Solution Structure of an Intramolecular DNA Triplex Containing an N^7-Glycosylated Guanine Which Mimics a Protonated Cytosine[†,‡]

Karl M. Koshlap,[§] Peter Schultze,[§] Helmut Brunar,[||] Peter B. Dervan,*[,||] and Juli Feigon*[,§]

Department of Chemistry and Biochemistry and Molecular Biology Institute, University of California, Los Angeles, California 90095, and Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Received September 27, 1996; Revised Manuscript Received November 26, 1996[⊗]

ABSTRACT: The three-dimensional structure of a pyrimidine−purine−pyrimidine DNA triplex containing an N^7-glycosylated guanine (^{7}G) in the third strand has been determined by NMR spectroscopy, restrained molecular dynamics calculations, and complete relaxation matrix refinement. Glycosylation of the guanine at the N7 position permits it to adopt a conformation such that the Hoogsteen face of the base mimics the arrangement of hydrogen bond donors seen in protonated cytosine. The NMR data confirm the previously proposed hydrogen bonding scheme of the ^{7}G·G·C triplet. The three-dimensional structure of the triplex accommodates the ^{7}G with less distortion of the phosphodiester backbone than would be required for an N^9-glycosylated guanine in the same sequence position, but some changes in the positions of the phosphodiester backbone are present compared to a C^+·G·C triplet. The structure provides a rationale for the observations that ^{7}G binds to Watson−Crick G·C base pairs with higher specificity and affinity than guanine, but with a lower stability at pH 5.2 than would be provided by a canonical C^+·G·C triplet.

DNA triplexes are formed when an oligodeoxynucleotide of the appropriate sequence binds in the major groove of homopurine−homopyrimidine tracts of DNA (Feigon *et al.*, 1995; Moser & Dervan, 1987; Sun & Hélène, 1993; Thuong & Hélène, 1993; Wells *et al.*, 1988). There are two major classes of DNA triplexes, both of which involve recognition of the duplex homopurine strand by the additional oligodeoxynucleotide. The two types of triplexes differ both in the orientation of the third strand and in the nature of the triplets formed. In one class of triplexes, the additional strand is oriented antiparallel to the Watson−Crick purine strand. Sequence specificity is achieved by the formation of reverse-Hoogsteen hydrogen bonds between G and G·C base pairs, and either A or T and A·T base pairs, resulting in G·G·C, A·A·T, or T·A·T triplets, respectively. These are often referred to as purine−purine−pyrimidine or purine motif triplexes, despite the inclusion of T·A·T triplets (Beal & Dervan, 1991; Broitman *et al.*, 1987; Cooney *et al.*, 1988; Dayn *et al.*, 1992; Kohwi & Kohwi-Shigematsu, 1988; Lipsett, 1963; Michel *et al.*, 1992; Radhakrishnan *et al.*, 1991; Thuong & Hélène, 1993). In the other class, usually referred to as pyrimidine−purine−pyrimidine or pyrimidine motif triplexes, the third strand orientation is parallel to the central purine strand. In this case, recognition is accomplished by the formation of Hoogsteen hydrogen bonds between T and A·T base pairs or C^+ and G·C base pairs, resulting in T·A·T and C^+·G·C triplets, respectively (de los Santos *et al.*, 1989; Hampel *et al.*, 1993; Lipsett, 1964; Mirkin *et al.*, 1987; Moser & Dervan, 1987; Rajagopal & Feigon, 1989a,b; Thuong & Hélène, 1993; Wells *et al.*, 1988). Noncanonical or mismatch triplets are also possible (Belotserkovskii *et al.*, 1990; Griffin & Dervan, 1989; Yoon *et al.*, 1992) but destabilize the triplex to a greater or lesser extent depending on nearest neighbors. With the ability to discriminate mismatches at a level comparable to that observed in duplex DNA (Best & Dervan, 1995; Greenberg & Dervan, 1995; Mergny *et al.*, 1991; Roberts & Crothers, 1991), triplex formation can be directed at a single site in genomic DNA (Strobel & Dervan, 1991; Strobel *et al.*, 1991). The remarkable selectivity demonstrated by even relatively short oligonucleotides has stimulated a great deal of interest. Potential applications include their use as antigene therapeutics (Beaucage & Iyer, 1993; Cooney *et al.*, 1988; Hélène & Toulme, 1990; Lu & Ferl, 1993; Maher *et al.*, 1989; Nagel *et al.*, 1993; Prins *et al.*, 1993) or as tools in chromosome mapping (François *et al.*, 1989; Strobel *et al.*, 1988, 1991).

The necessity that any cytosine in the third strand be protonated at N3 leads to a pH dependence of triplexes containing C^+·G·C triplets. In addition, repulsion between adjacent protonated cytosines destabilizes triplexes (Lee *et al.*, 1984). Many attempts have been directed at circumventing this unwanted dependence on acidic pH. For example, replacement of the proton at position 5 of cytosine with an alkyl group, e.g., methyl, facilitates protonation of N3 and results in greater stability at near physiological pH (Povsic & Dervan, 1989; Singleton & Dervan, 1992; Xodo *et al.*, 1991). More recent strategies have involved the design of nonnatural bases or base analogs that duplicate the hydrogen bonding characteristics of a protonated cytosine (Davison & Johnsson, 1993; Froehler & Ricca, 1992; Froehler *et al.*, 1992; Huang & Miller, 1993; Koh & Dervan, 1992; Krawczyk *et al.*, 1992; Miller *et al.*, 1992; Ono *et al.*, 1991; Xiang *et al.*, 1994; Young *et al.*, 1991).

It was recently demonstrated that when the glycosylated purine 7-(2-deoxy-β-D-*erythro*-pentofuranosyl)guanine (^{7}G) is incorporated into a pyrimidine oligodeoxynucleotide, it

[†] This work was supported by NIH Grant GM 37254 to J.F. and ONR Grant N00014-95-1075 to P.B.D.

[‡] Coordinates for these structures have been deposited in the Brookhaven Protein Data Bank (entry 1GN7).

[§] University of California.

[||] California Institute of Technology.

[⊗] Abstract published in *Advance ACS Abstracts,* February 1, 1997.

S0006-2960(96)02438-5 CCC: $14.00

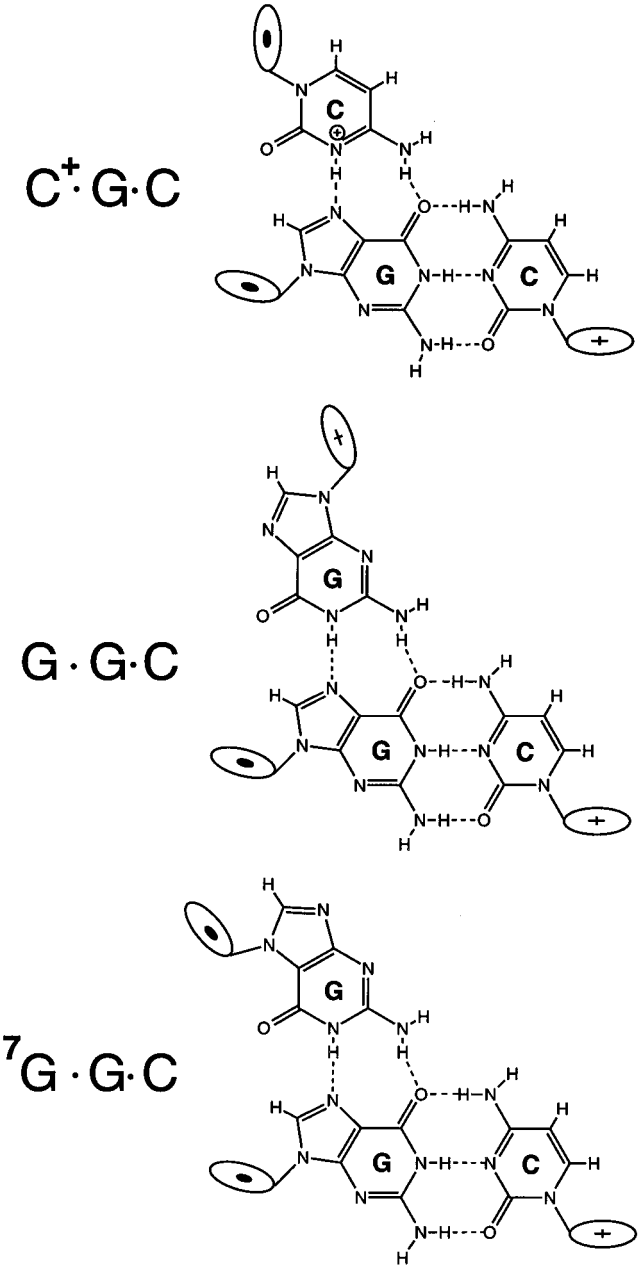

FIGURE 1: Schematic of C^+·G·C, pyrimidine motif (top), G·G·C, purine motif (middle), and ^{7}G·G·C, pyrimidine motif (bottom), triplets. Deoxyriboses are indicated by ellipses, and the relative phosphodiester backbone orientation is indicated by arrowheads (dots) and tails (crosses).

binds to Watson−Crick G·C base pairs with both high affinity and sequence specificity (Hunziker *et al.*, 1995). Attachment of the deoxyribose to the N7 position of the guanine base orients it such that the groups involved in hydrogen bond formation mimic those seen in both reverse Hoogsteen−Watson−Crick G·G·C triplets or Hoogsteen−Watson−Crick C^+·G·C triplets (Figure 1). Changing the point of attachment of the deoxyribose from the N9 to the N7 position should have important advantages with respect to the regularity of the phosphodiester backbone. Hélène and co-workers (Giovannangeli *et al.*, 1992; Sun *et al.*, 1991) have pointed out that a normal ^{9}G in the third strand can also bind to a G·C base pair in the Hoogsteen mode. However, the C1′ position of the third strand G can be expected to deviate significantly from the C1′ position of third strand cytosines or thymines. Thus the stability and orientation of a triplex with a third strand composed of

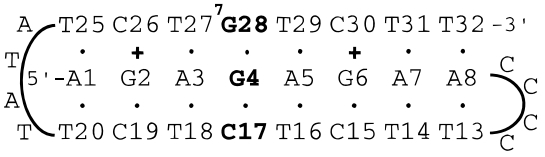

FIGURE 2: Folding scheme of the intramolecular triplex 5′-AGA**G**AGAACCCC TTCT**C**TCTTATATCT^7**G**TCTT (^{7}GGC triplex) containing the ^{7}G·G·C triplet used in these studies. The CCCC bases form the loop connecting the Watson−Crick purine and pyrimidine strands, while the TATA bases form the loop connecting the two pyrimidine strands.

guanines and thymines may be dependent on the number of GpT and TpG steps. In order to investigate the base pairing and structural effects upon the incorporation of ^{7}G into the Hoogsteen strand of a DNA triplex, we synthesized the DNA oligonucleotide 5′-AGA**G**AGAACCCCTTCT**C**TCTTATA-TCT^7**G**TCTT (loops are underlined and letters in bold are bases in the ^{7}G·G·C triplet), which is designed to fold into an intramolecular triplex (^{7}GGC triplex) (Figure 2). Using NMR spectroscopy, we show that this sequence forms a triplex as designed. We present the evidence for the proposed hydrogen bonding pattern for the ^{7}G·G·C triplet and the refined three-dimensional solution structure of the triplex.

MATERIALS AND METHODS

Sample Preparation. The phosphoramidite derivative of ^{7}G and the oligodeoxynucleotide 5′-AGAGAGAACCCCT-TCTCTCTTATATCT^{7}GTCTT were synthesized as previously described (Hunziker *et al.*, 1995). To remove triethylammonium acetate present from the FPLC purification protocol, the sample was chromatographed on Sephadex G15 and Bio-Rad AG 50W-X8 (Na^+ form) ion-exchange columns. The desired fractions were pooled and lyophilized to dryness. The final conditions of the NMR sample were 100 mM NaCl, 5 mM $MgCl_2$, and ∼2 mM DNA in 400 μL of either 99.999% D_2O or 90% H_2O/10% D_2O. The pH was adjusted to 6.0 with HCl/NaOH.

NMR Spectroscopy for Assignments. The NMR data were acquired at 500 MHz on General Electric GN500, Bruker AMX 500, or Bruker DRX500 spectrometers. Quadrature detection in the indirect dimension was accomplished by the method of States *et al.* (1982) on the General Electric spectrometer and with the States−TPPI mode (Marion *et al.*, 1989) on the Bruker spectrometers.

For assignment purposes, 2D NOESY experiments (Kumar *et al.*, 1980; Macura & Ernst, 1980) were obtained on the sample in D_2O at temperatures ranging from 20 to 40 °C and mixing times of 70−350 ms. At least 325 points were collected in t_1, 32 scans were acquired per increment, and the recycle delay was 2 s. For these and the following homonuclear 2D experiments, a spectral width of 5000 Hz was used in both dimensions, and 1024 points were collected in t_2. HOENOE spectra (Sklenář & Feigon, 1990b) were acquired at 30 and 35 °C with selective excitation of the cytosine H5 protons. Coherence transfer was accomplished by using an MLEV sequence, followed by the regular NOESY t_1 evolution period with NOESY mixing times of either 250 or 350 ms. HOHAHA spectra (Bax & Davis, 1985; Braunschweiler & Ernst, 1983) were acquired using the MLEV mixing sequence, 1.5 ms trim pulses for the spin lock, and isotropic mixing times of 50 and 100 ms at temperatures ranging from 20 to 40 °C. A minimum of 320

points were collected in t_1, with either 48 or 64 scans per t_1 block. P. COSY spectra (Marion & Bax, 1988; Mueller, 1987) with mixing pulses of 90° were acquired at 20, 30, and 35 °C. At least 550 points with 24 scans per block were collected in t_1. A $^1H-^{13}C$ HMQC spectrum (Bax *et al.*, 1983) was acquired with sweep widths of 5000 Hz in the ^1H dimension and 7545 Hz in the ^{13}C dimension. The size of the spectrum was 512 points in F_2, and 128 points in F_1, with 128 scans averaged per block.

One-dimensional spectra of the sample in 90% H_2O/10% D_2O were obtained with the $1\bar{1}$-spin echo pulse sequence as a function of temperature. The null was positioned at the frequency of the H_2O resonance, and the excitation maximum was centered in the imino region. 2D NOESY spectra of the sample in H_2O were obtained at 1, 5, and 10 °C. Water suppression was achieved using a $1\bar{1}$-spin echo read pulse as above. A spectral width of at least 12 048 Hz was used in both dimensions, either 1024 or 2048 points were collected in t_2, a minimum of 141 points with 64−160 scans per increment were collected in t_1, and the mixing times were either 100 or 150 ms.

NMR Spectra Used in Structure Calculations. NOE restraints for the nonexchangeable resonances were obtained from analysis of a series of phase-sensitive NOESY spectra (Kumar *et al.*, 1980; Macura & Ernst, 1980) with mixing times of 40, 80, 140, 200, and 260 ms which were obtained without removing the sample from the magnet. Except for the mixing times, the acquisition parameters were identical in all cases. Spectra were acquired with a sweep width of 5000 Hz in both dimensions, 1024 points in t_2, 350 points with 32 scans per block in t_1, and a recycle delay of 2 s, during which time the residual HDO signal was suppressed by irradiation. All spectra of the series were processed identically, with the t_2 apodization a 450 point, 55° phase shifted squared sine bell with first degree polynomial baseline flattening. The apodization in t_1 was a 350 point, 55° phase shifted squared sine bell with second degree polynomial baseline flattening. The spectra were zero-filled to 2048 by 2048 complex points to provide adequate digital resolution for integration of the cross-peak volumes. Processing was done on Silicon Graphics workstations using Felix 2.3 (Biosym Technologies).

The NOE restraints involving the exchangeable resonances were obtained from a NOESY spectrum of the sample dissolved in H_2O at 1 °C with a $\tau_m = 100$ ms. Solvent suppression was achieved using a $1\bar{1}$-spin echo read pulse (Sklenář & Bax, 1987). The carrier was centered on the water resonance, with the excitation maximum at the hydrogen-bonded imino region.

To obtain coupling constants of the deoxyribose protons, a P.COSY (Marion & Bax, 1988; Mueller, 1987) spectrum was collected at 35 °C with a 90° mixing pulse. The acquisition parameters were a sweep width of 5000 Hz in both dimensions, 1024 points in t_2, 550 points in t_1, 96 scans per t_1 block, and a recycle delay of 1.8 s. To subtract the dispersive component of the diagonal, a one-dimensional reference spectrum was collected, as described by Marion and Bax (1988). The same acquisition parameters were used as for the two-dimensional experiment, except that 1536 scans were collected, and the acquisition time was doubled. An additional delay equal to the acquisition time was added at the end of the P.COSY pulse sequence so that the total time for presaturation of the HDO signal and for relaxation would be equal to that of the reference spectrum. The apodization function in t_2 was a 700 point, 35° phase shifted squared sine bell, while that in t_1 was a 550 point, 60° phase shifted squared sine bell. The spectrum was zero-filled to 4096 points in both dimensions. Coupling constants were extracted from the spectrum by automatic optimization of the agreement between simulated and experimental H1′−H2′,H2″ cross peaks using the FORTRAN program CHEOPS (P. Schultze and J. Feigon, unpublished program) as previously described (Macaya *et al.*, 1992a,b). These coupling constants were then used with the program PSEUROT (de Leeuw & Altona, 1983) to determine the sugar conformations and to generate restraints on the dihedral angles ν_1 and ν_2.

Structure Determination and Analysis. The distance restraints between nonexchangeable protons that were used for the initial structure calculations were obtained from the 260 ms mixing time NOESY experiment. Cross peaks were picked and integrated using the AURELIA software package (Neidig *et al.*, 1995). The resulting NOESY peak lists were used for a semiautomatic assignment procedure. An expanded peak list was generated containing all possible assignments closer than 0.01 ppm in each dimension from the chemical shift list. As an additional decision criterion, the expanded list also contained the corresponding H−H distance from a preliminary model for each possible cross-peak assignment. This expanded peak list was checked by hand and then reimported into AURELIA for generation of distance and peak integral lists in X-PLOR format. The distances from the D_2O spectra were calibrated by setting the strongest observed H2′−H2″ cross peak to 1.9 Å. All lower limits were set to 2.0 Å, and the upper limits were set to the calibration distance plus a margin of 1.0 Å, except in the case of methyl groups, where an addition of 2.5 Å was made.

Quantification of cross-peak volumes for exchangeable proton resonances is not as accurate due to both the differing rates at which the resonances exchange with water and the $\sin^3$ excitation profile of the $1\bar{1}$-spin echo read pulse. Therefore, for the distances involving exchangeable protons all upper limits were set to 5 Å and the lower bound was 2.0 Å.

Repulsive NOE restraints were systematically derived to further refine structures that contained close proton−proton contacts but whose corresponding cross peaks were not observed in the experimental spectra. A list of all nonexchangeable proton−proton distances less than 8 Å was computed for the structure with lowest overall energy. This list, containing 6350 entries, was then combined with the chemical shift list and formatted as a peak list. Using AURELIA, this peak list was projected onto the 260 ms NOESY spectrum. Single data point intensities at each peak position were measured and written out as a list. The peak intensity list was sorted according to intensities, and all peaks below noise level, 360 in total, were combined and reformatted in X-PLOR format as repulsive distance restraints with a lower limit of 5 Å.

In the initial step of X-PLOR Version 3.1 (Brünger, 1992) calculations, 20 coordinate sets were obtained by metric matrix distance geometry embedding of all atoms. In the next step the initial structures were subjected to the distance geometry−simulated annealing (DGSA) procedure (Nilges *et al.*, 1988), consisting of template fitting to improve local geometry and simulated annealing. To improve the number of converging structures, the weight of the bond angle potential was increased 8-fold over the standard X-PLOR

force field, and dihedral angle restraints involving the four ligands of each chiral center were introduced to enforce the correct chiral configurations. In addition, the ceiling of the energy term for the distance restraints was set to 100, i.e., $^1/_{10}$ of the value in the published (Brünger, 1992) protocols. These changes led to an improved ratio of over 90% of acceptable structures compared to 50−60% obtained previously with the standard parameters. In the third step those modified parameters were returned to their standard values for further refinement by molecular dynamics and energy minimization based on the distance restraints. This was followed by the final relaxation matrix refinement, taking into account the peak intensities of the nonexchangeable protons. The geometry of the base triplets was restrained with standard hydrogen bond distances. In addition, planarity restraints with a low weight factor of 3 kcal/Å^2 were applied to each triplet at all refinement steps. This favors overall planarity of the triplets if no specific distance restraints cause out-of-plane tilting of particular bases.

All 20 initial structures were carried through to the third step of refinement. At this point the coordinate sets were sorted according to overall energy terms. The first eight of these coordinate sets were subjected to relaxation matrix refinement consisting of a short simulated annealing protocol under inclusion of the NOE potential. The nonexchangeable (^2H_2O) and repulsive restraints were converted to cross-peak volumes for the spectra at all five mixing times, while the exchangeable (H_2O) and hydrogen bond restraints were left as distances.

The helical parameters were calculated from the duplex portion of the triplex structures using the program CURVES (Lavery & Sklenar, 1988).

UV Spectroscopy. The melting behavior of the ^{7}GGC triplex and of 5′- AGA**G**AGAA CCCCTTCT**C**TCTTATA-TCT**C**TCTT (C^+GC triplex) was studied by using a Cary 1E UV−visible spectrophotometer to observe the absorbance at 260 nm as a function of temperature. The heating rate was 0.5 °C/min. The samples were 0.66−0.92 OD (260 nm) DNA, 100 mM NaCl, 5 mM $MgCl_2$, and 50 mM sodium acetate, pH 5.2.

RESULTS

^{1}H NOESY Spectrum of the Exchangeable Resonances of the ^{7}G·G·C Triplex. Information concerning the exchangeable protons of the ^{7}G·G·C triplex was obtained from NOESY spectra of the sample in H_2O. The imino−imino region of a NOESY spectrum of the sample in H_2O taken at 1 °C is shown in Figure 3a. Sequential imino−imino connectivities along the Watson−Crick and Hoogsteen paired strands are observed for all of the base pairs including the ^{7}G residue. The presence of these connectivities clearly indicates that the sequence folds into a triplex as designed. The imino−amino, aromatic region of the same spectrum is shown in Figure 3b. The cross peaks between the imino proton resonances and the amino and aromatic proton resonances characteristic of the particular type of triplet are present (Feigon *et al.*, 1995). Downfield-shifted C^+ amino protons, typical of triplexes containing third strand cytosines, are also present. The imino and amino resonances were assigned as previously described (Macaya *et al.*, 1991, 1992b; Rajagopal & Feigon, 1989a; Sklenář & Feigon, 1990a) and are given in Table 1.

The fact that the ^{7}G imino is included in the sequential imino−imino connectivities provides evidence for formation

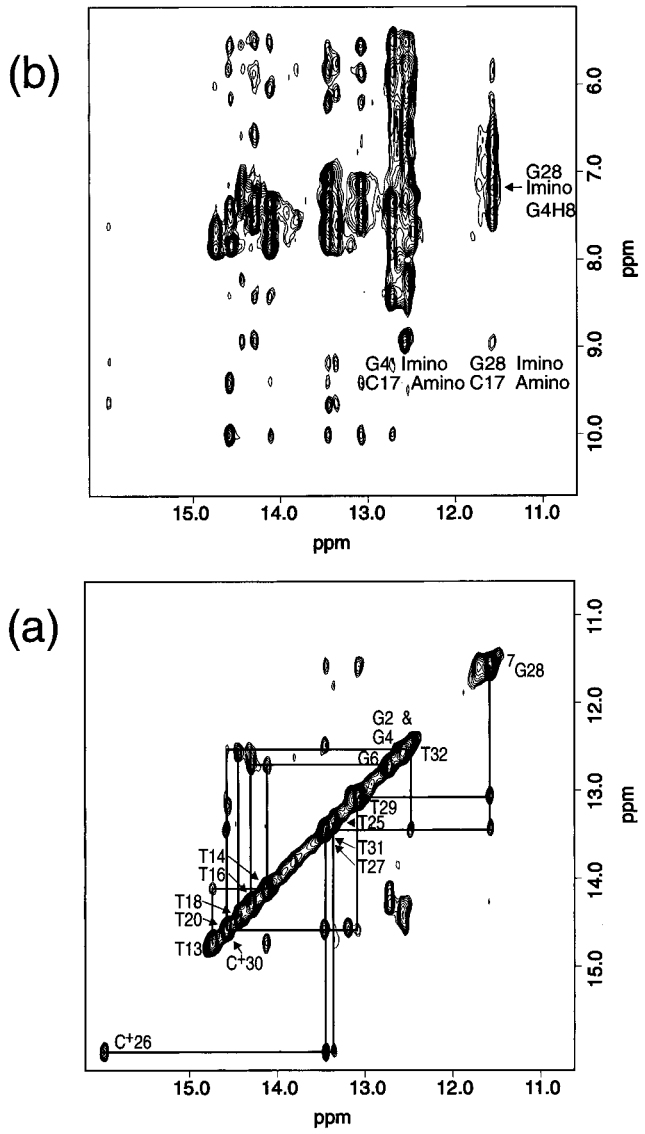

FIGURE 3: NOESY spectrum of the ^{7}GGC triplex at 1 °C and τ_m= 150 ms. (a) Imino−imino region. Sequential connectivities are observed along the entire lengths of the Watson−Crick bonded strands (indicated above the diagonal) and the Hoogsteen bonded strand (indicated below the diagonal), including the N^7-glycosylated guanine, G28. (b) Imino−amino, aromatic region. The cross peaks between the ^{7}G28 imino proton and both the G4H8 and the C17 amino protons and between the G4 imino proton and the C17 amino proton are indicated.

of the proposed ^{7}G·G·C triplet within the triple helix. The bonding scheme is definitively established by a number of cross peaks occurring in the H_2O NOESY spectrum. These include a strong cross peak between the ^{7}G28 imino proton and the H8 proton of G4. A weak cross peak between the ^{7}G imino proton and the hydrogen-bonded amino proton of C17 is also observed, which is probably the result of spin diffusion *via* the G28 amino proton(s). In the Watson−Crick G4·C17 base pair of this triplet, a strong cross peak is seen between the G4 imino proton and the hydrogen-bonded amino proton of C17. These are all consistent with the base triplet as previously proposed (Figure 1). Due to the broad line widths of the rapidly exchanging guanine amino resonances, it was not possible to assign any specific NOE interactions to these resonances.

^{1}H NOESY Spectrum of the Nonexchangeable Resonances of the ^{7}G·G·C Triplex. Figure 4 shows the H1′, H5− aromatic region of a NOESY spectrum of the ^{7}G·G·C triplex

Table 1: ^{7}GGC Triplex Chemical Shifts[a]

residue	imino	aminos	H6,H8	H2,H5,Me	H1′	H2′	H2″	H3′	H4′	H5′,5″
A1	—[b]	7.83, 7.50								
G2	12.54		7.57	−	6.04	2.62	2.96	5.08	4.34	
A3	−	7.39, 7.12	7.16	7.30	5.85	2.15	2.65	4.78	4.65	
G4	12.58		7.12	−	5.90	2.28	2.82	4.97	4.28	
A5	−	7.56, 7.47	7.18	7.40	5.83	2.08	2.69	4.79	4.50	
G6	12.72		7.33	−	5.86	2.39	2.86	4.84	4.39	
A7	−	7.83, 7.61	7.38	7.42	5.88	2.40	2.76	4.89		
A8	−	7.73, 7.61	7.77	7.93	6.11	2.43	2.54	4.84	4.44	
C9	−		7.55	5.73	5.92	2.07	2.36	4.68		
C10	−		7.54	5.90	5.94	1.93	2.29	4.51		3.85
C11	−		7.45	5.58	5.73	1.88	2.42	4.68		
C12	−		7.86	6.06	6.22	2.37	2.62	4.70	4.36	
T13	14.73	−	7.65	1.80	6.09	2.26	2.71	4.84		
T14	14.12	−	7.54	1.66	6.10	2.33	2.65	4.92		
C15	−	8.42, 7.61	7.59	5.55	5.94	2.26	2.55	4.89	4.18	
T16	14.30	−	7.52	1.72	6.06	2.21	2.62	4.78		
C17	−	6.56, 8.94	7.60	5.55	6.10	2.27	2.68	4.77		
T18	14.44	−	7.39	1.67	5.84	2.12	2.51	4.81	4.08	
C19	−	8.21, 6.96	7.55	5.55	6.03	1.97	2.43	4.72		
T20	14.56	−	7.48	1.61	6.15	2.42	2.45	4.87	4.24	
T21		−	7.34	1.72	5.99	1.75	2.25	4.75	4.19	
A22	−		7.99		5.92	2.49	2.41	4.67	3.96	3.96
T23		−	7.00	1.27	5.59	1.60	1.93	4.47	3.40	3.59, 3.44
A24	−		8.01		5.59	2.68	2.68	4.86	4.32	3.91
T25	13.37	−	7.56	1.82	6.13			4.82		
C26	15.96	9.64, 9.17	7.86	5.85	6.10	2.00	2.64	4.75	4.31	
T27	13.45	−	7.48	1.71	5.82	2.37	2.54	4.97	4.34	
G28	11.57		8.24	−	6.24	2.50	2.58	4.90	4.38	
T29	13.07	−	7.42	0.99	6.18	2.32	2.60	4.82	4.35	
C30	14.58	10.01, 9.40	7.67	5.63	5.92	2.16	2.61	4.67	4.35	
T31	13.46	−	7.69	1.67	6.24	2.28	2.67	4.90	4.29	
T32	12.48	−	7.47	1.60	6.31	2.23	2.29	4.53	4.10	4.27

[a] Chemical shifts of the exchangeable and nonexchangeable resonances at 1 and 35 °C, respectively. [b] Not applicable.

in D_2O at 35 °C. Sequential aromatic−H1′ connectivities are observed from G2 through T32, excluding breaks in the loop regions between C9 and C10 and between T23 and A24. No cross peak was identified for A1. Sequential connectivities are also present in the aromatic-H3′ and aromatic−H2′,2″ regions (data not shown) and confirmed the assignments. Information from NOESY spectra taken at several temperatures combined with HOHAHA (Bax & Davis, 1985; Braunschweiler & Ernst, 1983), HOENOE (Sklenář & Feigon, 1990b), and P.COSY spectra (Marion & Bax, 1988; Mueller, 1987) was used to obtain assignments of the nonexchangeable resonances, as previously described (Macaya *et al.*, 1992a). These assignments were aided by a natural abundance $^{13}C-^{1}H$ HMQC spectrum (Bax *et al.*, 1983) which is especially useful for unambiguous identification of AH2 resonances. Resonance assignments for all of the base and sugar protons are listed in Table 1. Based on the relative integrated intensities of the intraresidue H1′−aromatic cross peaks (Figure 4), all residues including ^{7}G28 are in the *anti* conformation.

In addition to the standard intranucleotide and sequential connectivities, interstrand cross peaks indicative of triplex formation are present, e.g., cross peaks between the H1′ resonances of third strand sugars and the purine H8 resonances of the neighboring triplet in the 5′ direction (with respect to the purine strand). Analogous cross peaks are seen between H1′ protons of the third strand and purine H3′ protons of the adjacent triplet (data not shown). These are particularly important in the structure determination since they represent interstrand restraints between the Hoogsteen-bonded strands. Interstrand contacts between the Watson−Crick duplex strands are indicated by NOESY cross peaks between AH2 protons of the purine strand and H1′ protons

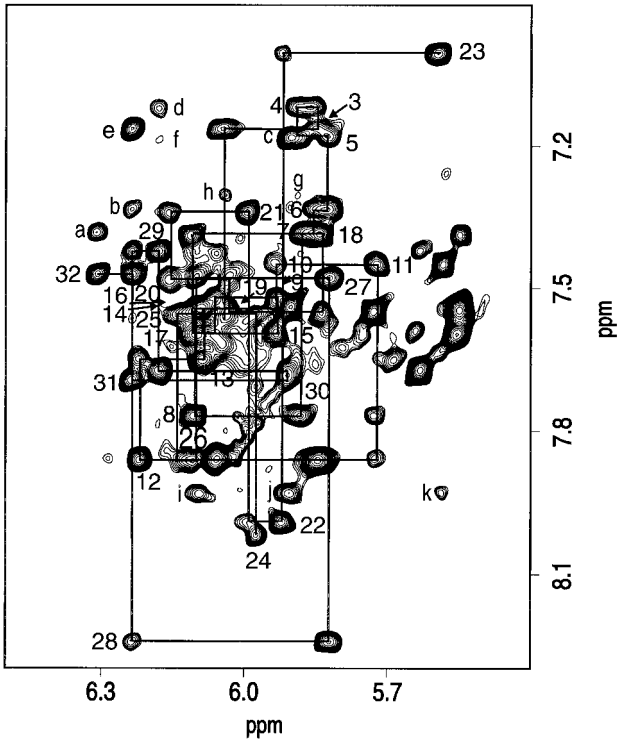

FIGURE 4: H1′, H5−aromatic region of a NOESY spectrum of the ^{7}GGC triplex at 35 °C and τ_m= 260 ms. Sequential connectivities are indicated by solid lines. Labeled interstrand NOEs are (a) T32H1′−A7H8, (b) T31H1′−G6H8, (c) C^+30H1′−A5H8 (overlapped with G4H1′−A5H8), (d) T29H1′−G4H8, (e) ^{7}G28H1′−A3H8, (f) T29H1′−A5H8, (g) G4H1′−A3H2, (h) C19H1′−A3H2, (i) T14H1′−A8H2, (j) C10H5−A8H2, and (k) C11H−A8H2.

of the Watson−Crick pyrimidine residue of the neighboring triplet (also in the 5′ direction with respect to the purine

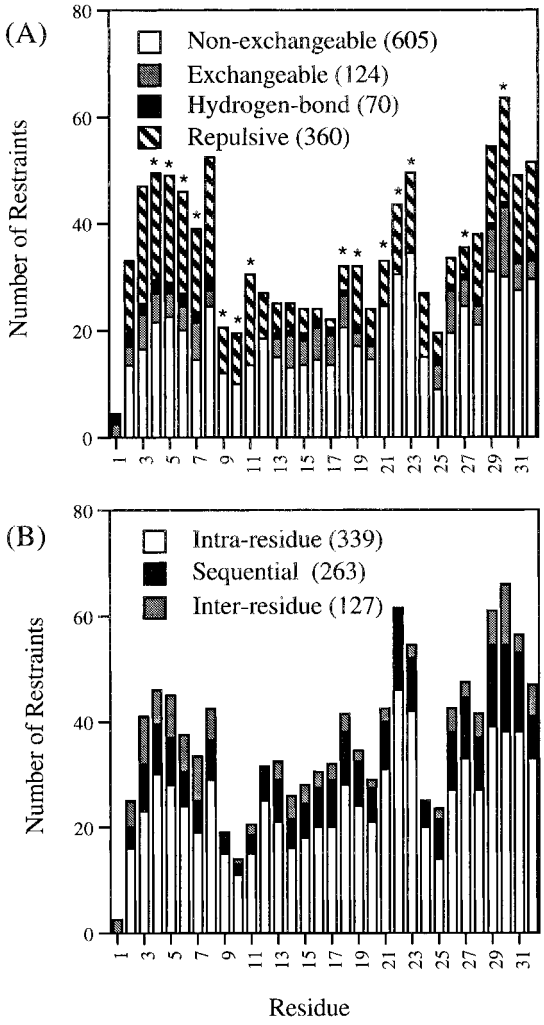

FIGURE 5: Number of NOE distance restraints per residue. (A) Repulsive, hydrogen bond, exchangeable, and nonexchangeable. Residues with ν_1 and ν_2 torsion angle restraints are labeled with an *. (B) Sequential, intraresidue, and interresidue.

strand). In the case of A8H2, NOESY cross peaks are also observed to CH5 protons of the cytosine loop.

NOE and Torsion Angle Restraints. A total of 1159 distance restraints were used in the initial stages of the structural calculations, prior to relaxation matrix refinement. These include 70 hydrogen bond restraints, 605 restraints from the nonexchangeable NOESY spectra, 124 restraints from the exchangeable NOESY spectrum, and 360 repulsive restraints derived from the absence of cross peaks for some short $^1H-^1H$ distances in a preliminary round of structure calculations. Figure 5A illustrates how these distance restraints are distributed over the individual residues. In Figure 5B the distance restraints for each residue are categorized as being either intraresidue, sequential, or interresidue (nonsequential).

For those parts of the triplex that were already well-defined (i.e., the triplet core and the CCCC loop), inclusion of repulsive NOE restraints did not result in a change of conformation nor notably improve the precision. The less well-defined TATA loop, however, is considerably better restrained and shows less conformational variability after inclusion of repulsive NOE restraints in the structure calculations. This was determined by refining the structures (prior to the complete relaxation matrix refinement step) both with and without the repulsive NOE restraints. The average pairwise RMSD values for both sets of structures are compared in Table 2.

Table 2: Effect of Repulsive NOE Restraints on Average RMSD

residues considered	without repulsive restraints	with repulsive restraints
all residues	2.00	1.55
triplex core	1.03	1.01
triplex core + CCCC loop	1.20	1.10
triplex core + TATA loop	1.97	1.52

Table 3: Coupling Constants of the Sugar Protons of the ^{7}GGC Triplex

	$J_{1'2'}$	$J_{1'2''}$	$J_{2'2''}$	$J_{2'3'}$	$J_{2''3'}$
G4	8.7	3.6	−12.0	6.4	4.1
A5	9.0	3.8	−12.7	5.8	3.8
G6	9.4	4.2	−13.3	6.3	3.8
A7	7.7	6.0	−14.8	6.5	1.9
C9	6.7	6.8	−14.9	6.2	3.0
C10	6.9	6.1	−13.1	5.8	3.0
C11	7.6	5.8	−15.3	4.6	0.9
T18	6.4	7.2	−16.4	8.0	2.4
C19	4.5	7.9	−16.4	6.8	3.7
T21	8.0	5.7	−14.2	6.1	3.4
A22	6.9	5.5	−14.4	5.4	4.1
T23	8.2	5.6	−14.5	5.9	2.8
T27	9.3	4.1	−12.8	5.9	2.9
C30	5.2	6.5	−15.9	6.8	4.0

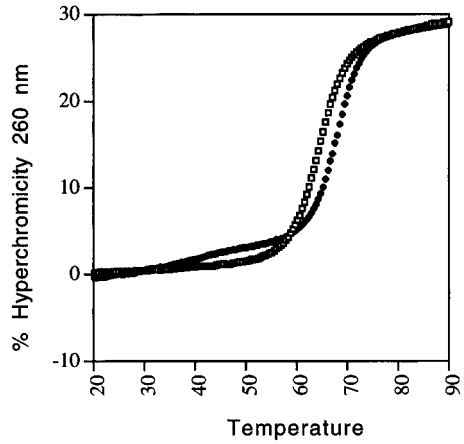

FIGURE 6: UV melting curves of the ^{7}GGC (open squares) and the C^+GC (filled diamonds) triplexes. The T_m's determined from the first derivatives are 65 and 69 °C, respectively.

In the ^2H_2O P.COSY spectrum (not shown), the H1′−H2′ and H1′−H2″ cross peaks of 14 residues were sufficiently resolved to be simulated using CHEOPS. The coupling constants thus obtained were used to determine sugar conformations, i.e., ν_1 and ν_2 torsion angle restraints for the sugar puckers and consequently the δ backbone angles (Table 3).

UV Spectroscopy. The melting behavior of the ^{7}GGC triplex and of the C^+GC triplex, in which the ^{7}G·G·C triplet is replaced with a C^+·G·C triplet, was studied by UV spectrophotometry. The plots of hyperchromicity vs temperature are shown in Figure 6. Under the conditions of ionic strength and pH 5.2 used in this experiment, it can be seen that both undergo a single cooperative transition but that the melting temperature of the ^{7}GGC triplex is 4 °C lower than that of the C^+GC triplex.

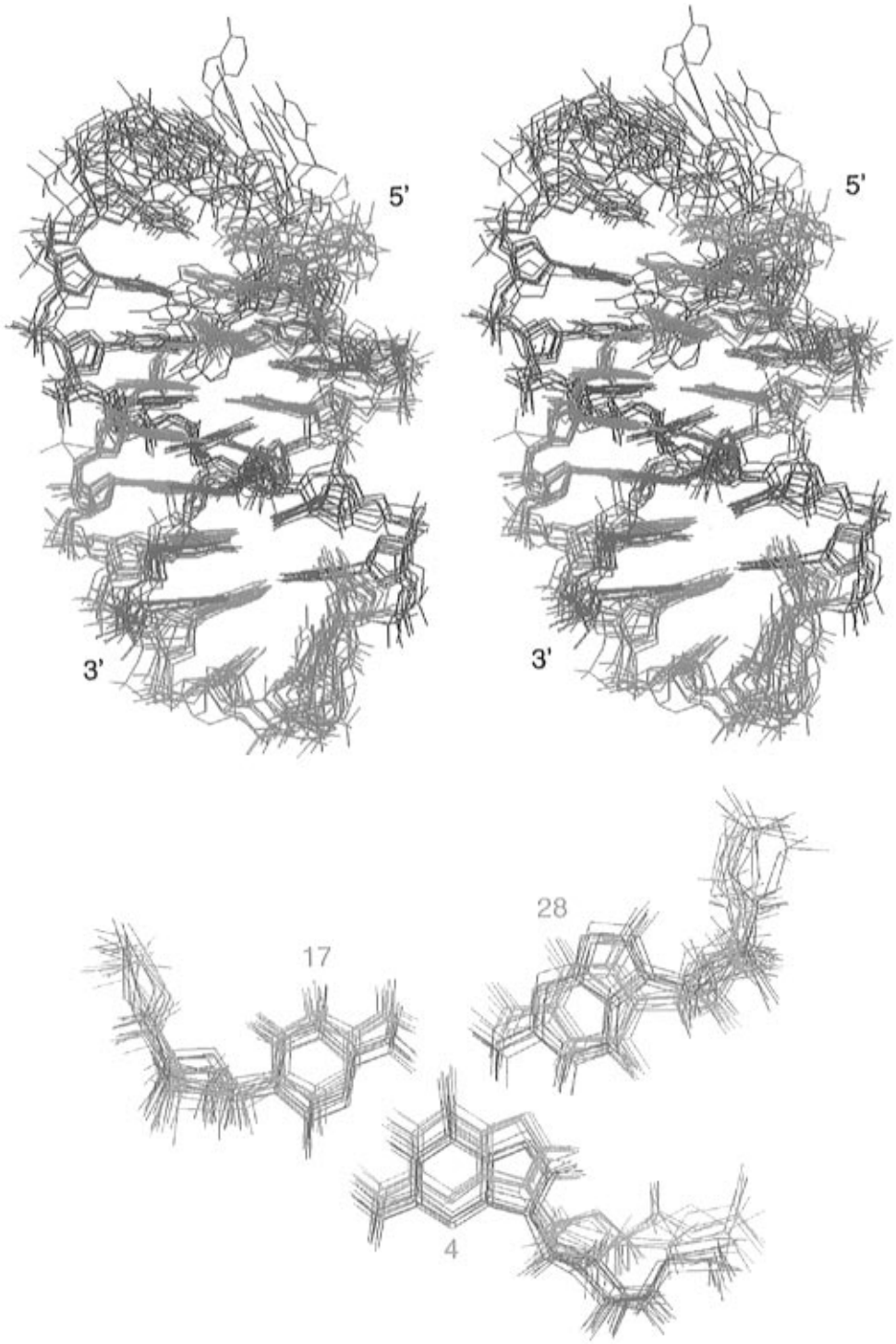

FIGURE 7: (A, top) Superposition of the eight lowest energy structures of the ^{7}GGC triplex, looking into the major groove. The central purine strand and the CCCC loop are in green, the Watson−Crick pyrimidine strand and the TATA loop are in red, the third strand is in blue, and ^{7}G is in light blue. (B, bottom) ^{7}G·G·C triplet (blue) superimposed on a C^+·G·C triplet (purple). The triplets were positioned relative to one another by superimposing their respective flanking triplets; i.e., the T27·A3·T18 and the T29·A5·T16 triplets were superimposed on the T29·A5·T16 and the T31·A7·T14 triplets, respectively.

DISCUSSION

Triplex Formation and Stability. In this work, we have studied the effect of incorporation of an N^7G on base triplet formation, structure, and stability of an intramolecular triplex. Triplex formation was initially assayed both from analysis of the UV melting profile and from analysis of the imino and amino proton resonances observed in NMR spectra of the sample in H_2O. The NMR spectra indicate that the oligonucleotide folds as designed into an intramolecular triplex with ^{7}G·G·C as a central triplet. Imino resonances are observed for each of the base pairs in both the Watson−Crick and Hoogsteen paired strands, and two sets of sequential imino−imino connectivities can be traced.

At pH 7, ^{7}G binds a target Watson−Crick G·C base pair with an equilibrium association constant greater than or equal to 1.6×10^8 M^{-1} (Hunziker *et al.*, 1995). This is an affinity equal to that shown by a $^{me}C^+$ for a G·C base pair in a similar context. In contrast, our UV melting study at pH 5.2 shows that the ^{7}GGC triplex melts 4 °C lower than the C^+GC triplex, in which the ^{7}G·G·C triplet is replaced by a C^+·G·C

Table 4: Refinement Statistics for the ^{7}GGC Triplex

NOE violations	3−7 > 0.5 Å
dihedral angle violations	none > 5
$R_{1/6}$ factor	11.3 (±0.2)%
average pairwise RMSD (all residues)	1.5 (±0.3) Å
average pairwise RMSD (triplex core)	1.1 (±0.2) Å
average RMSD from ideal covalent geometry	
bond lengths	0.038 Å
bond angles	7.4°

triplet. The destabilizing effect of a ^{7}G·G·C triplet relative to a C$^+$·G·C triplet at pH 5.2 can be rationalized in terms of its localized effect on base stacking interactions and the phosphodiester backbone and is discussed below. At pH 7, destabilization as a result of the difficulty of protonating the C$^+$·G·C triplet (Hunziker *et al.*, 1995) apparently outweighs any destabilization of the triplex caused by the inclusion of ^{7}G.

Triplex Structure and Helix Morphology. A superposition of the eight best (lowest energy) structures of the ^{7}GGC triplex is presented in Figure 7A. The single-stranded oligodeoxynucleotide containing ^{7}G folds into a well-defined, right-handed triple helical structure. Structure statistics for the refinement are summarized in Table 4. The average pairwise RMSDs are 1.5 ± 0.3 Å for the whole structure and 1.1 ± 0.2 Å for the triplex core, i.e., the stack of triplets without the loops. The triplex forms a regular helical structure with relatively little distortion at or around the site of the ^{7}G·G·C triplet. In order to analyze the structure and obtain a reasonable comparison to A-DNA and B-DNA helices, helical parameters for the Watson−Crick duplex part of the triplex were calculated (Figure 9). This analysis essentially treats the Hoogsteen paired strand as a ligand in the major groove. One of the characteristic parameters for the distinction between A- and B-DNA, the X-displacement of the Watson−Crick base pairs, varies from about −2.3 to −3.5 Å with an average of −2.9 Å. The extremes of the X-displacement occur near the ends of the triplex, where the regular helical structure may be affected somewhat by constraints from the loops. In the central part of the triplex, the range of helical displacement is much smaller, between about −2.8 and −3.2 Å, and alternates around the average from A·T to G·C base pairs. The X-displacement is about halfway between that for standard A-DNA and B-DNA duplexes. The rise per residue ranges from 2.4 to 3.5 Å, with an average of about 3.1 Å. The rise is probably the least well-defined of the helical parameters, since it is probably the most affected by the limited precision of the distance restraints. The helical twist varies from 18° to 41°, with an average of 29°. There is a clear alternation in twist values along the helix, with the ApG and ApA steps having a relatively low twist and the GpA steps having a relatively high twist. The alternation in twist values is inversely correlated with an alternation of rise values. This pattern of rise and twist values has been found in crystal structures of B-DNA decamers or dodecamers; i.e., ApG and ApA steps have a relatively low twist and a relatively high rise, while GpA steps have a relatively high twist and a relatively low rise (Quintana *et al.*, 1992; Yanagi *et al.*, 1991). The variation in twist values has been attributed to maximizing of stacking of the DNA bases (Calladine, 1982).

^{7}G·G·C Triplet Structure and Effect of Its Inclusion on the Helix. As expected from analysis of the NOESY cross peaks, the ^{7}G28 forms a base triplet with the Watson−Crick G4·C17 pair using the hydrogen bonding scheme proposed by Hunziker *et al.* (1995) (Figure 1). All three bases have standard *anti* glycosidic conformations. Although inclusion of a standard G to form a G·G·C triplet would require much greater distortions of the backbone than the ^{7}G, the ^{7}G·G·C triplet is still not completely isomorphous with a C$^+$·G·C triplet. In order to assess how the helix accommodates ^{7}G28·G4·C17, we superimposed the ^{7}G·G·C triplet on a C$^+$·G·C triplet as it would be if it were positioned at the same location in the triplex. This was done by superimposing the respective flanking triplets of ^{7}G28·G4·C17 and C$^+$-30·G6·C15 for the family of eight best structures. If the ^{7}G·G·C triplet were accommodated in the triplex with the same positioning of the Watson−Crick G·C base pairs as for a C$^+$·G·C triplet, this would result in a reasonably large distortion of the phosphodiester backbone in the third strand around the ^{7}G. Instead, it can be seen that in the triplex the ^{7}G base is approximately centered on the position where a C$^+$ base would be, with the result that the positions of the C1′ atoms are nearly isomorphous. The Watson−Crick G·C base pair is displaced from the position that it would occupy if it were part of a C$^+$·G·C triplet. The greatest displacement occurs in the position of the Watson−Crick G4 nucleotide, which is shifted to a position further out from the helix axis. Thus, the helix accommodates the new triplet by a slight displacement of all of the residues over their positions in the C$^+$GC triplex containing a C$^+$·G·C triplet at that site, with the largest change in the position of the phosphodiester backbone at G4. The resultant decrease in stacking interactions, along with a slight distortion of the backbone [perhaps coupled with electrostatic effects resulting from the loss of the positive charge on the protonated cytosine (Giovannangeli *et al.*, 1992)], accounts for the observed reduction in melting temperature discussed above.

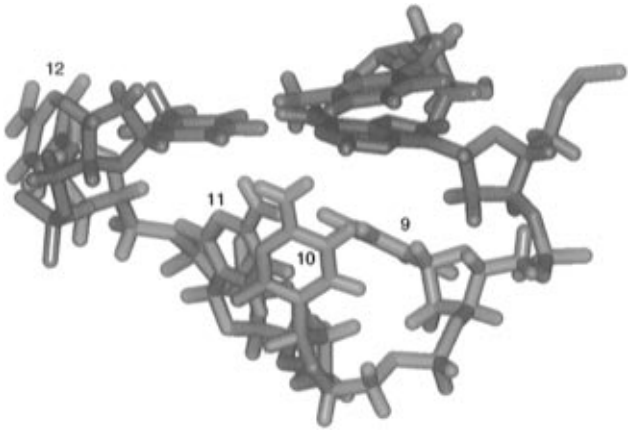

FIGURE 8: Lowest energy structure of the CCCC loop and its adjacent triplet. T32 is in blue, A8 and the CCCC loop are in green, and T13 is in red.

Loop Conformations. In contrast to other NMR studies of intramolecular pyrimidine motif DNA triplexes (Macaya *et al.*, 1992b; Radhakrishnan & Patel, 1994a,b; Wang *et al.*, 1996), we were able to assign most of the resonances in both loops. The CCCC loop which connects the Watson−Crick paired purine and pyrimidine strands of the ^{7}GGC triplex is nearly as well-defined as the triplex core (Figure 7A). The first residue of the loop, C9, stacks on A8 of the adjacent triplet. Following C9, the backbone makes a sharp turn, and C10 and C11 are positioned in the interior of the loop, approximately perpendicular to the planes formed by the triplets. C9 is positioned closest to C11, and C11 is stacked loosely on C10. C12 is turned out into the solution, and

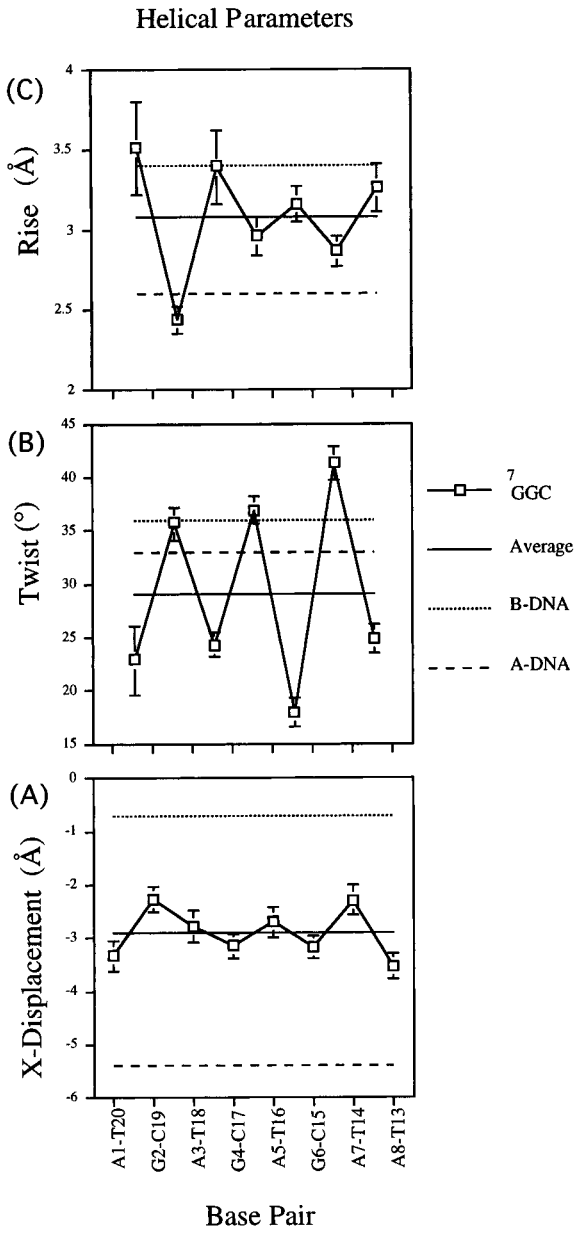

FIGURE 9: Helical parameters of the ^{7}GGC triplex and comparison to standard A-DNA and B-DNA: (A) X-displacement, (B) twist, and (C) rise. A-DNA and B-DNA refer to the standard DNA duplex conformations (Saenger, 1984). Error bars indicate the standard deviation of the measurement from the eight lowest energy structures.

there is another sharp turn of the backbone between C12 and the first base of the helix T13 (Figure 8). Although the cytosine bases in the loop could be protonated at the pH of the NMR sample, we find no evidence for formation of a C^+·C base pair in the loop.

In contrast to the CCCC loop, the TATA loop connecting the two pyrimidine strands is less well-defined. This may be attributable to the fact that there are fewer restraints on this loop due to ambiguous assignments. Whereas the triplex core and the CCCC loop are almost not affected, the conformational variability of the TATA loop is clearly diminished (cf. Table 2) by the inclusion of repulsive distance restraints based on absent NOE peaks. This method excludes conformations that would give rise to unobserved cross peaks and thus yields an ensemble of conformers giving a more accurate representation of the dynamics of this triplex molecule in solution. In all of the triplexes we have studied which contain a TATA loop, we have been unable to identify cross peaks arising from the first base in the helix (A1) (Macaya *et al.*, 1991; Wang *et al.*, 1992).

Comparison to Other Triplexes. Although for many years it was thought that triplexes adopted an A-DNA helical structure (Arnott & Selsing, 1974), more recent spectroscopic and structural studies of DNA triplexes have shown that in general they are more like B-DNA (Liquier *et al.*, 1991; Liu *et al.*, 1994; Macaya *et al.*, 1992a). There are no crystal structures of DNA triplexes, but NMR-derived models and structures of several pyrimidine motif intramolecular DNA triplexes have been reported (Bornet & Lancelot, 1995; Macaya *et al.*, 1992b; Radhakrishnan & Patel, 1994a,b; Wang *et al.*, 1996). In these structures, most of which contain a mismatch or nonnatural base in the third strand, the average X-displacement varied from −1.4 Å (Wang *et al.*, 1996) to −2.1 Å (Radhakrishnan & Patel, 1994a). The ^{7}GGC triplex presented here has a somewhat larger average X-displacement of −2.9 Å (Figure 9A). All of these values are intermediate to those of B-form DNA (−0.7 Å) and A-form DNA (−5.4 Å). The average rise for the reported triplex structures is 3.1−3.4 Å. Within the accuracy of the structures these values are essentially the same. These values are close to or identical to the canonical B-DNA rise of 3.4 Å and are larger than the A-DNA rise of 2.6 Å. Of these triplex structures, the ^{7}GGC triplex has the lowest average twist (29°) (Figure 9C), while the triplex containing a single noncanonical TCG triplet has the highest (32°). The variation in twist from step to step is within the ranges found in both B-DNA and A-DNA crystal structures. Finally, all of the NMR-derived triplex structures have sugar puckers that are predominantly S-type. While all of these triplex structures were refined using somewhat different methods, and the refinement protocols have improved substantially in the more recent structures, there are trends in the helical parameters. All of these pyrimidine−purine−pyrimidine triplexes have an axial rise similar to B-DNA, an X-displacement intermediate between A- and B-DNA, and sugars that are mostly S-type in the range generally found in B-DNA. Thus, binding of a third strand in the major groove of B-DNA induces a structural rearrangement that moves the Watson−Crick base pairs slightly off the helix axis but otherwise retains most of the features of B-DNA.

REFERENCES

Arnott, S., & Selsing, E. (1974) *J. Mol. Biol. 88*, 509−521.
Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−360.
Bax, A., Griffey, R. H., & Hawkins, B. L. (1983) *J. Magn. Reson. 55*, 301−315.
Beal, P. A., & Dervan, P. B. (1991) *Science 251*, 1360−1363.
Beaucage, S. L., & Iyer, R. P. (1993) *Tetrahedron 49*, 6123−6194.
Belotserkovskii, B. P., Veselkov, A. G., Filippov, S. A., Dobrynin, V. N., Mirkin, S. M., & Frank-Kamenetskii, M. D. (1990) *Nucleic Acids Res. 18*, 6621−6624.
Best, G. C., & Dervan, P. B. (1995) *J. Am. Chem. Soc. 117*, 1187−1193.
Bornet, O., & Lancelot, G. (1995) *J. Biomol. Struct. Dyn. 12*, 803−814.
Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.
Broitman, S. L., Im, D. D., & Fresco, J. R. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5120−5124.
Brünger, A. T. (1992) *X-PLOR (Version 3.1) Manual*, Yale University Press, New Haven and London.
Calladine, C. R. (1982) *J. Mol. Biol. 161*, 343−352.
Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J., & Hogan, M. E. (1988) *Science 241*, 456−459.

Davison, E. C., & Johnsson, K. (1993) *Nucleosides Nucleotides 12*, 237−243.
Dayn, A., Samadashwily, G. M., & Mirkin, S. M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 11406−11410.
de Leeuw, F. A. A. M., & Altona, C. (1983) *J. Comput. Chem. 4*, 428−437.
de los Santos, C., Rosen, M., & Patel, D. (1989) *Biochemistry 28*, 7282−7289.
Feigon, J., Koshlap, K. M., & Smith, F. W. (1995) *Methods Enzymol. 261*, 225−255.
François, J.-C., Saison-Behmoaras, T., Barbier, C., Chassignol, M., Thuong, N. T., & Hélène, C. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9702−9706.
Froehler, B. C., & Ricca, D. J. (1992) *J. Am. Chem. Soc. 114*, 8320−8322.
Froehler, B. C., Wadwani, S., Terhorst, T. J., & Gerrard, S. R. (1992) *Tetrahedron Lett. 33*, 5307−5310.
Giovannangeli, C., Rougee, M., Garestier, T., Thuong, N. T., & Hélène, C. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 8631−8635.
Greenberg, W. A., & Dervan, P. B. (1995) *J. Am. Chem. Soc. 117*, 5016−5022.
Griffin, L. C., & Dervan, P. B. (1989) *Science 245*, 967−971.
Hampel, K. J., Burkholder, G. D., & Lee, J. S. (1993) *Biochemistry 32*, 1072−1077.
Hélène, C., & Toulme, J. J. (1990) *Biochim. Biophys. Acta 1049*, 99−125.
Huang, C. Y., & Miller, P. S. (1993) *J. Am. Chem. Soc. 115*, 10456−10457.
Hunziker, J., Priestley, E. S., Brunar, H., & Dervan, P. B. (1995) *J. Am. Chem. Soc. 117*, 2661−2662.
Koh, J. S., & Dervan, P. B. (1992) *J. Am. Chem. Soc. 114*, 1470−1478.
Kohwi, Y., & Kohwi-Shigematsu, T. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 3781−3785.
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.
Kumar, A., Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1−6.
Lavery, R., & Sklenar, H. (1988) *J. Biomol. Struct. Dyn. 6*, 63−91.
Lee, J. S., Woodsworth, M. L., Latimer, L. J., & Morgan, A. R. (1984) *Nucleic Acids Res. 12*, 6603−6614.
Lipsett, M. N. (1963) *Biochem. Biophys. Res. Commun. 11*, 224−228.
Lipsett, M. N. (1964) *J. Biol. Chem. 239*, 1256−1260.
Liquier, J., Coffinier, P., Firon, M., & Taillandier, E. (1991) *J. Biomol. Struct. Dyn. 9*, 437−445.
Liu, K., Miles, H. T., Parris, K. D., & Sasisekharan, V. (1994) *Nat. Struct. Biol. 1*, 11−12.
Lu, G. H., & Ferl, R. J. (1993) *Int. J. Biochem. 25*, 1529−1537.
Macaya, R. F., Gilbert, D. E., Malek, S., Sinsheimer, J., & Feigon, J. (1991) *Science 254*, 270−274.
Macaya, R. F., Schultze, P., & Feigon, J. (1992a) *J. Am. Chem. Soc. 114*, 781−783.
Macaya, R. F., Wang, E., Schultze, P., Sklenář, V., & Feigon, J. (1992b) *J. Mol. Biol. 225*, 755−773.
Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95−117.
Maher, L. J., III, Wold, B., & Dervan, P. B. (1989) *Science 245*, 725−730.
Marion, D., & Bax, A. (1988) *J. Magn. Reson. 80*, 528−533.
Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989) *J. Magn. Reson. 85*, 393−399.
Mergny, J.-L., Sun, J.-S., Rougée, M., Montenay-Garestier, T., Barcelo, F., Chomilier, J., & Hélène, C. (1991) *Biochemistry 30*, 9791−9798.
Michel, D., Chatelain, G., Herault, Y., & Brun, G. (1992) *Nucleic Acids Res. 20*, 439−443.
Miller, P. S., Bhan, P., Cushman, C. D., & Trapane, T. L. (1992) *Biochemistry 31*, 6788−6793.
Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A., & Frank-Kamenetskii, M. D. (1987) *Nature 330*, 495−497.
Moser, H. E., & Dervan, P. B. (1987) *Science 238*, 645−650.
Mueller, L. (1987) *J. Magn. Reson. 72*, 191−196.
Nagel, K. M., Holstad, S. G., & Isenberg, K. E. (1993) *Pharmacotherapy 13*, 177−188.
Neidig, K. P., Geyer, M., Gorler, A., Antz, C., Saffrich, R., Beneicke, W., & Kalbitzer, H. R. (1995) *J. Biomol. NMR 6*, 255−270.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317−324.
Ono, A., Ts'o, P. O. P., & Kan, L.-s. K. (1991) *J. Am. Chem. Soc. 113*, 4032−4333.
Povsic, T. J., & Dervan, P. B. (1989) *J. Am. Chem. Soc. 111*, 3059−3061.
Prins, J., de Vries, E. G., & Mulder, N. H. (1993) *Clin. Oncol. (R. Coll. Radiol.) 5*, 245−252.
Quintana, J. R., Grzeskowiak, K., Yanagi, K., & Dickerson, R. E. (1992) *J. Mol. Biol. 225*, 379−395.
Radhakrishnan, I., & Patel, D. J. (1994a) *J. Mol. Biol. 241*, 600−619.
Radhakrishnan, I., & Patel, D. J. (1994b) *Structure 2*, 17−32.
Radhakrishnan, I., de los Santos, C., & Patel, D. J. (1991) *J. Mol. Biol. 221*, 1403−1418.
Rajagopal, P., & Feigon, J. (1989a) *Biochemistry 28*, 7859−7870.
Rajagopal, P., & Feigon, J. (1989b) *Nature 339*, 637−640.
Roberts, R. W., & Crothers, D. M. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9397−9401.
Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
Singleton, S. F., & Dervan, P. B. (1992) *Biochemistry 31*, 10995−11003.
Sklenář, V., & Bax, A. (1987) *J. Magn. Reson. 75*, 378−383.
Sklenář, V., & Feigon, J. (1990a) *Nature 345*, 836−838.
Sklenář, V., & Feigon, J. (1990b) *J. Am. Chem. Soc. 112*, 5644−5645.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286−292.
Strobel, S. A., & Dervan, P. B. (1991) *Nature 350*, 172−174.
Strobel, S. A., Moser, H. E., & Dervan, P. B. (1988) *J. Am. Chem. Soc. 110*, 7927−7929.
Strobel, S. A., Doucette-Stamm, L. A., Riba, L., Housman, D. E., & Dervan, P. B. (1991) *Science 254*, 1639−1642.
Sun, J. S., & Hélène, C. (1993) *Curr. Opin. Struct. Biol. 3*, 345−356.
Sun, J. S., De Bizemont, T., Duval-Valentin, G., Montenay-Garestier, T., & Hélène, C. (1991) *C. R. Acad. Sci., Ser. III 313*, 585−590.
Thuong, N. T., & Hélène, C. (1993) *Angew. Chem., Int. Ed. Engl. 32*, 666−690.
Wang, E., Malek, S., & Feigon, J. (1992) *Biochemistry 31*, 4838−4846.
Wang, E., Koshlap, K. M., Gillespie, P., Dervan, P. B., & Feigon, J. (1996) *J. Mol. Biol. 257*, 1052−1069.
Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M., & Wohlrab, F. (1988) *FASEB J. 2*, 2939−2949.
Xiang, G. B., Soussou, W., & McLaughlin, L. W. (1994) *J. Am. Chem. Soc. 116*, 11155−11156.
Xodo, L. E., Manzini, G., Quadrifoglio, F., Van der Marel, G. A., & Van Boom, J. H. (1991) *Nucleic Acids Res. 19*, 5625−5631.
Yanagi, K., Privé, G. G., & Dickerson, R. E. (1991) *J. Mol. Biol. 217*, 201−214.
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.
Young, S. L., Krawczyk, S. H., Matteucci, M. D., & Toole, J. J. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 10023−10026.

BI962438A