# Solution Conformation of an Intramolecular DNA Triplex Containing a Nonnucleotide Linker: Comparison with the DNA Duplex†

John P. Bartley,‡ Tom Brown,§ and Andrew N. Lane*,‖

*Centre for Instrumental and Developmental Chemistry, Department of Chemistry, Queensland University of Technology, Brisbane, Australia, Department of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, U.K., and Division of Molecular Structure, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, U.K.*

*Received March 26, 1997; Revised Manuscript Received September 17, 1997*⊗

ABSTRACT: The solution properties of the parallel intramolecular DNA triplex d(GAGAGA-oct-TCTCTC-oct-CTCTCT) (oct = -O-$(CH_2)_8$-O-$PO_2$-O-$(CH_2)_8$-O-$PO_2$-) and the duplex d(GAGAGA-oct-TCTCTC) have been examined by UV melting and high-resolution nuclear magnetic resonance spectroscopy (NMR). All nucleotides were primarily in the *S* conformation (i.e. near C2′-*endo*) in both the duplex and the triplex. However, the sugars of the Hoogsteen pyrimidine strand had a lower fraction of the *S* state than the Watson−Crick strands. Glycosidic torsion angles derived from nuclear Overhauser effect (NOE) build-up curves were found in the range −103° to −133°, with a clear alternation in magnitude along the GAGAGA strand in the triplex, whereas the glycosidic torsion angles were more similar in the duplex. Internucleotide NOEs were also consistent with an overall B-like geometry, rather than the A family. However, particularly in the Hoogsteen strand, some sequential NOE intensities were intermediate between those of the B and A forms. Distance and torsion constraints derived from NMR experiments were used to generate structures and were refined by restrained molecular dynamics. Extensive chemical shift differences between residues in the triplex and duplex were found for the purine strand, and there were remarkable differences in the pattern of shift differences for the A and G residues that correlated with differences in glycosidic torsion angles. Although there are differences in structure between the free duplex and that in the triplex, they are in important respects similar, indicating that only small conformational adjustments are needed to make parallel triple helices.

DNA triplexes have been known for almost 40 years (*1*). Recently, interest in such structures has been rekindled as the possibility of using a third DNA strand to hybridize to B-DNA was recognized as a potential antigene agent. Because intermolecular triplexes are relatively unstable (*2*−*4*), much of the published structural work has been concerned with intramolecular triplexes consisting of a single DNA oligonucleotide that folds back on itself, forming an intrinsically stable double hairpin (*5*−*13*). The requirement for loops of 4−5 nucleotides to avoid strain has made analysis by NMR difficult, as such triplexes typically contain 30 or more residues. The relatively large number of nucleotides and the stringent sequence requirements for the formation of triplexes via Hoogsteen or reversed Hoogsteen pairing of the third strand have resulted in very crowded NMR spectra. Nevertheless, solution structures of such triplexes have been reported (*5*−*8*, *12*, *13*).

Triple helices can be formed from DNA strands, RNA strands, PNA[1] strands, and mixtures thereof. There are conflicting reports in the literature regarding the conformation of the Watson−Crick duplex within the triplexes, with some describing them as B-like (*7*, *10*) and others as A-like (*14*−*16*), though the structures solved showed some characteristics that were intermediate between the canonical A and B structures (*8*). Recently, a crystal structure of a DNA·$PNA_2$ triplex was reported (*17*), in which the Watson−Crick DNA·PNA moiety was clearly more like A-form than B-form DNA, though it was sufficiently different from either to be given a new structural designation, namely, the P-form. Furthermore, although they appear rather similar globally, there are significant local conformational differences between the antiparallel YR·R and parallel YR·Y DNA triplexes (*5*−*8*, *18*). In addition, it has been suggested that there may be two classes of triplex conformation, determined by the sugar composition of the strands (*19*, *20*). It therefore seems likely that there are at least three structural classes of triplex, determined by the chemical nature of the strands (DNA, RNA, or PNA) and the base composition of the third strand.

We have chosen to use an intrinsically flexible nonnucleotide linker between the strands, either oct or hexa(ethylene glycol) (*21*−*23*). Optimal stability seems be achieved with heptakis(ethylene glycol) (*23*), though the reported difference between 6 and 7 monomers was slight, and hexa(ethylene glycol) linkers have been successfully used in structural studies (*18*, *21*, *22*). These linkers have the dual advantages of not imposing steric stress on the DNA strands and having proton NMR resonances that do not overlap with those of the DNA. We have examined in detail the solution structure of d(GAGAGA-X-TCTCTC-X-CTCTCT) using both linkers. This is the first triplex of the YR·Y kind not containing a mismatch to be solved. At pH 5, where the cytosines on

† This work was supported by the Medical Research Council of the UK and Oswel Research Products Co.

* To whom correspondence should be addressed: FAX 0181 906 4477; e-mail alane@nimr.mrc.ac.uk.

‡ Queensland University of Technology.

§ University of Southampton.

‖ National Institute for Medical Research.

⊗ Abstract published in *Advance ACS Abstracts,* November 1, 1997.

[1] Abbreviations: PNA, peptide nucleic acid; heg, hexa(ethylene glycol); oct, -O-$(CH_2)_8$-O-$PO_2$-O-$(CH_2)_8$-O-$PO_2$-; DQF-COSY, double quantum filtered correlation spectroscopy; rmsd, root mean square difference; NOESY, nuclear Overhauser effect spectroscopy; ROESY, rotating frame Overhauser effect spectroscopy.

S0006-2960(97)00710-1 CCC: $14.00 

the third strand become protonated, Hoogsteen base-pairing with the DNA duplex is stabilized (*2*, *24*). We have undertaken a comparison of the conformational properties of the triplex and the duplex under similar conditions. As the duplex part of the molecule consists of an oligopurine·oligopyrimidine sequence, it is possible that the duplex structure is significantly different from that of mixed-sequence DNA. There are few examples of duplex structures of this kind. However, it is known that $d(A)_n \cdot d(T)_n$ has a significantly different structure from that of mixed-sequence DNA (*25*), and alternating GA/CT sequences have been the object of interest as they can form a variety of different structures (*26*, *27*). It is also expected that the conformation of the DNA duplex would be altered by binding a DNA strand in the major groove (*28*).

## MATERIALS AND METHODS

### *Materials*

Heg and oct linker phosphoramidites for oligonucleotide synthesis were synthesized as previously reported (*29*). Oligonucleotides were synthesized by standard solid-phase phosphoramidite methods and purified by HPLC (*30*). Both the triplex d(GAGAGA)-X-(TCTCTC)-X-(CTCTCT) and the duplex d(GAGAGA)-X-(TCTCTC) were independently synthesized and analyzed by NMR spectroscopy (X = heg or oct). The underline denotes the Watson−Crick pairs in the triplexes. The nucleotides are numbered sequentially from G1 to T18.

### *Methods*

*Thermodynamic Stability.* Thermodynamic stabilities of intramolecular triplexes were determined as a function of pH by UV melting as previously described (*31*). $T_m$ values were determined from the maximum of the derivative of the melting curve. At low pH there was a single transition, but at higher pH (above pH 6.5) two transitions were observed, with the lower-temperature transition shifting to lower temperatures as the pH was increased. At high pH, the triplex-to-duplex and duplex-to-strand transitions were well-resolved (see Results), so that the parameters for the two transitions could be determined independently. The duplex-to-strand transition was independent of pH, whereas the triplex was stabilized by protons.

*NMR Spectroscopy.* $^1H$ spectra were recorded at 9.4, 11.75, and 14.1 T on Bruker AM400, Varian UnityPlus, and Varian Unity spectrometers, respectively. Oligonucleotide samples were made up in 0.15 M KCl and 20 mM $MgSO_4$, and the pH was adjusted to 5 to a final triplex concentration of 2 mM. Spectra in $^1H_2O$ were recorded using the gradient echo Watergate read pulse (*32*). NOESY spectra were recorded in both $^1H_2O$ and $^2H_2O$. DQF-COSY spectra were recorded in $^2H_2O$. On the Bruker instrument, the time-proportional phase incrementation scheme was used to obtain pure absorption mode spectra, whereas on the Varian spectrometers, the hypercomplex method was used (*33*). Zero-quantum coherences in NOESY spectra were suppressed using a variable 180° pulse in the mixing time (*34*).

Sums of coupling constants were measured from cross-sections of DQF-COSY spectra. The H1′−H4′ distance was determined from the ratio of the areas of the H1′−H4′/H1′−H2″ cross-peaks through H1′ in NOESY spectra. The distance was calculated by extrapolating back to zero mixing time (*35*). An isotropic rotational correlation time of 2.4 ns was determined from the time dependence of the cytosine H6−H5 vectors (*36*, *37*). Cross-peak volumes were determined using the peak-fitting routine within Felix (Molecular Simulations Inc., San Diego, CA), whereby the relative volumes were estimated as the product of the peak areas along *F*2 and *F*1, divided by the mean peak height. These volumes were then normalized to the average cross-peak volume of the cytosine H6−H5 pairs at each mixing time.

*Molecular Structures.* Sugar puckers were analyzed using sums of coupling constants ($\sum_{1'}$, $\sum_{2'}$, and $\sum_{2''}$) derived from DQF-COSY, NOESY, and in some cases from 1D spectra. The scalar coupling information was supplemented where possible with the H1′−H4′ distance determined from NOESY build-up rates, and the data were analyzed with the program pfit as previously described (*35*). Standard deviations for the coupling data were typically 0.5 Hz for $\sum_{1'}$ and 1 Hz for $\sum_{2'}$ and $\sum_{2''}$. In some instances, an estimate of $\sum_{3'}$ was obtained, which was used only qualitatively to define possible ranges of sugar conformation. Sugar conformations were systematically searched with *P* values from 0° to 198°, pucker amplitudes from 30° to 42°, and fraction of the *S* state from 0.4 to 1 in steps of 5°, 3°, and 0.02, respectively. Glycosidic torsion angles were determined from the time-dependent intranucleotide NOEs using the program NUCFIT (*38*), which compares NOEs calculated for nucleotides embedded in a short duplex with the experimental NOEs. The target nucleotide is considered to exist as a mixture of conformational states, namely, $\chi$(S)*P*(S) and $\chi$ (N)*P*(N), with the mole fraction of the *S* state $f_s$. Different initial values of $\chi$ and *P* were chosen, and for each starting mixture of conformations, the difference between the calculated and experimental NOEs is used to drive conformational readjustments using the least-squares Marquardt−Levenberg algorithm (*39*). The calculation therefore accounts for spin diffusion and conformational averaging of the nucleotides, within the framework of a two-state equilibrium.

*N*3-protonated cytosine was created within Insight II (Molecular Simulations Inc., San Diego, CA) from cytosine by rehybridizing the *N*3 atom, adding a proton, and setting the formal charge to +1 (Figure 1). The fragment was minimized and stored in a fragment library under the name CH. The modified nucleotide was incorporated into the triplex by replacing an existing C with CH, where required.

Initial triplex structures were made starting with either a standard B-DNA or A-DNA duplex and docking the third strand (in either the A or B conformations) under the influence of the NOE-derived distances between this strand and the duplex. A more efficient approach was found to be docking the third strand to duplexes that had been refined starting from either A or B structures using the experimental constraints appropriate to the duplex part. Initial structures were first energy-minimized, followed by restrained molecular dynamics (rMD) at 1000 K including all of the experimental constraints for 10 ps. The system was cooled in steps to 300 K, with final rounds of 10 ps of rMD and 1000 steps of restrained conjugate gradient energy minimization.

The separate intramolecular DNA duplex was refined using a simple rMD protocol starting from either A or B DNA structures, which were partially randomized at 500 K for

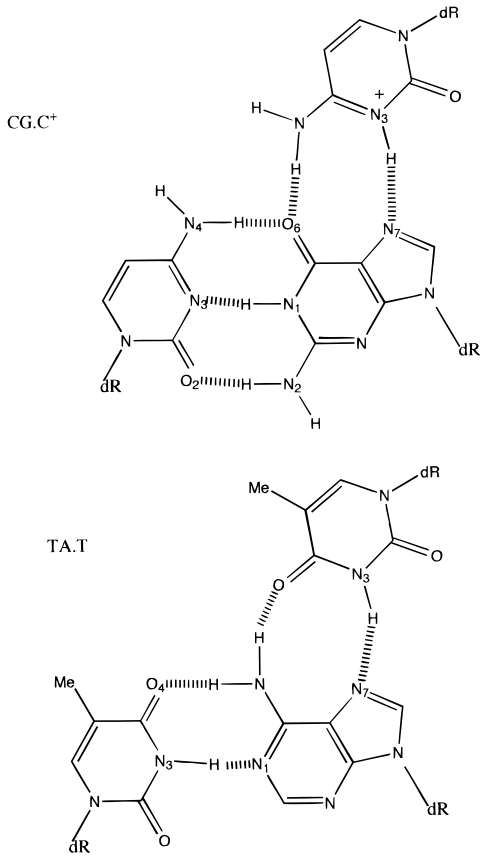


FIGURE 1: Structures of TA·T and CG·C$^+$ triplets.

1−2 ps in the absence of constraints, using the Amber force field with a distance-dependent dielectric ($\epsilon = 4r$), the electrostatics scaled down by a factor of 4, and the Lennard−Jones potential scaled down by a factor of 2. The experimental constraints were then applied and the system was subjected to 10−20 ps of rMD at 300 K, followed by 5000 steps of conjugate gradient minimization. Energies were calculated with the full force field (i.e., all scaling factors set to unity). Parameters were retained with full and low force field scaling for comparison. With the density and precision of the constraints, the effect of the force field was rather minor for the duplex. Torsion angles for the refined structures were obtained from DISCOVER, and additional parameters describing the helical properties were obtained using Curves 5.1 (*40*). Convergence was monitored by the low residual restraint energy (<1 kcal mol$^{-1}$), low potential energy, and good stereochemistry assessed by the potential energy of the bond length and bond angle deviations compared with simple energy-minimized structures.

## RESULTS

*Thermodynamic Stability*. UV melting curves were recorded for numerous triplexes as a function of cation concentration and pH for both the heg and oct linkers. At pH 5, only one transition was observed, but at pH values above 6, the melting resolved into two transitions. The lower-temperature transition was assigned to the melting of the triplex to the duplex plus strand state (*4*). The dependence of the $T_m$ for the triplex-to-duplex transition on pH is shown in Figure 2. At all pH values tested, the oct linker gave a slightly more stable triplex than the heg linker. The triplexes are stabilized by protonation, which we will show later is mainly at the N3 of C15. However, Figure 2 shows that the $T_m$ decreases essentially linearly with increasing pH, and there is no indication of an inflection point up to pH 7.5. Hence, the apparent p$K_a$ must be much greater than 7.5. Because the $T_m$ at pH 7.5 is only 10 °C, it is not possible to measure the melting at higher pH, and therefore a p$K_a$ value cannot be determined for this triplex. Furthermore, it is unlikely that the three Hoogsteen cytosines will have identical p$K_a$ values; the terminal residue will have a substantially lower p$K_a$ value than the internal cytosine residues (J. L. Asensio Alvarez, A. N. Lane, and T. Brown, unpublished data; and see below). However, to a first approximation, the apparent p$K_a$ for such a curve will be the arithmetic mean of the p$K_a$ of the cytosine in the strand state (ca. 4.5) and in the triplex state. As the apparent p$K_a$ > 7.5, this implies a p$K_a$ in the triplex state of >10. This is consistent with recent results on a different intramolecular triplex, where the p$K_a$ in the triplex state was shown to be >8.5 (*4*). Given a p$K_a$ value of ca. 4.5 for free cytosine (*41*), even an apparent p$K_a$ of >7.5 implies an increase of at least 3 units over the strand state, which corresponds to a difference in the free energy of dissociation of the protonated triplex and the unprotonated triplex of 7*RT* (4.2 kcal mol$^{-1}$ at 298 K).

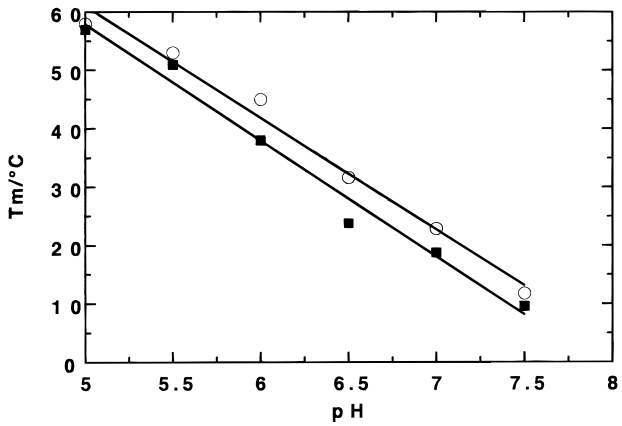


FIGURE 2: Dependence of $T_m$ on pH for the oct and heg linkers. Ultraviolet melting experiments were carried out as described under Methods. $T_m$ values were determined from the maxima in plots of d$A$(260)/d$T$ versus temperature. (■) heg linker; (□) oct linker. Lines are linear regression fits.

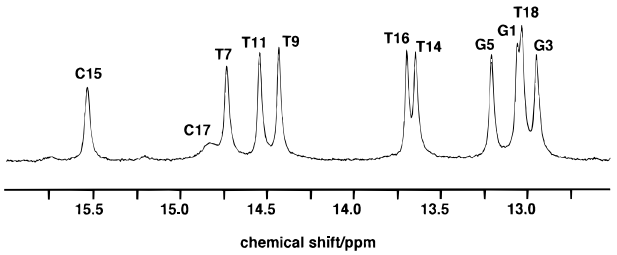


FIGURE 3: NMR spectrum of d(GAGAGA-oct-TCTCTC-oct-CTCTCT) in $H_2O$. The spectrum was recorded at 600 MHz, 2 °C and pH 4.5 as described under Methods. The low-field (imino proton) region is shown with the assignments (see text).

*NMR Spectroscopy of the Triplex d(GAGAGA-X-TCTCTC-X-CTCTCT)*. We have focused on the triplex with the oct linker, as it is slightly more stable than that with the heg linker (Figure 2) and also gave slightly better dispersed NMR spectra (not shown). Figure 3 shows the low-field NMR spectrum in $^1H_2O$. There are numerous resonances between

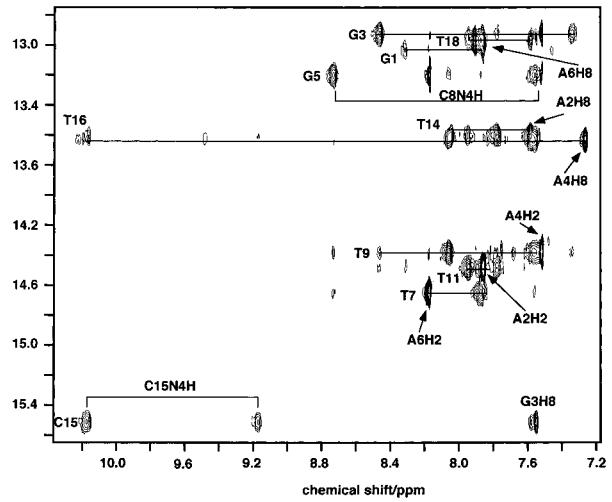


FIGURE 4: NOESY spectrum of d(GAGAGA-oct-TCTCTC-oct-CTCTCT) in $H_2O$. The spectrum was recorded at 10 °C with a mixing time of 300 ms. The acquisition times were 0.34 s in $t_2$ and 0.028 s in $t_1$. The data matrix was zero-filled to 16 384 by 2048 complex points and apodized with a Gaussian window function.

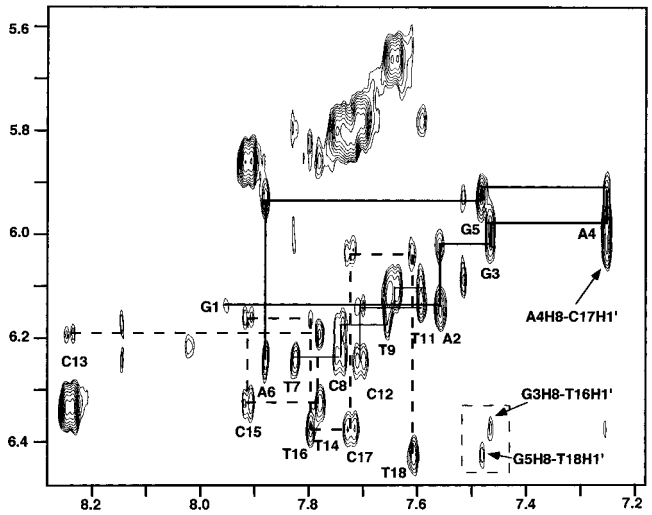


FIGURE 5: NOESY spectrum of d(GAGAGA-oct-TCTCTC-oct-CTCTCT). The spectrum was recorded at 600 MHz and 30 °C with a mixing time of 250 ms. The acquisition times were 0.37 s in $t_2$ and 0.046 s in $t_1$. The data matrix was zero-filled to 8192 by 2048 complex points and apodized with a Gaussian function.

15 and 9 ppm. As there is a maximum of 6 Watson−Crick base pairs, whose imino protons resonate between 12 and 14 ppm, there must be additional imino and amino proton resonances that can be attributed to the formation of Hoogsteen base pairs (Figure 1). Resonances near 1 and 3 ppm (not shown) correspond to linker protons, which do not overlap resonances of the nucleotides. The exchangeable protons were assigned using NOESY, as shown in Figure 4. The NH assignments are shown above the spectrum in Figure 3. These were obtained making use of the $NH_2$ of both the $CN4H_2$ and the $AN6H_2$, which in the TA·T triples are sharp owing to hydrogen-bonding interactions with both protons. NOEs between the two nonequivalent amino protons and to the imino protons and other amino protons allowed unambiguous assignments to be made for the Hoogsteen strand (Figures 3 and 4). We note that only C15 has an observable $N3H^+$, which showed NOEs to both of its N4 amino protons and also to the N3H of the two neighboring Thy residues, T14 and T16. There were no remaining unassigned resonances. Further, there were no additional peaks to the water resonance in the NOESY and ROESY spectra, indicating that the remaining $CN3H^+$ protons would have to be exchanging extremely rapidly. On lowering the pH to 4.5 and the temperature to 2 °C, the broad resonance at 14.85 ppm appeared (Figure 3) that integrated to less than one proton. We presume that this is from protonation of C17N3. This resonance shows the effects of exchange as it is much broader than the $N3H^+$ of C15, which implies a larger dissociation rate constant and presumably a lower p$K_a$ value. We found no evidence of any other broad peaks that might be attributed to the terminal $C13N3H^+$. This could either be due to a very fast exchange with solvent or to a p$K_a$ value close to that of free cytosine, i.e. 4.3−4.5 (*41*), or both. A low p$K_a$ for the terminal cytosine would be expected, as it is exposed to solvent. The presence of the Hoogsteen amino protons at their characteristic shift indicates that the terminal C does form a hydrogen-bonded triplet, however. We tentatively conclude that C13 is not fully protonated at pH 5. The AdeN6H were both sharp and exhibited NOEs to the two hydrogen-bonded TN3H. In duplex DNA, these amino protons are broad and difficult to observe owing to exchange by rotation about the C6N6 bond. In this instance, the hydrogen bonding of both amino protons to the thymidine carbonyl oxygen prevents exchange either by rotation or with solvent and is characteristic of the Hoogsteen TA·T base triplet. This is clearly visible in Figure 4. The figure also shows additional NOEs that cannot be assigned to exchangeable protons. These include NOEs between the Hoogsteen imino protons and nonexchangeable protons, which were assigned to the H8 of the purines (see below).

Nonexchangeable protons were assigned using NOESY, TOCSY, and DQF-COSY. The cytosine H6 and H5 resonances were first identified in the COSY spectrum, and the thymidine H6 and methyl groups from their characteristic NOE. Sequential connectivities for T C were identified by the NOEs TH6−CH5, and C T connectivities were identified from the CH6−TMe NOEs. These were then connected in the sequence using the H6($i$)−H1′($i$)−H6($i$ + 1)pathway (Figure 5). In this manner, the two pyrimidine strands could be unambiguously assigned. The remaining singlet peaks in the 7−8 ppm region belong to the H8 of the A and G residues and the H2 of the A residues. The latter were identified by their narrow lines, strong NOEs to the Watson−Crick-paired TNH (see below), NOEs to H1′($i$ + 1) on both Watson−Crick strands, and the absence of NOEs to the H2′/H2″ region. The purines were then sequentially assigned by the H8−H1′−H8 pathway. H2′ and H2″ resonances were assigned using the intraresidue NOEs H1′−H2″ and H8/H6−H2′ and further verified as H1′−H2′ and H1′−H2″ cross-peaks in the DQF-COSY spectrum. H3′ resonances were assigned using the H8/H6($i$)−H3′($i$)−H8/H6($i$ + 1) pathway and by the COSY H2′−H3′ cross-peaks. Proton assignments are available as Supporting Information.

NOEs were observed between the H8 of the purine residues and the H1′ of the $i$ + 1 residue in the third strand (Figure 5). Together with the NOEs between H8 and imino protons (Figure 4) and between the $NH_2$ protons of the two cytosine residues in each base triple, we conclude that the third strand forms Hoogsteen base pairs to the purine strand.

*Nucleotide Conformations.* A few H1′ resonances are resolved in 1D spectra. They showed either triplet or double-doublet splitting patterns, with $\sum_{1'} > 14$ Hz, which indicates

Table 1: Nucleotide Conformations[a]

| | triplex | | | | | duplex | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| base | $P_s$ (deg) | $f_s$ | % E | $-\chi$ (deg) | *R* | $P_s$ (deg) | $f_s$ | % E | $-\chi$ (deg) | *R* |
| G1 | 135 | 0.88 | 0.05 | 110 | 0.12 | 159 | 0.94 | 2.1 | 124 | 0.062 |
| A2 | 144 | 0.90 | 0.54 | 124 | 0.095 | 162 | 0.96 | 1.8 | 116 | 0.056 |
| G3 | 141 | 0.83 | 1.6 | 108 | 0.13 | 164 | 1.0 | 1.4 | 114 | 0.039 |
| A4 | 154 | 0.86 | 0.39 | 133 | 0.126 | 164 | 0.95 | 1.0 | 118 | 0.018 |
| G5 | 151 | 0.88 | 0.63 | 110 | 0.071 | 164 | 1.0 | 1.0 | 118 | 0.013 |
| A6 | 192 | 0.76 | $3 \times 10^{-4}$ | 107 | 0.071 | 183 | 0.96 | 0.36 | 123 | 0.053 |
| T7 | 151 | 0.91 | 0.043 | 111 | 0.092 | 167 | 0.98 | 1.5 | 119 | 0.082 |
| C8 | 166 | 0.69 | 0.047 | 113 | 0.19 | 164 | 0.85 | 0.61 | 126 | 0.025 |
| T9 | 148 | 0.74 | 0.21 | 107 | 0.18 | 141 | 0.88 | 0.26 | 122 | 0.01 |
| C10 | 113 | 0.59 | 0.094 | 122 | 0.176 | 127 | 0.7 | 0.1 | 117 | 0.089 |
| T11 | 133 | 0.83 | 0.11 | 106 | 0.09 | 151 | 0.83 | 0.24 | 117 | 0.089 |
| C12 | 80−190 | 0.8 | $2.4 \times 10^{-4}$ | 107 | 0.089 | 90−180 | 0.75 | nd | anti | nd |
| C13 | 151 | 0.69 | 0.57 | 110 | 0.16 | | | | | |
| T14 | 148 | 0.8 | 0.28 | 103 | 0.12 | | | | | |
| C15 | 151 | 0.64 | 0.02 | 125 | 0.3 | | | | | |
| T16 | 158 | 0.89 | 0.9 | 112 | 0.068 | | | | | |
| C17 | 121 | 0.53 | 0.7 | 122 | 0.09 | | | | | |
| T18 | 136−195 | 0.81 | $5.5 \times 10^{-4}$ | 117 | 0.25 | | | | | |

[a] Sugar parameters were determined from coupling constants using the program pfit and glycosidic torsion angles were determined using the program NUCFIT (*38*) (see text). $P_s$ is the pseudorotational phase angle of the major (*S*) conformer, $f_s$ is its mole fraction, and % E is the percent error between calculated and observed coupling contants. $\chi$ is the glycosidic torsion angle and *R* is the NOE residual index.

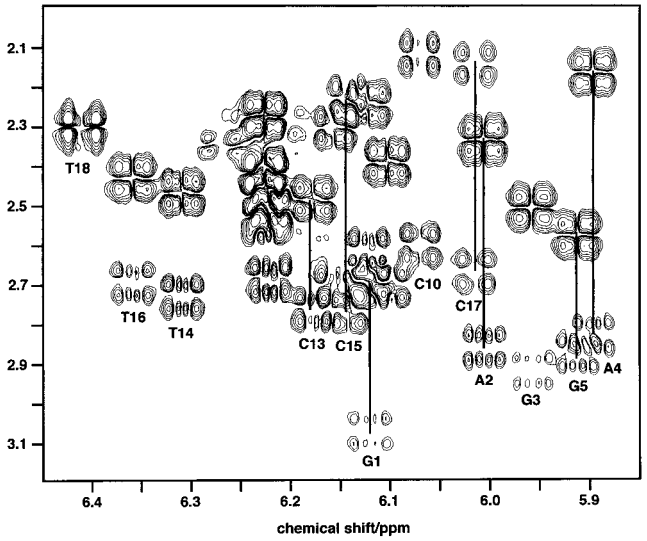


FIGURE 6: DQF-COSY spectrum of d(GAGAGA-oct-TCTCTC-oct-CTCTCT). The spectrum was recorded at 500 MHz and 30 °C with acquisition times of 0.455 s in $t_2$ and 0.07 s in $t_1$. The data matrix was zero-filled once in $t_2$ and twice in $t_1$ and apodized using a slightly resolution-enhancing Gaussian function. The final digital resolution was 1.1 Hz/point in *F*2 and 4.39 Hz/point in *F*1.

that the sugars of these residues are in the *S* conformation. The DQF-COSY spectrum (Figure 6) showed characteristic antiphase patterns for the H1′−H2′ and H1′−H2″ cross-peaks, from which values of $\sum_{1'}$ could be measured. For all residues, $\sum_{1'}$ was >13 Hz, which indicates that the dominant conformation is *S*. This observation effectively rules out models of this triplex that are in the A conformation.

At 30 °C, the correlation time determined from the NOE build-up curve of the cytosine H6−H5 vectors was 2.4 ns. Under these conditions, the line widths are smaller than the coupling constants, so that the DQF-COSY spectra can be used for analysis of the sugar conformations (*35*). The estimated line widths for H1′ and for H2′ and H2″ are 1.5−2 and 5−6 Hz, respectively. The cross-peaks in Figure 6 show considerably different fine structure for some residues, indicating significantly different sugar conformations. For example, the cross-peaks for H1′−H2′ for A4, A2, T14, and T16 show a clear pattern of ++ − − in the fine structure, whereas the H1′−H2″ peaks show the expected alternating + − + − pattern typical of *S* conformations containing a small (<20%) admixture of the C3′-*endo* conformation. In contrast the H1′−H2′ cross-peaks of C13, C10, and C17 shows a simple + − pattern with a large gap in between them owing to cancellation of the antiphase components. Similarly, the H1′−H2″ cross-peaks of these residues appear very similar to the H1′−H2′ cross-peaks, indicating that $^3J_{1'2'} \approx {}^3J_{1'2''}$. It is notable also that the apparent value of $\sum_{3'}$ is largest for the cytosine residues. The sums of the coupling constants and the H1′−H4′ distance were analyzed as previously described (*35*). In many instances, a unique sugar conformation was not consistent with the data, whereas a two-state model was. Assuming that the minor conformer has a value of *P* of ca. 9°, we have analyzed the conformations of the sugar puckers in terms of the pseudorotation angle of the major (*S*) conformer *P*s, and the mole fraction of this conformer, $f_s$. As Table 1 shows, the residues in the purine strand have a high fraction of the *S* conformation ($f_s$ > 0.8) with pseudorotation phase angles in the range 130−150°. In the Watson−Crick pyrimidine strand, the conformational purity is typically lower, in the range 60−90% *S*. The Hoogsteen strand showed the lowest conformational purity, especially for the three cytosine residues, and somewhat lower pseudorotation phase angles.

The trends in sugar conformations observed in this triplex are similar to those reported for a different intramolecular triplex (*10*), where the lowest $f_s$ values were also observed for the cytosine residues, in both the Hoogsteen and the Watson−Crick strands. This is in contrast with published structures that indicate pure conformational states for all nucleotides (*5*−*8*, *18*). When the populations of the *S* and *N* states are nearly equal, numerous conformationally sensitive NOEs are strongly affected, which significantly influences the structure determination when tight distance bounds are used (*42*).

The glycosidic torsion angles were determined by fitting NOE build-up curves using the program NUCFIT (*38*) with mixing times of 50, 100, 150, 200, and 250 ms. Sugar

conformations were searched in the neighborhood of those found from analysis of the coupling constants (Table 1), with nonlinear regression at each starting conformation (typically 48 starting conformations/nucleotide). In all cases, the best fit to the data (lowest *R* factor) was found for a conformational blend in which the glycosidic torsion angle for the *S* pucker was in the range −100° to −130° and that for the *N* pucker was near −160°. The calculations converged independently of the starting structure, and the high ratio of data (ca. 20 observations) to parameters (3) means that the nucleotide conformations are well-determined. The nucleotide conformations determined in this way are given in Table 1. The glycosidic torsion angles are typical of nucleotides found in the B family of conformations, with values of the purines on average −112° ± 10°, and −111° ± 6° and −115° ± 8° for the Watson−Crick and Hoogsteen pyrimidine strands, respectively. Hence, the nucleotides are predominantly found in the range typical of B-DNA, though there is a substantial fraction of A-like conformations for the cytosines in the Hoogsteen strand. The high *R* value obtained for Cyt 15 needs comment. A lower value was actually achieved for a *syn* conformation; however, this structure was inconsistent with internucleotide NOEs (see below). Some additional conformational averaging could affect the observed NOE intensities for this residue, which is best accommodated in the three-dimensional structural model as a normal *anti* glycosidic torsion angle (see below). Interestingly, there is an alternation in the magnitude of the glycosidic torsion angles that parallels the alternating sequence.

Further information about the nucleotide conformations was obtained from measurements of $\sum_{4'}$. The rotamer for the torsion angle $\gamma$ was estimated from the value of $\sum_{4'}$ and relative NOE intensities H8/H6 to H5′/H5″. For most of the nucleotides, $\sum_{4'}$ was estimated as <10 Hz, which, given a coupling constant $^3J_{3'4'}$ of 3−4 Hz, requires that both $J_{4'5'}$ and $J_{4'5''}$ are less than 6−7 Hz. This shows that the dominant torsion about $\gamma$ is near 60° ($g^+$). One exception was noted, namely, T7, where $\sum_{4'} \approx 16$ Hz. For this nucleotide, rotamers other than $g^+$ must be significantly occupied.

*NMR Spectroscopy of the Duplex d(GAGAGA-oct-TCTCTC).* In order to obtain information about possible changes in conformation between the parent duplex and the intramolecular triplex, we have analyzed the intramolecular duplex d(GAGAGA-oct-TCTCTC) under identical conditions of salt concentration, pH, and temperature. The spectra were well dispersed (see below), and assignment of the proton signals was straightforward. The proton assignments are available as Supporting Information.

There were numerous large changes in chemical shifts of the purine strand in the duplex compared with the triplex (Figure 7), as would be expected as the third strand hydrogen-bonds directly to the purine strand. In contrast, the changes for the Watson−Crick pyrimidine strand were much smaller, reflecting their greater distance from the Hoogsteen bases. The difference between the shifts for the adenines and the guanines was particularly pronounced. Thus, the chemical shifts of H8, H2″ and H3′ of the Ade residues showed substantial negative changes, whereas the largest changes for the Gua residues occurred for H1′ and H2″, which were positive. Protonation of the cytosines should reduce their aromaticity and therefore ring current fields. This may in part account for the differences in $\Delta\delta$(H8) between the A and G residues. However, the large differential shifts of the minor-groove protons suggests conformational changes rather than direct ring current effects and suggests that the conformational changes are not uniform along the sequence. This is perhaps not surprising as the mode of stacking of a protonated C is likely to be very different from that of an unprotonated C or a T.

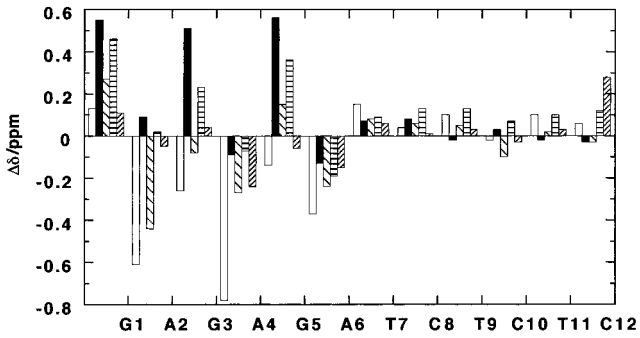


FIGURE 7: Chemical shift differences between d(GAGAGA-TCTCTC-CTCTCT) and d(GAGAGA-TCTCTC). The chemical shift differences of selected protons were calculated for triplex − duplex using the data tabulated in the supporting information. Positive differences are downfield. Open bars, H8/6; solid bars, H1′; widely hatched bars, H2′; horizontally striped bars, H2″; narrowly hatched bars, H3′.

The nucleotide conformations were analyzed in the same way as described for the triplex helix. The values of $\sum_{1'}$ in the purine strand were essentially identical in the duplex and triplexes. Small differences in $\sum_{2'}$ indicate that there may be some change in the average conformation of the sugars. The sugars of the purines have essentially pure conformations, which are nearer C2′-*endo* in the duplex than in the triplex (Table 1). The pyrimidine strand shows somewhat larger differences in sugar conformation, especially for the cytosine residues, with apparently a decrease in order in the triplex compared with the duplex.

The glycosidic torsion angles in the duplex are slightly larger (ca. 10°) on average than in the triplex and are more uniform along the sequence in the duplex than in the triplex (Table 1); the alternation along the sequence seen in the triplex is not present in the duplex. This may be related to the strong alternating pattern of chemical shift changes that were observed between the duplex and triplex (Figure 7). However, the changes in conformation of the Watson−Crick nucleotides are quite small on adding the third strand.

*Molecular Modeling.* In the triplex, the sequential NOEs in general were found to be typical of B-like helical structures, e.g., H8($i$)−H2″($i$ − 1) ≫ H8($i$)−H2′($i$ − 1), except for some steps in the Hoogsteen strand, where these NOE intensities were almost equal. This could indicate an intermediate structure or be the result of conformational averaging. Calculations have shown that the ensemble averages of 50:50 mixtures of *N* and *S* states produces sequential H2′ and H2″−H8/6($i$ + 1) NOEs of similar intensity and therefore similar apparent distances (*42*).

In addition to the expected sequential NOEs that typify right-handed helices, we have identified numerous NOEs between the third strand and the two Watson−Crick strands. These NOEs confirm the orientation of the third strand and the Hoogsteen base-pairing. In addition, they provide valuable restraints for the determination of the three-dimensional structure of the triplex.

Table 2: Statistics of the Structure Refinement[a]

| triplex (docked structures) | | | | rmsd values (Å) duplex (refined structures) | | | | | refined triplexes | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | BA | AB | AA | | Bini | Aini | Bf | Af | | AAini | BB | BBf |
| BB | 1.13 | 1.03 | 1.06 | Bini | | 3.21 | 1.35 | 1.48 | BBf | 2.13 | 1.33 | 0.51 |
| BA | | 1.09 | 1.22 | Aini | | | 2.97 | 2.91 | | | | |
| AB | | | 0.99 | Bf | | | 0.37 | 0.69 | | | | |
| | | | | Af | | | | 0.44 | | | | |

| | $U_{pot}$ (kcal mol$^{-1}$) | $U_f$ (kcal mol$^{-1}$) | $U_b$ (kcal mol$^{-1}$) | $U_{ang}$ (kcal mol$^{-1}$) |
|---|---|---|---|---|
| triplex | $-120.9 \pm 7$ | $0.32 \pm 0.06$ | $4.3 \pm 0.5$ | $58.9 \pm 1.2$ |
| duplex | $-90.5 \pm 2.6$ | $0.57 \pm 0.16$ | $2.7 \pm 0.3$ | $35.7 \pm 0.4$ |

[a] Pairwise rmsd values were calculated for all atoms for refined structures starting from different initial structures. For the triplex, BB refers to B-DNA duplexes docked with a B DNA third strand, BA refers to B-DNA duplexes docked with an A-DNA strand etc. BBf are the final refined triplex structures, BB are the intermediate B-like triplex structures, and AAini is a triplex in the A conformation for all strands (see text). For the duplex, ini and f refer to initial and final (refined) structures starting from A or B conformations. Bf versus Bf and Af versus Af are pairwise comparisons between different MD runs using different random number seeds. $U_{pot}$, $U_f$, $U_b$, and $U_{ang}$ are the potential energy, the residual restraint energy, the bond energy, and the bond angle energy, respectively.

The NOE intensities in the duplex are typical of a right-handed B-like helix, as expected from the analysis of the nucleotide conformations (Table 1).

*Structure of the Triplex.* Restraints for the nucleotides were taken from the detailed analysis of the time-dependent NOEs and coupling constants as described above. The nucleotides were considered to consist of a mixture of two conformational states, namely, $\chi(S)P(S)$ and $\chi(N)P(N)$. Within the context of this model, we have derived the parameters corresponding to the more highly populated *S* state. Intranucleotide distances and torsion angles ($\chi$ and $\delta$, plus $\gamma$ for some residues) were calculated and used as relatively tight restraints for the structure calculations. After initial rounds of calculations had produced a family of consistent structures, the intranucleotide constraints were tightened further, and addition rounds of rMD were carried out.

Internucleotide NOE intensities were converted into loose distance constraints according to relative NOE cross-peak intensities measured from NOESY spectra and were classified as s(trong) (<2.6 Å), m(edium) (<3.5 Å), and w(eak) (3−5 Å). Hydrogen-bond constraints based on the chemical shifts and exchange behavior of the imino and amino protons and cross-strand NOEs were used for the Watson−Crick base pairs, and additional constraints for the Hoogsteen pairs were supplied. In total, we have quantified 295 NOEs. For the structure calculations, 279 NOE-derived conformationally sensitive distance constraints were used, which included 90 intranucleotide, 152 sequential, and 29 cross-strand distances. Including 27 hydrogen-bond constraints, a total of 306 distances (17/residue) was used. In addition, torsion constraints were supplied as follows: 12 $\gamma$, 16 $\chi$, and 18 $\delta$, for a total of 352 restraints (19.5/residue). There are 132 undetermined torsion angles in the structure (ignoring the linkers), which gives a ratio of experimental data to degrees of freedom of 2.67:1.

Initial triplex structures were generated as described in the Methods section. Table 2 gives the statistics of the refinements. The initial docked structures refined to an average rmsd of $1.1 \pm 0.1$ Å. The agreement with the data and between different starting structures shows that the models are stereochemically reasonable and quite well-determined. The nucleotide conformations in particular are well-determined, but large fluctuations in position or torsion angles were found for backbone atoms, particularly the backbone atoms in the Hoogsteen strand. These models ignore the effects of conformational averaging in the nucleotides, which is especially prevalent in the cytosine residues and in the third strand in general. Because we have determined the conformations of the nucleotides as a mixture of *S* and *N* states, the models have accounted for averaging of each nucleotide but not directly for the internucleotide interactions. We are currently calculating ensembles of structures to comply directly with the experimental data.

Further starting points were A and B duplexes minimized in the presence of constraints appropriate to the duplex. These two initial structures refined to B-like duplexes that had an rmsd of 0.8 Å, indicating that the constraints determine the conformation. These structures were used for docking of the third strand, followed by rMD and energy minimization. This procedure was found to be more efficient and gave a higher yield of structures having good stereochemistry and low residual violations. After an acceptable number of structures had been obtained that satisfied the experimental restraints, the best eight, chosen according to their potential energy, residual energy violations, and maximum distance or torsion violations, were subjected to further refinement. The intraresidue constraints were tightened to reflect the observed NOE intensities and coupling constants, and the system was heated to 1000 K for 1 ps without constraints, with the electrostatic potential scaled by 0.2 and the Lennard−Jones term by 0.5, minimized for 500 cycles with conjugate gradients with constraints, followed by 10 ps of rMD at 300 K and 4000 steps of energy minimization using conjugate gradients. The scaling factors were then returned to unity for determination of the energy of the system. This resulted in the final energies and rmsd values as shown in Table 2. The pairwise rmsd values for these structures was $0.51 \pm 0.12$ Å. These structures had rmsd values of $1.33 \pm 0.11$ Å to the minimized, docked structures and 2.13 Å to a triplex in the A conformation for all three strands. Furthermore, the restraint violations were few and very small (no violations >0.1 Å or 2°) (final total restraint energy $U_f = 0.32 \pm 0.06$ kcal mol$^{-1}$). The bond length ($U_b$) and bond angle ($U_{ang}$) energies were also small (Table 2), i.e., no larger than those obtained from unrestrained energy minimization of DNA triplexes. Hence the structures agree well with the experimental data and have good stereochemistries.

Table 3: Averaged Torsion Angles and Helicoid Parameters in the Triplex and Duplex[a]

| | α | δ | *P* | γ | χ | twist | rise | X-disp | pr tw | inc |
|---|---|---|---|---|---|---|---|---|---|---|
| triplex | −70 ± 3 | 126 ± 6 | 125 | 74 ± 2 | −123 ± 5 | 35 | 3.15 | −3 | −19 | −5.7 |
| duplex | −69 ± 1 | 126 ± 4 | 130 | 58 ± 2 | −123 ± 4 | 36 | 3.1 | −2.5 | −22 | −0.7 |
| B-DNA | −41 | 139 | 192 | 38 | −98 | 36 | 3.38 | −0.71 | 3.7 | −5.9 |
| A-DNA | −50 | 79 | 13 | 41 | −154 | 32.7 | 2.56 | −5.4 | 13.7 | 19.0 |
| YR·R[b] | na | na | S | na | −70/−110 | 30 | 3.6 | −1.9 | na | na |
| RY·Y[c] | na | na | 124 | na | −130 | 30 | 3.4 | −1.9 | na | na |
| YR·R[d] | −131 | 128 | 142 | 118 | −121 | 38 | 3.3 | −2.9 | −11 | −5.4 |
| YR·Y[e] | na | 90−120 | S | na | na | 31 | 3.2 | −1.4 | na | na |
| YR·Y[f] | na | na | S | na | na | 29 | 3.1 | −2.9 | na | na |

[a] pr tw is the propeller twist and inc is the base-pair inclination to the helix axis. na, not available. [b] Reference *6*. [c] Reference *7*. [d] Reference *18*. [e] Reference *12*. [f] Reference *13*.

As many of the backbone torsion angles are determined by default and are highly dependent on the details of the force field, we have concentrated on those that are directly determined by the experimental data or that are strongly correlated with these torsions, namely, $\gamma$, $\chi$, $\delta$, and $\alpha$. Table 3 shows a summary of the torsion angles averaged over the refined structures. In addition, we have used the Curves program to calculate additional helical parameters, which are also summarized in Table 3. In general, the twist and rise are reasonably well determined by the experimental data (*38*), but other parameters such as X-displacement and inclination are less well determined by the data and depend to some extent on the modeling process. However, these latter parameters are characteristic of global helix morphologies and are commonly quoted. It is notable that the variation of torsion angles is greatest in the Hoogsteen strand, both between residues, and between structures.

The duplex part of the triple helix is characterized by relatively high values of $\chi$, a low rise, and a substantial displacement of the helix axis toward the minor groove. The latter is unexpected, as a large ligand has been placed in the major groove. However, triple helices are also unusual in that they are based on a polypurine·polypyrimidine sequence, which does not necessarily have the same structure as a mixed sequence duplex. The structure of $d(A)_n \cdot d(T)_n$, for example, is significantly different from that of mixed-sequence duplexes (*25*), and alternating GA sequences have been reported to form alternative structures under some conditions (*26*). It is therefore important to compare the structure with that of the appropriate duplex.

*Structure of the Watson−Crick Duplex.* Restraints for the duplex were determined as described above, and comprised in total 140 distances (60 intraresidue, 50 sequential, 15 exchangeable proton distances, and 15 H-bonds) plus 36 torsion ($\chi$, $\gamma$, and $\delta$) constraints (i.e., 14.7 restraints/nucleotide on average). Because the nucleotides had a high conformational purity according to the detailed analysis of the coupling constants (Table 1), the structures were analyzed as a single family of conformations. Structures starting from A or B DNA converged to a pairwise rmsd over all atoms of about 0.44 ± 0.11 Å and 0.37 ± 0.17 Å, respectively, and for all structures 0.51 ± 0.12 Å (Table 2). The average torsion angles and helicoid parameters are shown in Table 3. These data indicate that the duplex is in the B family of conformations. However, its structure differs significantly from the canonical B structure, as shown by the parameters in Table 3 and the energy of restraint violations of 199 kcal $mol^{-1}$ for the B-DNA structure. For comparison, the refined duplexes gave a final restraint violation energy of 0.57 ± 0.16 kcal $mol^{-1}$. The values of $\chi$ are somewhat larger than for mixed sequence B-DNA in solution, and the propeller twist is large and negative. Although the twist and inclination are typical of the B-form, the rise is significantly smaller, and the displacement of the helix axis into the minor groove is relatively large compared with standard B-DNA. However, although the convergence from A and B DNA duplex was good, the data probably do not determine the value of the axis displacement very well in the duplex, as simple energy minimization also tends to move the helix axis into the minor groove. Nevertheless, the axis displacement is larger in the triplex than in the duplex. Similarly, the magnitude of the propeller twist is not well-determined, though for the present molecules, a large value was consistently found.

The structures of the duplex in the free and triplex state are in many respects very similar; the rmsd between the free duplex and the duplex part of the triplex was 0.88 ± 0.17 Å, which is only slightly larger than the rmsd values for the refined triplexes and the refined duplexes. This suggests that the duplex is predisposed to accepting the pyrimidine strand into its major groove and only minor readjustments of the duplex are needed to form this triplex.

## DISCUSSION

Ten superimposed structures of the refined triplex are shown in Figure 8. The Hoogsteen strand lines the major groove of the Watson−Crick duplex and makes contacts with the purine strand. The positions of the bases are well-determined by the data, and a significant amount of the rmsd arises from relatively poor definition of some of the backbone positions, especially where no restraints on $\gamma$ could be obtained (see above). The major groove of the duplex is divided into two unequal grooves by the Hoogsteen strand, with widths approximately in the ratio 1:2 (purine side versus pyrimidine side), which is more clearly shown in the space-filling picture (Figure 9). The Watson−Crick bases show a considerable propeller twist (Table 3) that is consistent with values observed in other DNA triple helices (*12*, *18*) and DNA duplexes (*41*); the bases of the Hoogsteen strand lie further out of the mean planes formed by the Watson−Crick base pairs. This in fact brings the pyrimidines of the Watson−Crick and Hoogsteen strands relatively close. Figure 10 shows the stacking pattern of the $CG \cdot C^+$ triple on the 3′ TA·T triple. There is extensive base overlap between the two purines. Considerably less overlap is present in the Watson−Crick TC step or the Hoogsteen $C^+T$ step. However, the stacking interactions for the two

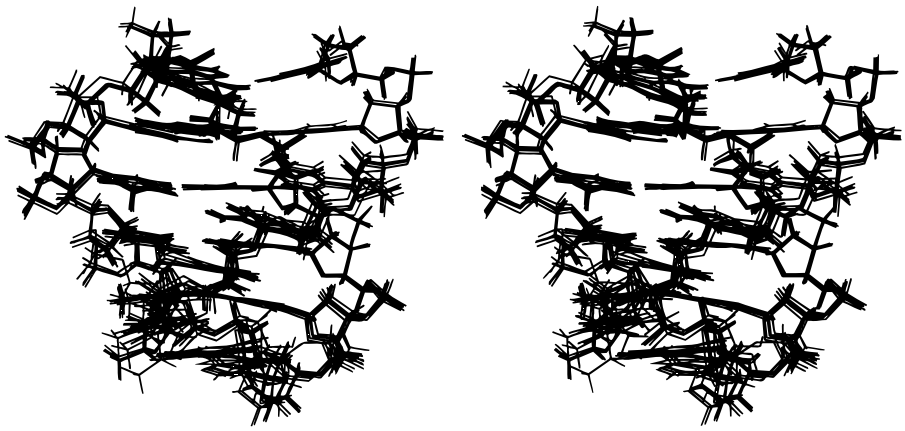

FIGURE 8: Solution structure of d(GAGAGA-oct-TCTCTC-oct-CTCTCT): stereoview of 10 superposed triplex structures. The structures are shown looking at the backbone of the Hoogsteen strand in the major groove of the Watson−Crick duplex.

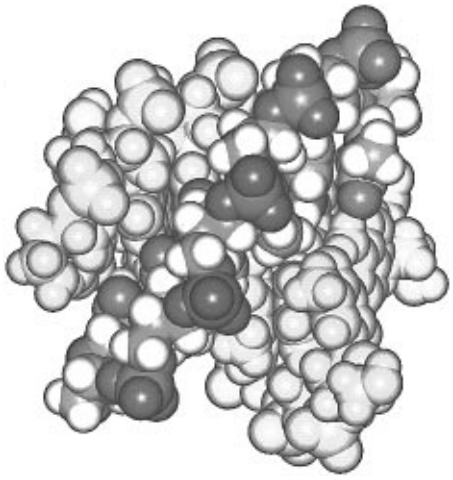

FIGURE 9: Space-filled structure showing the bisection of the major groove into two unequal grooves by the Hoogsteen strand.

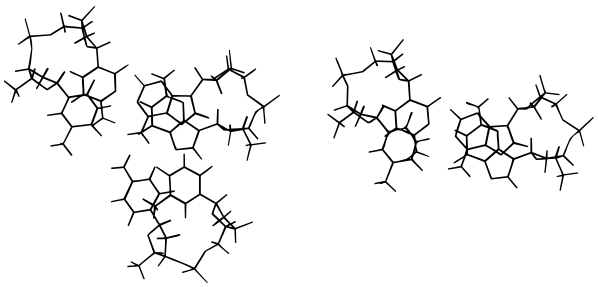

FIGURE 10: Base stacking in the C10G3·C15$^+$/T9A4·T16 steps (left) and in the C10G3/T9A4 of the free duplex (right).

pyrimidine strands are different. In the Watson−Crick TC step, the methyl group points out of the major groove toward the solution and makes no direct interactions, while the O2 lies over the face of the cytosine ring. In contrast, in the C$^+$ T (Hoogsteen) step, the methyl group lies under the face of the C$^+$ residue. It is also clear that the base stacking interactions in the Watson−Crick duplex are somewhat different in the free duplex compared with the triplex (Figure 10), which parallels the differences determined in some of the helicoid parameters (Table 3) and the different nucleotide conformations (Table 1).

Although overall the structures of the duplex are very similar in free and complexed duplexes, there are notable differences in detail. These seem particularly clear in the nucleotide conformations (cf. Table 1). It is also instructive to compare the present results with those obtained for other DNA triplexes. The averaged values of some parameters are shown in Table 3. One of the YR·R triplexes and the YR·Y triplexes contain a mismatch in the center of the molecule. Furthermore, we have analyzed our data according to a two-state conformational model for each nucleotide, and constraints were applied that are appropriate for the major conformer. This is likely to lead to differences in detail compared with other methods of data analysis. However, it is clear that many features of the triple helices are quite similar, with a main difference being the slightly lower rise that we find in our structure. The major result is that in all of the DNA triple helices so far examined, the sugars are largely in the *S* domain, the glycosidic torsion angles are near −120°, and $\gamma$ is primarily in the common $g^+$ rotamer. However, the nucleotides show extensive averaging in the pyrimidine strands, which necessarily means that some of the helicoid features are not particularly well-determined. Nevertheless, it seems clear that in the YR·Y motif, the axis displacement into the minor groove is large (2−3 Å) and the helical rise is lower than in canonical B-DNA but larger than in the A conformation, with 10.5−12 base pairs/turn.

The YR·R and YR·Y motifs can be superimposed surprisingly well (*7*, *8*), despite the nonisomorphous third strands running in opposite directions. Nevertheless, there are significant differences in local conformations, notably for the nucleotides. In the YR·R structure, the glycosidic torsions angles cover a wide range from −60° to −115°, with pseudorotation angles near 108° (C1′-*exo*), whereas in our YR·Y structure the glycosidic torsion angles are more uniform and near −120°, with the pseudorotation angles near 126−144°. This difference may in part determine the large difference in stability of triplexes containing parallel pyrimidine strands (stable) and antiparallel purine strands (unstable). The orientation of the bases in the third strand will determine the stacking and optimization of hydrogen-bonding interactions with the Watson−Crick purine strand. In this investigation we have shown that the pyrimidine-rich third strand slots into the major groove in a parallel orientation without major distortion of the preformed duplex. In general the stability of a triplex will depend on the propensity of the third strand to adopt a conformation that can fulfil this criterion. Design of third-strand oligonucleotides containing "triplex stabilizing" modified bases must take into account this constraint.

## SUPPORTING INFORMATION AVAILABLE

Seven tables (NMR assignments, torsion angles, sugar conformations, cross-strand NOEs, curves analyses, and thermodynamic data) and three figures (NOESY spectra of the triplex and the duplex and DQF-COSY spectrum of the duplex) (13 pages). Ordering information is given on any current masthead page.

## REFERENCES


1. Felsenfeld, G., Davies, D. R., and Rich, A. (1957) *J. Am. Chem. Soc. 79,* 2023−2024.
2. Plum, G. E., Park, Y.-W., Singelton, S. F., Dervan, P. B., and Breslauer, K. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87,* 9436−9440.
3. Rougee, M., Faucon, B., Mergny, J. L., Barcelo, F., Giovanangeli, C., Garestier, T., and Hélène, C. (1992) *Biochemistry 31,* 9269−9278.
4. Plum, G. E., and Breslauer, K. J. (1995) *J. Mol. Biol. 248,* 679−695.
5. Radhakrishnan, I., Patel, D. J., Veal, J. M., and Gao, X. (1992) *J. Am. Chem. Soc. 14,* 6913−6915.
6. Radhakrishnan, I., and Patel, D. J. (1993) *Structure 1,* 135−152.
7. Radhakrishnan, I., and Patel, D. J. (1994) *Structure 2,* 17−32.
8. Radhakrishnan, I., and Patel, D. J. (1994) *Biochemistry* 33, 11405−11415.
9. Sklenar, V., and Feigon, J. (1990) *Nature 345,* 836−838.
10. Macaya, R. F., Schultze, P., and Feigon, J. (1992) *J. Am. Chem. Soc. 114,* 781−783.
11. Macaya, R. F., Wang, E., Schultze, P., Sklenar, V., and Feigon, J. (1992) *J. Mol. Biol. 225,* 755−773.
12. Wang, E., Koshlap, K. M., Gillespie, P., Dervan, P. B., and Feigon, J. (1996*) J. Mol. Biol. 257,* 1052−1069.
13. Koshlap, K. M., Schultze, P., Brunar, H., Dervan, P. B., and Feigon, J. (1997) *Biochemistry 36,* 2659−2668.
14. Arnott, S., and Selsing, E. (1974) *J. Mol. Biol. 88,* 509.
15. Uemoto, K., Sarma, M. H., Gupta, G., Luo, J., and Sarma, R. H. (1990) *J. Am. Chem. Soc. 112,* 4539−4545.
16. Weersinghe, S., Smith, P. E., Mohan, V., Cheng, Y. K., and Pettitt, B. M. (1995) *J. Am. Chem. Soc. 117,* 2147−2158.
17. Betts, L., Joesy, J. A., Veal, J. M., and Jordan, S. R. (1995) *Science 270,* 1838−1841.
18. Ji, J., Hogan, M. E., and Gao, X. (1996) *Structure 4,* 425−435.
19. Han, H., and Dervan, P. B. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90,* 3806−3810.
20. Han, H., and Dervan, P. B. (1994) *Nucleic Acids Res. 22,* 2837−2844.
21. Durand, M., Peloille, S., Thuong, N. T., and Maurizot, J. C. (1992) *Biochemistry 31,* 9197−9204.
22. Dittrich, K., Gu, J., Tinder, R., Hogan, M., & Gao, X. (1994) *Biochemistry 33,* 4111−4120.
23. Rumney, S., and Kool, E. T. (1995) *J. Am. Chem. Soc. 117,* 5635−5646.
24. Sun, J-S., and Hélène, C. (1993) *Curr. Opin. Struct. Biol. 3,* 345−356.
25. Nelson, H. C. M., Finch, J. T., Luisi, B. F., and Klug, A. (1987) *Nature 330,* 221−226.
26. Lavelle, L., and Fresco, J. R. (1995) *Nucleic Acids Res. 23,* 2692−2705.
27. Subirana, J. A., and Faria, T. (1997) *Biophys. J. 73,* 333−338.
28. Chandler, S. P., Strekowski, L., Wilson, W. D., and Fox, K. R. (1995) *Biochemistry 34,* 7234−7242.
29. Grzybowski, J., McPhilips, F., and Brown, T. (1995) in *PCR: essential data* (Newton, C. R., Ed.) pp 93−98, Wiley, New York.
30. Brown, T., and Brown, D. J. S. (1992) *Methodss Enzymol. 211,* 20−35.
31. Ebel, S., Lane, A. N., and Brown, T. (1992) *Biochemistry 31,* 12083−12086.
32. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. Struct. 2,* 661−665.
33. States, D. J., Haberkorn, R. A., and Ruben, D. J. (1982) *J. Magn. Reson. 48,* 286−292.
34. Rance, M., Bodenhausen, G., Wagner, G., Wüthrich, K., and Ernst, R. R. (1985) *J. Magn. Reson. 62,* 497−504.
35. Conte, M. R., Bauer, C. J., and Lane, A. N. (1996) *J. Biomol. NMR 7,* 190−206.
36. Lane, A. N., Lefèvre, J.-F., and Jardetzky, O. (1986) *J. Magn. Reson. 66,* 201−218.
37. Birchall, A. J., and Lane, A. N. (1990) *Eur. Biophys. J. 19,* 73−78.
38. Lane, A. N. (1990) *Biochim. Biophys. Acta 1049,* 189−204.
39. Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1986) *Numerical Recipes,* Chapter 14, Cambridge University Press, Cambridge, England.
40. Lavery, R., and Sklenar, H. (1996) Curves 5.1 Manual, CNRS, Paris.
41. Saenger, W. (1984) *Principles of Nucleic Acid Structure,* Chapter 5. Springer Verlag, New York.
42. Lane, A. N. (1996) *Magn. Reson. Chem. 34,* S3−S10.


BI970710Q