand suggests that, under the conditions used for NMR studies, the G C2–N2 bond is in fast rotation and that the amino protons are in slow exchange with solvent H_2O. In another NMR study of a triplex of the RRY-type, similar G imino–amino proton NOEs are also present (Radhakrishnan *et al.*, 1991b). This different behavior of the G amino resonances of the third strand compared to those of the purine strand in the duplex can be attributed to their C–N bonding character and the local environment. It is interesting to note that the third strand G amino protons are located in the major groove and are hydrogen-bonded to the duplex G(O6) and that those of the duplex G amino protons are located in the minor groove and are hydrogen-bonded to the C(O2) in RRY triplexes.

Sugar Conformation. The sugar conformation of RRY6 is probed by proton coupling constants derived from various

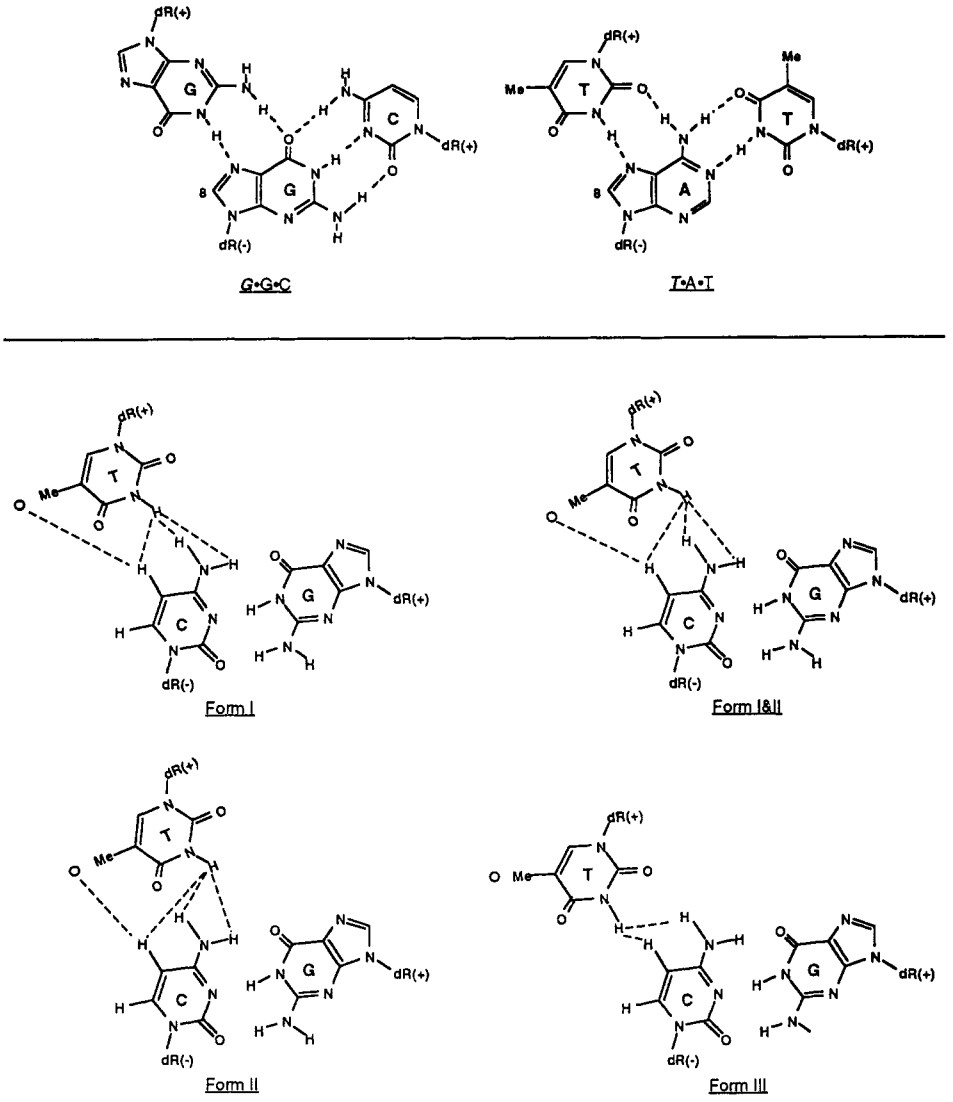

FIGURE 3: Alignment of the RRY triplets *G*·G·C and *T*·A·T and the inverted *T*·C·G. Three forms are considered for the T·C·G triplet. Form I contains no H-bond, but T(H3) and C(H4 nb) are aligned; form II contains a potential H-bond between T(O4) and C(H4 nb); form III contains a potential H-bond betweenT(O2) and C(H4 nb). An intermediate form of I and II is shown as form I&II. For comparison purposes, some of the interproton distances in the *T*·C·G triplet are indicated as dashed lines. An open circle approximates the pseudocenter of the T methyl group.

COSY experiments. A summary of these results is listed in Table S1 of the supplementary material. The H1′–H2′ and H1′–H2″ couplings, which are derived from the antiphase multiplets in COSY-35 spectra, are in the range of 8.0 ± 1.5 and 6.5 ± 1.0 Hz, respectively. These results and the analysis of the coupling patterns of the H2′–H3′, H2″–H3′, and H3′–H4′ cross peaks from TOCSY data (Table S1, supplementary material) indicate that all sugar moieties, including those of the *T*·C·G inversion site, adopted a conformation that is closely related to the C2′-endo ring pucker (Majumdar & Hosur, 1992). As is often observed in DNA duplexes, the sugar rings of purine residues of RRY6, which are associated with very weak or null $J_{3'4'}$, are more typical of C2′-endo than those of the pyrimidine residues, which are associated with moderate $J_{3'4'}$. This is because the H3′ proton in a rigid C2′-endo pucker is orthogonal to H4′, resulting in negligible coupling constants between the two protons.

A close examination of the difference between $J_{1'2'}$ and $J_{1'2''}$, as well as the value of $J_{3'4'}$, indicates that sugar protons of the third strand exhibit the largest difference in $\Delta(J_{1'2'}-J_{1'2''})$ and nearly null $J_{3'4'}$ (Table S1, supplementary material). This suggests that the contribution of the C2′-endo conformation to the sugar conformation in each strand increases in the order of pyrimidine strand < duplex purine strand < third purine strand. This observation implies a preferred C2′-endo conformation for the third purine strand that binds to a G-rich duplex. Previously, the homo-YRY triplexes have been found to adopt predominantly the C2′-endo sugar pucker (Macaya *et al.*, 1992). The insertion of a *G*·T·A inversion in a YRY triplex resulted in a C3′-endo-type conformation (C2′-exo) of the G sugar (Radhakrishnan *et al.*, 1992a). Further studies of triplexes are needed to understand the preference of sugar conformation of the third strand and the unusual base triplet in the framework of RRY triplexes.

Perturbation of the Phosphate Backbone. The assignments of ^{31}P resonances (Table 1) are obtained from a combined analysis of ^{1}H–^{31}P and ^{1}H coupling correlation spectra. The RRY6 triplex contains a total of 25 phosphorus resonances, including 21 internucleotide O3′–P–O5′ phosphates and two O5′–P–O_{linker} and two O3′–P–O_{linker} linker phosphates. Three of the four hexaethylene glycol ^{31}P resonances are shifted downfield relative to the internucleotide ^{31}P resonances (Figure 5). The internucleotide phosphate resonances of RRY6 are detected over a frequency range of 2.2 ppm, in contrast to the 0.5-ppm ^{31}P chemical shift dispersion for the parent duplex. Chemical shift analysis of ^{31}P resonances reveals a sequence-dependent pattern: the ^{31}P resonances centered at the *T*·C·G inversion site, such as T4–P–G5, G9–P–G10, and G10–P–

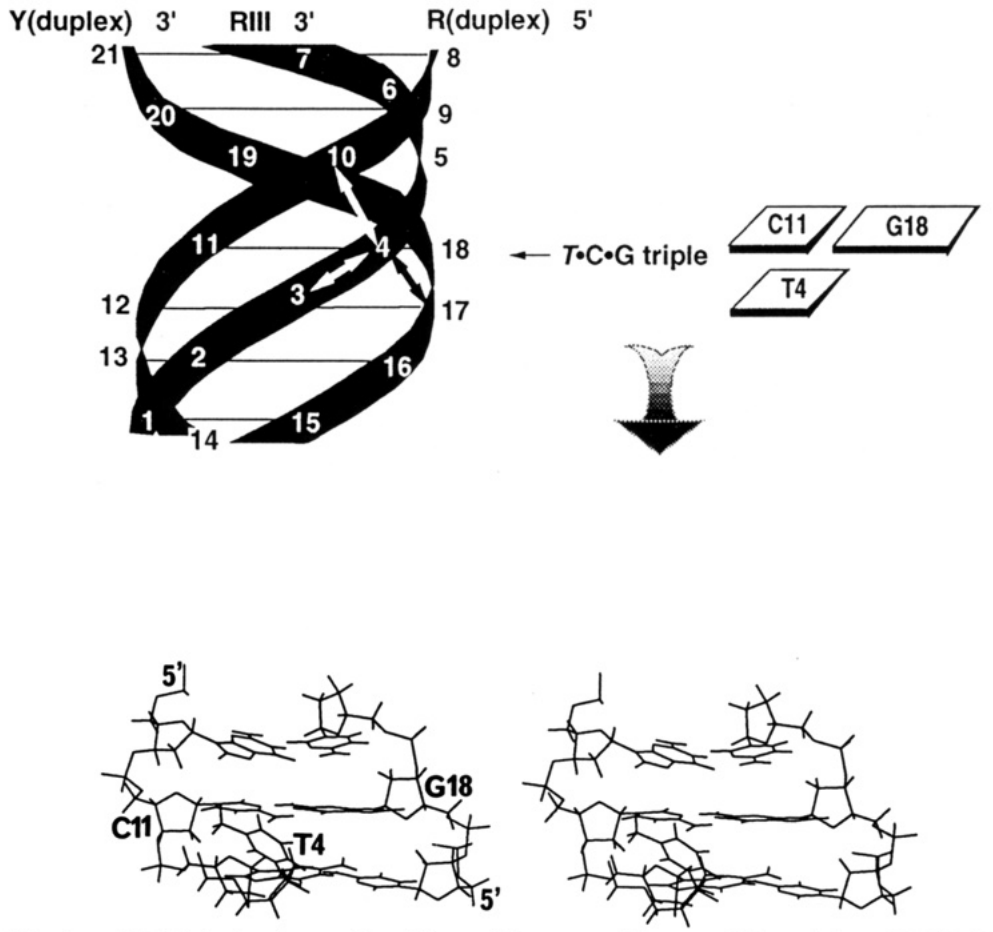

FIGURE 4: Helical drawing of RRY6 (top) and stereoview of the model structure of the central three triplets of RRY6. For clarity, only T4 is shown for the third strand.

C11, are shifted upfield, while the ^{31}P resonances from the third strand, such as G2–P–G3, G3–P–T4, G6–P–G7, are shifted downfield. The most perturbed linkages, T4–P–G5 and G9–P–G10, are confined to a local site, consistent with the antiparallel helical alignment of the third strand (Figure 4, the schematic helical drawing). Besides the influence of chemical and electronic environment, ^{31}P chemical shifts are most sensitive to bonding geometry, such as parameters of α and ζ (bond angle O3′–P-O5′) and β and ϵ (dihedral angles P–O5′–C5′–C4′ and C4′–C3′–O3′–P) (Giessner-Prettre *et al.*, 1984; Gorenstein, 1984; Lankhorst *et al.*, 1984). For instance, the resonances shifted downfield from the duplex ^{31}P region are attributed to a BI to BII form transition (*ca.* –4.5 to –4.0 ppm) or strand unwinding in intercalation complexes (*ca.* –4.5 to –2.0 ppm; Giessner-Prettre *et al.*, 1984; Gorenstein, 1984; Gao & Patel, 1989). Thus, the observed ^{31}P chemical shifts from the third strand and the *T*·C·G base triplet are likely to be associated with the variation in backbone conformation of the third strand compared to that of the unperturbed form (*vide infra*) and with a unique local environment at these sites.

The coupling cross peaks in the ^{1}H–^{31}P correlation spectrum (Figure 5) can be divided into three types that differ in spectral coupling pattern and intensity (Table 3; only those of nonterminal and reasonably well-resolved peaks are considered). Types I and II include residues of the third strand and those adjacent to the *T*·C·G site. In most cases, each ^{31}P is coupled to three protons, the H3′, H4′, and one of the H5′ protons (the assignments of H5′ and H5″ are arbitrary). The H4′–P cross peaks are well-resolved from those of the H5′–P in type I and type II residues. We have analyzed the spectral patterns using the assumptions that sugar conformations are in the C2′-endo region, which corresponds to ($J_{2'3'} \sim 6$ Hz) $\gg$ ($J_{2''3'} \sim 0$ Hz) and $4 > J_{3'4'} > 0$ Hz, and that the third strand sugar moieties are more C2′-endo-like compared to those of the duplex motif (*vide supra*). Additionally, the backbone H5′ and H5″ protons are considered to project away from the sugar ring, since base–H5′,5″ cross peaks are absent from the NOESY spectra at 80- and 100-ms mixing times. Given the above conditions and the observed H–P coupling pattern (Table 3), the coupling of H3′-P is estimated to be comparable to $J_{2'3'}$ (~6 Hz) for type I residues. According to the heteronuclear Karplus equations (Majumdar & Hosur, 1992), this H3′–P coupling confines the ϵ torsion angle (C4′–C3′–O3′–P) in the range of 80–95° or 265-280°.

Compared to the type I cross peaks, the H3′–P cross peaks of types II and III exhibit reduced intensities and slightly larger splittings (Table 3 and Figure 5). This trend is consistent with a relatively smaller $J_{3'P}$ and the increased contribution from $J_{3'4'}$ and $J_{2''3'}$, reflecting a combined variation in backbone ϵ and sugar torsion angles in residues of types II and III. Type I H–P cross peaks are well-resolved in the H4′–P and H5′–P regions. Thus, we can evaluate the relative orientations of the β (P–O5′–C5′–C4′) and γ (O5′–C5′–C4′–C3′) torsion angles of type I residues (Table 3). The H4′ and H5′ protons, which are coupled to ^{31}P by 3–6 Hz, do not show coupling constants in DQF-COSY (data not shown). This and the NOESY results mentioned above position the γ angle in the range of 60–90°. According to the observed H–H and H–P coupling patterns, a γ angle range of 60–90° is correlated to a β angle range of 210–180°. While ^{1}H–^{31}P correlations provide much of the information about the type I residues (Table 1), the H4′–P cross peaks of type III residues are not well-resolved, and their H5′–P cross peaks exhibit stronger

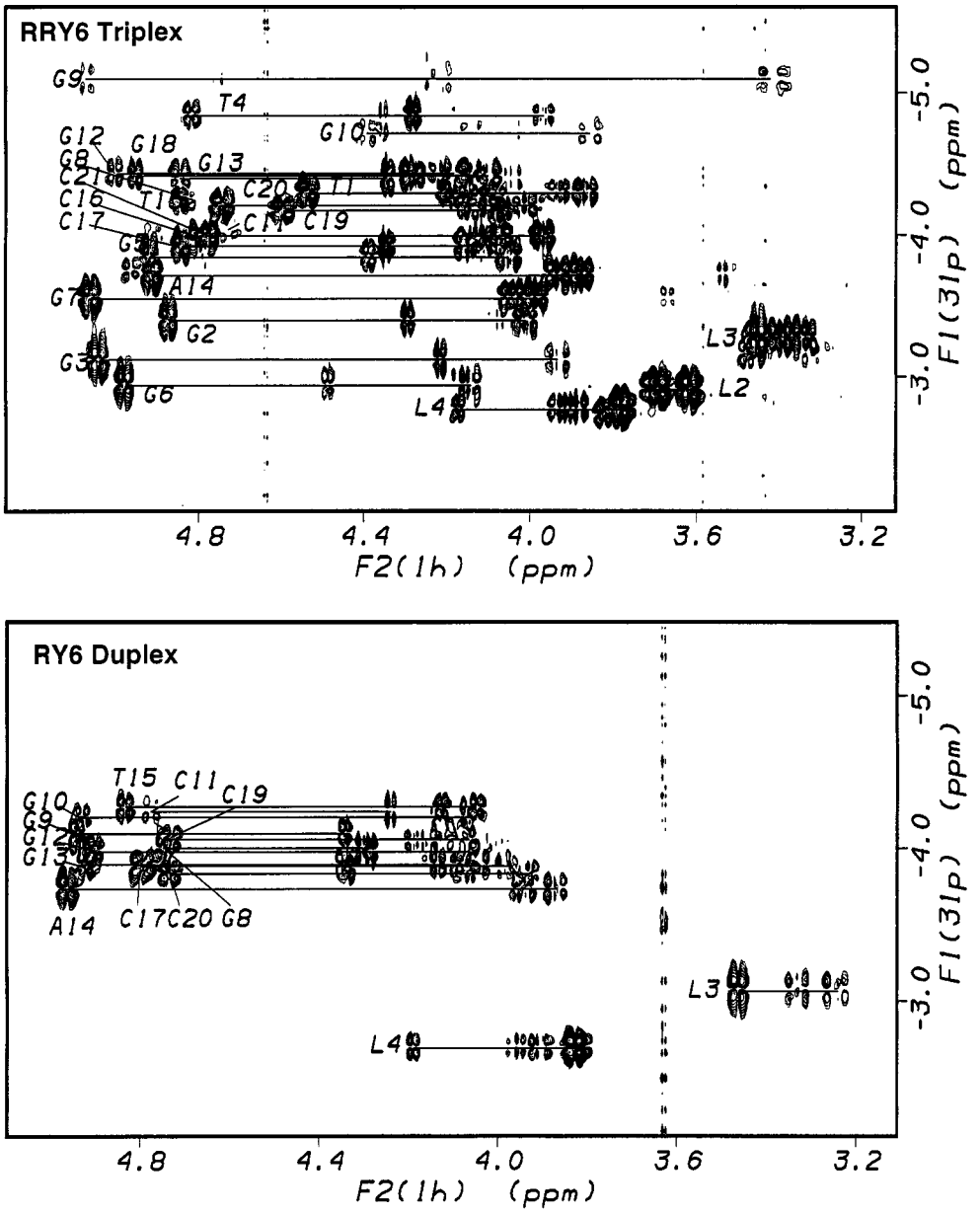

FIGURE 5: ^{1}H–^{31}P correlated spectra of the RRY6 triplex (top) (32 °C, sample conditions are the same as were given for Figure 2) and the RY6 duplex (bottom) (20 °C, pH 6.2). The ^{1}H–^{31}P spectrum of RRY6 collected at 20 °C was essentially identical to that shown here, but the intensities of the cross peaks of G9 and G10 were very weak. Assignments of ^{31}P resonances, including those from linkers, are labeled in the spectra, and their couplings to the sugar H3′, H4′, and H5′ or H5″ protons are shown in solid lines. Except for linker ^{31}P resonances, ^{31}P of the RY6 duplex exhibits much narrower chemical shift dispersion compared to that of the RRY6 triplex.

Table 3: Summary of ^{1}H–^{31}P Correlations Observed from ^{1}H–^{31}P Spectral Analysis[a]

	typeI: G2, G3, G6			type II: G9, G10			type III: others[b]		
2D peak	int	pattern	J_{app}	int	pattern	J_{app}	int	pattern	J_{app}
H3′–P	s	+–	11.0	w	+–	13.7	m	+–	12.4
H4′–P	m	+–	8.4				m	+–	9.1
H5′,H5″–P	m	+–	21.7	w	+–	24.5	s	+–+–	20.3[c]

[a] See Figure 5 for spectral conditions. int is cross-peak intensities: s, strong; m, medium; w, weak. + and – represent positive and negative phases of the multiple peaks. J_{app} is the sum of all coupling constants measured from the ^{1}H dimension. [b] Others: G12, G13, C16, C17, G18, C19, and C20. T4 shows a pattern similar to those of others, but the H5″–P peak showed lower intensity compared to that of the H4′–P peak. [c] Intensity taken from G13. Spectra were extensively overlapped in this region.

intensities (Figure 5 and Table 3). These differences in the spectral coupling patterns of the residues of the third strand and those from the duplex motif are consistent with a noncanonical backbone conformation in the third strand of RRY6.

Unusual NMR Parameters and Local Conformations of the Central T·C·G Triplet. The T4 imino proton H3 of the *T*·C·G triplet appears at 11.05 ppm, which is within the frequency range for non-hydrogen-bonded T imino protons. Under neutral pH conditions, T4(H3), as well as other third strand imino protons, exchanges with H_2O more slowly than exposed terminal imino protons but faster than their duplex counterparts. The third strand imino protons are observable in the spectra of NOESY using gradient echo pulses or jump-return pulses for water suppression (Plateau & Guéron, 1982; Piotto *et al.*, 1992). NOEs correlating T4 with the surrounding residues are summarized in Table 2.

(a) Proton Contacts between T4 and G3 (Table 2a and Figure 4). In addition to the expected sequential connectivities for the third strand sequence, an unusually intense NOE cross peak is detected at the resonance frequencies of the T4(H5) methyl and G3(H1′). Furthermore, T4(H5) is making

Table 4: Recognition and NMR Spectral Patterns of RRY *vs* YRY Triplexes

R·R·Y (*k*, *i*, *j*)[a]	*Y*·R·Y (*k*, *i*, *j*)[a]
third strand antiparallel to the purine duplex strand identified by the $k-(i-1)$ directionality of cross-strand NOEs	third strand parallel to the purine duplex strand identified by the $k-(i+1)$ directionality of cross-strand NOEs
"minor groove" side of the third strand facing away from the major groove of the purine strand; no third strand H1′–duplex base proton contacts detected	"minor groove" side of third strand facing the major groove of the purine strand; third strand H1′–duplex base proton contacts can be detected
third strand T iminos may be shifted slightly upfield; third strand G iminos shifted slightly downfield	third strand T iminos shifted slightly upfield
NOEs involving the third strand:	NOEs involving the third strand:
sequential imino–imino proton connectivities	sequential imino–imino proton connectivities
in the same triple, cross-strand imino–amino protons	in the same triple, cross-strand imino–amino protons
$R(\text{amino})_k$–$Y(\text{amino})_j$; $R(\text{amino})_k$–$Y(\text{amino})_{j-1}$	
$R(\text{imino})_k$–R(base H)$_i$; $R(\text{imino})_k$–R(base H)$_{i-1}$	$Y(\text{imino})_k$–R(base H)$_i$; $Y(\text{imino})_k$–R(base H)$_{i+1}$
$R(\text{imino})_k$–R(H2′,2″)$_i$; $R(\text{imino})_k$–R(H2′,2MM)$_{i-1}$	$Y(\text{imino})_k$–R(H2′,H2MM)$_i$; $Y(\text{imino})_k$–R(H2′,2″)$_{i+1}$

[a] *i*, *j*, and *k* are residue numbers of R, Y, and the third strands (de los Santos *et al.*, 1989; Rajagopal & Feigon, 1989; Radhakrishnan *et al.*, 1991b, 1992b).

contacts with other sugar protons of G3 (Table 2a). This set of NOEs positions the T4 methyl in close vicinity to the sugar of G3 in an orientation distinctly different from that of a canonical helical structure.

(b) Proton Contacts between T4 and G10. An NOE between T4(H5) and G10(H8) is identified (Table 2), whereas the expected proton contacts for a triplex between T4(H6) and G10(H8) are absent.

(c) Proton Contacts between T4 and C11. The identification of NOEs linking T4 with C11 is critical for defining the formation of the *T*·C·G triplet. The *T*·C·G triplet can have three orientations, which can be distinguished by interproton distances (Figure 3). Typically, hydrogen-bond formation through T4(O2) to the C11(H4) non-hydrogen-bonded amino proton (form III) would result in a strong NOE between T4-(H3) and C11(H5). Conversely, hydrogen-bond formation between T4(O4) and C11(H4) (form II) would result in a moderate NOE between T4(H5) and C11(H5) and weak NOEs between T4(H5) and non-hydrogen-bonded C11(H4) and between T5(H5) and C11(H6) (Figure 3). Experimentally, T4 imino proton H3 is relatively closer to the non-hydrogen-bonded C11(H4) amino proton than to the hydrogen-bonded amino proton (Figure 3 and Table 2). Only weak contacts are observed between T4(H3) or the T4(H5) methyl and C11(H5) protons. NOEs linking T4(H3) or T4(H5) with C11(H6) and T4(H5) with the C11(H4) non-hydrogen-bonded proton are too weak to be detected. These observations are inconsistent with the hydrogen-bonding form III, but suggest an alignment similar to that of form I (Figure 3). Our results do not completely rule out an equilibrium between that of forms I and II due to the detection of a weak NOE between T4(H5) and C11(H5). An averaged form (form I&II) is shown in Figure 3.

(d) Proton Contacts between T4 and C17. Previous studies indicated that the third strand in both the RRY and YRY types is close to the duplex purine strand and distant from the duplex pyrimidine strand (Radhakrishnan *et al.*, 1991). The observation of a pair of NOEs between the T4(H3) of the third strand and C17(H4) amino protons of the pyrimidine strand (Table 2) indicates an unusual local conformation that permits the third strand to approach the pyrimidine strand. This analysis is in line with the observation of other T4-related NOEs, suggesting a major distortion at the *T*·C·G site.

A structural model of the central three base triplets is constructed using restrained molecular dynamics based on NOEs discussed above in a–d and sequential interproton connectivities. This model shows the alignment and the local conformation of the *T*·C·G inversion site. T4 is stacked in the third strand helix in an in-plane orientation shown in Figure 4. The unusual proton–proton interactions between T4–C17 and T4–G10 (Table 2) indicate a possible propeller twist accompanied by a rotation of the T4 base away from the helical center (Figure 4).

Triplex and Duplex Comparison. The observation of unusually upfield-shifted base proton resonances of G10 and C11 and sugar protons of G10 in RRY6 (Table 1) prompted us to investigate a free duplex RY6, which contains a sequence identical to the duplex motif of RRY6, including a bis-(hexaethylene glycol) linker between residues A14 and T15 (Figure 1). The combined analysis of NOEs and chemical shifts permits the assessment of structural features unique to triplex formation. Although this information about the direct comparison of a triplex with the related duplex is important for understanding the sequence dependence of DNA recognition, there is a lack of such comparisons in the literature.

The comparison of the relative NOE intensities of base-to-sugar protons, especially base to H2′ and H2″ protons, at a 100-ms mixing time for RY6 and RRY6 (Figure 6) reveals interesting patterns. (a) In general, the observed trends of relative NOE intensities in RY6 and RRY6 are comparable, that is, the intraresidue base to H2′ NOEs are stronger than the NOEs of intraresidue base to H2″ and the interresidue base to H2″. The latter two are comparable in intensity, while the interresidue base to H2′ NOEs are much weaker. In either an A- or a B-form duplex, a base proton is always closer to its own H2′ than to the H2″; a base proton is closer to the adjacent H2′ than to the H2″ in the A-form, and the reverse is true for the B-form duplex. Judging from these basic patterns, the observed base to H2′ and H2″ NOEs in RY6 and RRY6 suggest a deviation in the global structures from typical A- or B-type helices. (b) Interestingly, a breakdown of the continuous pattern is observed at the G10 and C11 sites of RRY6 in contrast to RY6 (Figure 6). The two base protons show NOEs to their own H2′ protons at reduced intensities compared to other intraresidue base to H2′ NOEs and showed no cross peaks to their own H2″. Furthermore, the intensities of the sequential base to H2′ NOEs are comparable to those of the sequential base to H2″ and intraresidue base to H2′ (Figure 6). These results and the large chemical shift differences associated with the G10 and C11 residues (Table 1b) suggest a major conformational variation in the duplex motif of RRY6 that is induced by third strand binding. Since the sugar conformation of G10 and C11 is confined to the C2′-endo region (Table S1, supplementary material), this conformational variation is most likely to be related to the change in base–sugar stacking and/or the torsion angles of the glycosidic linkages and the phosphate backbone. Indeed, there are large phosphate chemical shifts and varied H–P coupling patterns for G9–P–G10 and G10–P–C11 shown in Table 1b and Figure 5. (c)

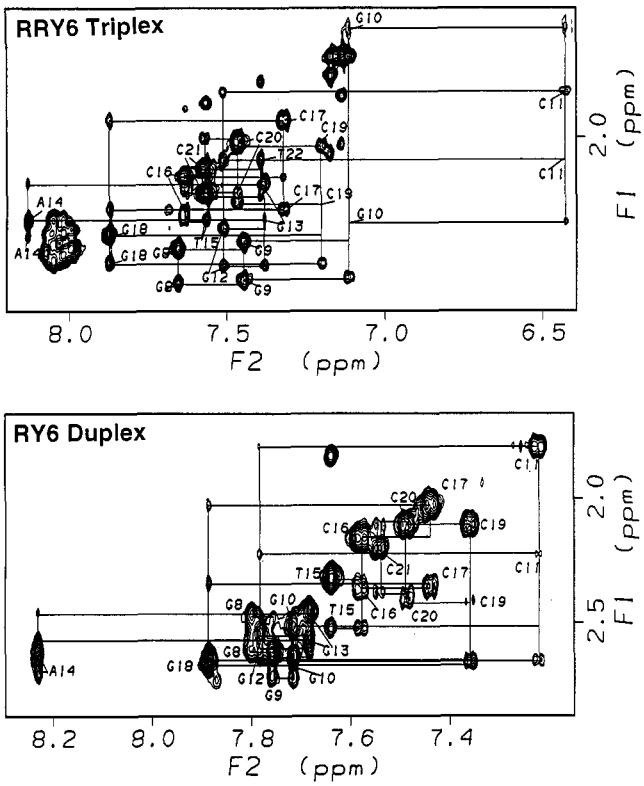

FIGURE 6: 100-ms NOESY spectra of RRY6 (top) (15 °C, for sample conditions, see the legend for Figure 2) and RY6 (bottom) 20 °C, in the presence of 20 mM LiCl and 28 mM $MgCl_2$, pH 6.2) recorded in D_2O. Intraresidue and sequential connectivities of base (6.4–8.1 ppm) to H2′ and H2″ (1.6–2.9 ppm) protons are shown by solid lines and residue numbers.

The sequential NOEs connecting base to H2′ and H2″ protons at the T–G steps of the third strand are much weaker than their counterparts in the duplexes, while the related intraresidue NOEs are poorly resolved (Figure 6). This observed NOE pattern of the third strand of RRY6 reflects the feature of the helical conformation of the third strand that differs from canonical duplex helices.

Although qualitative, the observed chemical shift differences between RY6 and RRY6 are indicators of the perturbations in local environment and conformation upon triplex formation. (a) The chemical shift differences are most evident at the purine sequence, to which the third purine strand binds (Table 1b). These large shifts are centered around the *T*·C·G site and gradually diminish at the third base triplets on either side of the inversion site. (b) The resonances of G10 in the G5·G10·C19 triplet exhibit larger chemical shift differences compared to those of G12 in the G3·G12·C17 triplet. Moreover, the observed NOE patterns (Table 2, cross-strand NOEs of third strand imino to base and sugar protons) show more pronounced spectral irregularities at the G5·G10·C19 triplet compared to the G3·G12·C17 triplet. Noting that the G3·G12·C17 triplet flanks the 5′-direction of the *T*·C·G inversion site and that the G5·G10·C19 triplet is on the immediate 3′-side, these results are consistent with a sequence-dependent effect on structural perturbation. (c) The amino protons of C residues in the pyrimidine strand of RRY6 exhibit large downfield shifts (Table 1a,b). This trend is not observed in YRY triplexes, which show comparable C amino proton chemical shifts in the triplex and duplex (Radhakrishnan *et al.*, 1989). The C amino protons are directly involved in Watson–Crick as well as reverse Hoogsteen base pairing. Thus, the observed downfield shift of the triplex amino resonances indicates structural and/or hydrogen-bonding features unique to the RRY6 triplex.

Comparison with Literature. Regular RRY triplexes consisting of *T*·A·T, *A*·A·T, and *G*·G·C base triplets have been well-characterized by NMR and structural calculations (Pilch *et al.*, 1990; Radhakrishnan *et al.*, 1991b, 1993; Cheng *et al.*, 1992). A distinct feature of T and G mixed third strand is the increase in base rise at the T–G step to an average value of 5.6 Å. In the NOESY spectrum, no sequential NOEs between these residues are observed and the H1′ resonances for the G residues are unusually upfield-shifted (Radhakrishnan *et al.*, 1993). NMR spectra of RRY6 exhibit a similar pattern: no sequential NOEs are detected between the T1–G2 and T4–G5 steps, and only moderately upfield-shifted H1′ resonances from G2 and G5 residues are observed (Table 1). These results indicate that there is sufficient structural similarity between the unperturbed RRY triplex and the *T*·C·G-containing RRY6. On the other hand, the inclusion of the *T*·C·G triplet is clearly accompanied by structural perturbation at adjacent sites. This is supported by the facts that in RRY6 there are fewer and weaker third strand imino-to-duplex sugar H2′ and H2″ NOEs and imino-to-imino NOEs and that the third strand imino protons resonate at lower field compared to their counterparts reported in the literature (Radhakrishnan *et al.*, 1991b, 1993). In RRY6, a set of altered chemical shifts is observed in the sugar H2′ and H2″ protons of all third strand G residues, with H2′ protons appearing downfield relative to H2″ protons (H2′ exhibits larger *J* constants to H1′, and H2″ is closer to H1′ in short mixing time NOESY). The inverted chemical shifts in the third strand G residues have been observed in another RRY triplex (Radhakrishnan & Patel, 1993) and may be a general characteristic typical of the third strand G residues in antiparallel triplexes. The underlined course of the inversion of the H2′ and H2″ chemical shifts remains to be investigated.

The opposite strand orientation of the antiparallel RRY and parallel YRY triplexes gives rise to a set of distinctively different NMR spectral features, which is summarized in Table 4. These spectral patterns serve as markers for the identification of strand orientation, base triplet alignments, and interstrand interactions in RRY or YRY triplexes. Typically, the antiparallel orientation of RRY triplexes results in cross-strand NOEs from the third strand imino H_k to duplex base proton of the $i - 1$ residue. In the parallel YRY triplex, the third strand imino H_k exhibits NOEs to the base proton of the $i + 1$ residue (Radhakrishnan *et al.*, 1992b). RRY triplexes adopt a conformation with the third strand minor groove side facing away from the major groove of the purine strand, and thus, no NOE from the third strand H1's and duplex base protons can be detected. In contrast, YRY triplexes adopt a conformation with the third strand minor groove side facing the major groove of the purine strand, and thus, NOEs from the third strand H1's and duplex base protons can be detected.

CONCLUSION

NMR analysis reveals the general alignment of the *T*·C·G inverted triplet flanked by *G*·G·C triplets and a set of molecular interactions that is responsible for the stabilization of triplex formation. A sequence-dependent perturbation around the *T*·C·G site is described, with the *G*·G·C triplet well-defined on the 5′-side of T4 but somewhat interrupted on the 3′-side of T4. The spectral comparison of the RY6 duplex and the RRY6 triplex reveals a structural impact on the duplex around the *T*·C·G inversion site. The third purine strand exhibits unique NMR patterns, such as the large chemical shift dispersion of ^{31}P resonances and their characteristic H–P couplings, and the close contacts between the T4 methyl and

imino protons and adjacent G3, G10, C11, and C17 residues. These results suggest that the DNA binding oligonucleotides possess unique structural features that are different from those of the canonical DNA helices. Understanding the recognition of a C·G inverted base pair in an otherwise homo-RY duplex is important for designing nucleotide analogs that can specifically bind to such inversion sites. We are in the process of incorporating NMR results into molecular dynamics simulations for the elucidation of a three-dimensional structural model of the RRY6 triplex. This should provide insight into the molecular basis for the *T*·C·G triplet formation and lead to the construction of a general model for specific recognition of the C·G inversion in hetero-RY duplexes.

ACKNOWLEDGMENT

The authors thank Dr. Kenneth Johnson for initial technical help, Dr. Montgomery Pettitt and Jason Cheng for many stimulating discussions, and Dr. Jeffrey Rice for careful reading of the manuscript.

SUPPLEMENTARY MATERIAL AVAILABLE

Table giving the *J* couplings of the RRY6 triplex and the RY6 duplex and figure containing spectra of RRY6 as a function of [Mg^{2+}] (3 pages). Ordering information is given on any current masthead page.

REFERENCES

Asseline, U., & Thuong, N. T. (1993) *Tetrahedron Lett. 34*, 4173–4176.
Beal, P. A., & Dervan, P. B. (1992a) *Nucleic Acids Res. 20*, 2773–2776.
Beal, P. A., & Dervan, P. B. (1992b) *J. Am. Chem. Soc. 114*, 4976–4982.
Cheng, Y. K., & Pettittt, B. M. (1992) *J. Am. Chem. Soc. 114*, 4465–4474.
Chubb, J. M., & Hogan, M. E. (1992) *Tibtech 10*, 132–136.
Cooney, M., Czernuszewicz, G., Postel, E., Flint, S. J., & Hogan, M. E. (1988) *Science 245*, 456–459.
de los Santos, C., Rosen, M., & Patel, D. (1989) *Biochemistry 28*, 7282–7289.
Derome, A. (1987) *Modern NMR techniques for chemistry research*, Pergamon Press, New York.
Durlan, R. H., Rao, T. S., Revankar, G. R., Tinsley, J. H., Myrick, M. A., Seth, D. M., Rayford, J., Singh, P., & Jayaraman, K. (1993) *Nucleic Acids Res.* (submitted for publication).
Gao, X., & Patel, D. J. (1989) *Q. Rev. Biophys. 22*, 751–762.
Giessner-Prettre, C., Pullman, B., Cheng, D. M., Iuorno, V., & T'so, P. O. P. (1984) *Biopolymers 23*, 377–388.
Gorenstein, D. G. (1984) *Phosphorus-31 NMR. Principles and applications*, Academic Press, New York.
Gorenstein, D. G., Schroeder, S. A., Fu, J. M., Metz, J. T., Roongta, V., & Jone, C. R. (1988) *Biochemistry 27*, 7223–7237.
Griffin, L. C., & Dervan, P. B. (1989) *Science 245*, 967–971.
Krawczyk, S. H., Milligan, J. F., Wadwani, S., Moulds, C., Froehler, B. C., & Matteucci, M. D. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 3761–3764.
Lankhorst, P. P., Haasnoot, C. A. G., Erkelens, C., & Altona, C. (1984) *J. Biomol. Struct. Dyn. 1*, 1387–1405.
Le Doan, T., Perrovault, L., Praseuth, D., Habhoub, N., Decout, J., Thuong, T., Lhomme, J., & Hélene, C. (1987) *Nucleic Acids Res. 15*, 7749–7760.
Macaya, R. F., Schultze, P., & Feigon, J. (1992) *J. Am. Chem. Soc. 114*, 781–783.
Majumdar A., & Hosur, R. V. (1992) *Prog. Nucl. Magn. Reson. Spectrosc. 24*, 109–158.
McBride, L. J., & Caruthers, M. H. (1983) *Tetrahedron Lett. 24*, 245–248.
Mergny, J. L., Sun, J. S., Rougee, M., Montenay-Garestier, T., Barcelo, F., Chomilier, J., & Hélene, C. (1991) *Biochemistry 30*, 9791–9798.
Mohan, V., Cheng, Y.-K., Marlow, G., & Pettitt, B. M. (1993) *Biopolymers* (in press).
Mooren, M. M., Pulleyblank, D. E., Wizmenga, S. S., Bloomers, M. J., & Hilbers, C. W. (1990) *Nucleic Acids Res. 18*, 6523–6529.
Moser, H. E., & Dervan, P. B. (1987) *Science 238*, 645–650.
Ono, A., T'so, P. O. P., & Kan, L. S. (1991) *J. Am. Chem. Soc. 113*, 4032–4033.
Pilch, D. S., Levenson, C., & Shafer, R. H. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 1942–1946.
Piotto, M., Saudek, V., & Sklenar, V. (1992) *J. Biol. Nucl. Magn. Reson. 2*, 661–665.
Plateau, P., & Guéron, M. (1982) *J. Am. Chem. Soc. 104*, 7310–7311.
Plum, G. E., Park, Y. W., Singleton, S. F., Dervan, P. B., & Breslauer, K. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9436–9440.
Radhakrishnan, I., & Patel, D. J. (1993) *Structure 1*, 135–152.
Radhakrishnan, I., Gao, X., de los Santos, C., Live, D., & Patel, D. J. (1991a) *Biochemistry 30*, 9022–9030.
Radhakrishnan, I., de los Santos, C., & Patel, D. J. (1991b) *J. Mol. Biol. 221*, 1403–1418.
Radhakrishnan, I., Patel, D., Veal, J. M., & Gao, X. (1992a) *J. Am. Chem. Soc. 114*, 6913–6915.
Radhakrishnan, I., Patel, D., & Gao, X. (1992b) *J. Am. Chem. Soc. 114*, 2514–2523.
Rajagopal, P., & Feigon, J. (1989) *Nature 339*, 637–640.
Roberts, R. W., & Crothers, D. M. (1992) *Science 258*, 1463–1465.
Sen, D., & Gilbert, W. (1988) *Nature 334*, 364–366.
Sklenar, V., & Feigon, J. (1990) *Nature 345*, 836–838.
Sklenar, V., Miyashiro, H., Zon, G., Miles, H. T., & Bax, A. (1986) *FEBS Lett. 208*, 94–98.
Stilz, H. U., & Dervan, P. B. (1993) *Biochemistry 32*, 2177–2185.
T'so, P. O. P., Aurelian, L., Chang, E., & Miller, P. S. (1992) *Antisense Strategies 660*, 159–177.
Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M., & Wohlrab, F. (1988) *FASEB J. 2*, 2939–2949.
Yoon, K., Hobbs, C. A., Koch, J., Sardaro, M., Kutny, R., & Weis, A. L. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 3840–3844.