Conformational and Thermodynamic Properties of Parallel Intramolecular Triple Helices Containing a DNA, RNA, or 2′-OMeDNA Third Strand

Juan Luis Asensio,[†,§] Reuben Carr,[‡] Tom Brown,[‡] and Andrew N. Lane*[,†]

Contribution from the Division of Molecular Structure, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK, and Department of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, UK

Received June 10, 1999

Abstract: The thermodynamic stability and solution conformational properties of three intramolecular triple helices based on the sequence AGAAGA-x-TCTTCT-x-TCTTCT (x is a non-nucleotide linker) comprising a DNA duplex and DNA, RNA, or 2′-OMeDNA third strands have been compared. The most stable triple helix contains the 2′-OMe third strand, followed by the triplex containing RNA in the third strand. Comparison of the NMR spectroscopic data for the RNA hybrid triplex with those of the all-DNA triplex shows that the duplex parts of the structure are very similar; the major difference is that the RNA strand is characterized by C3′-endo sugars (except the two terminal residues). In the all-DNA triplex γ has a substantial fraction of the trans rotamer for both of the internal adenine residues (A3 and A4), whereas in the free duplex γ is g^+ for these residues. In the RNA-containing triplex, only A3 shows the presence of $\gamma(t)$, and in the 2′-OMe state, both A3 and A4 are $\gamma(g^+)$. In addition, the 2′-OMe triplex shows conformational heterogeneity. Thus, there are sugar-dependent differences in the degree of distortion in the purine strand imposed by the third strand binding. The helical parameters for the underlying duplexes are very similar in all three triplexes. However, the helical parameters for the third strands are different for the DNA versus the RNA and 2′-OMe strands, reflecting their different sugar conformations. The lower degree of distortion of the underlying duplex in the presence of the 2′-OMe third strand is consistent with higher thermodynamic stability of this triplex compared with the greater distortion of the duplex induced by both DNA and RNA third strands.

Introduction

Oligonucleotides that would target the major groove of DNA, forming a triple helix and prevent transcription are possible antigene agents.[1] As such, they would be very effective agents for external regulation of gene activity, since this strategy bypasses the amplification steps inherent in transcription and translation that lead to the more traditional drug targets.[1,2] Triplexes can be assembled from mixtures of nucleic acid strands ranging from all-DNA to all-RNA structures. As DNA and RNA duplexes adopt radically different conformations in solution, there may be different global conformations of triple helices, depending on the composition of their strands. However, not all of the possible "hybrid" structures are stable.[3,4] Of the eight possible combinations of DNA (D) and RNA (R) strands of a given sequence, two of them (rPudPy·dPy and rPurPy·dPy) were unstable. Furthermore, the triplexes rPurPy·rPy and rPudPy·rPy were less stable than the other four possibilities, which were of comparable stability.[3] On the basis of the observed affinity cleavage patterns, it was proposed that there are two major families of conformation, one including dPudPy·dPy, dPudPy·rPy, dPurPy·dPy, dPurPy·rPy and the other comprising rPudPy·rPy + rPurPy·rPy.[4] These results may reflect the conformational properties of the target duplexes. It has been shown for short duplexes that dPu·dPy is in the B form and rPu·rPy is in the A form. In addition, the DNA·RNA hybrids have distinct properties intermediate between A and B forms, with dPu·rPy less A-like than rPu·dPy.[5,6]

NMR studies of RNA triple helices (RR·R) have indicated that the sugar puckers are C3′-endo and that these triplexes have features most in common with an A form structure.[7] In contrast, several studies of all-DNA triplexes have shown that they adopt a conformation more similar to the B-form, with C2′-endo sugar puckers, although with some A-like characteristics.[8−16] Thus,

* Corresponding author.
† National Institute for Medical Research.
‡ University of Southampton.
§ Present address: Dr. Juan Luis Asensio, Instituto de Quimica Organica General, Juan de la Cierva 3, Madrid 28006, Spain.

(1) Sun, J.-S.; Hélène, *Curr. Opin. Struct. Biol.* **1993**, *3*, 345−356.
(2) Gowers, D. M.; Fox, K. R. *Nucleic Acids Res.* **1999**, *27*, 1569−1577.
(3) Han, H.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3806−3810.
(4) Han, H.; Dervan, P. B. *Nucleic Acids Res.* **1994**, *22*, 2837−2844.
(5) Gyi, J. I.; Conn, G. L.; Lane, A. N.; Brown, T. *Biochemistry* **1996**, *35*, 12538−12548.
(6) Gyi, J. I.; Conn, G. L.; Lane, A. N.; Brown, T. *Biochemistry* **1998**, *37*, 73−80.
(7) Holland, J. A.; Hoffman, D. W. *Nucleic Acids Res.* **1996**, *24*, 2841−2848.
(8) Macaya, R. F.; Schultze, P.; Feigon, J. *J. Am. Chem. Soc.* **1992**, *114*, 781−783.
(9) Radhakrishnan, I.; Patel, D. J. *Structure* **1993**, *1*, 135−152.
(10) Radhakrishnan, I.; Patel, D. J. *Structure* **1994**, *2*, 17−32.
(11) Radhakrishnan, I.; Patel, D. J. *Biochemistry* **1994**, *33*, 11405−11415.
(12) Bartley, J. P.; Brown, T.; Lane, A. N. *Biochemistry* **1997**, *36*, 14502−14511.
(13) Tarköy, M.; Phipps, A. K.; Schultze, P.; Feigon, J. *Biochemistry* **1998**, *37*, 5810−5819.
(14) Asensio, J. L.; Brown, T.; Lane, A. N. *Structure* **1999**, *7*, 1−11.
(15) Nunn, C. M.; Trent, J. O.; Neidle, S. *FEBS Lett.* **1997**, *416*, 86−89.

10.1021/ja991949s CCC: $18.00
Published on Web 11/13/1999

the all-DNA and all-RNA triplexes probably do form two distinct classes of structures. A DNA.DNA·RNA hybrid triplex (DD·R) was shown to be stable, with the RNA strand in the expected position in the major groove.[17] In a recent study it was concluded that the underlying duplex was essentially in the same conformation for either the DNA or the RNA third strand, and that the sugars of the RNA third strand were mixed C2′-endo/C3′-endo.[18] Furthermore, the triplex containing the RNA third strand was slightly ($\Delta T_m = 2.5$ K) more stable than the analogue containing a DNA third strand.[18] In contrast, molecular mechanics calculations[19] indicated a potential energy for the DD·D triplex lower than that for DD·R, and differences in conformation. Thus, the relationship between strand composition, conformation, and stability is not clear.

One of the problems with parallel triplexes is their relatively low thermodynamic stability at physiological pH. A simple modification that is generally thermodynamically stabilizing is methylation of the 2′-O of ribose. O2′-methylation of one of the strands of an RNA duplex is slightly stabilizing or destabilizing,[20,21] whereas methoxylation of a DNA strand in a DNA duplex or the DNA strand in a DNA·RNA hybrid duplex is generally stabilizing.[21,22] The stabilization is greatest when a DNA pyrimidine strand is methoxylated.[21] Methoxylation at the C2′ in the Hoogsteen strand of a parallel triple helix is stabilizing and more so than replacing a Hoogsteen DNA strand with the RNA analogue.[23] Hence, this simple modification can significantly enhance the stability of parallel triplexes.

Although there is a considerable body of information about the conformation of all-DNA triple helices from a variety of methodologies,[3−16] relatively little is known about the structures of the hybrid triple helices.[4,17−19,23] To understand the relationship between strand composition, conformation and thermodynamic stability of parallel triple helices, and the influence of the 2′-OMe modification, we have measured the pH dependence of the thermodynamic stability of three intramolecular parallel triple helices of the general sequence AGAAGA-x-TCTTCT-x-TCTTCT where x is the nonnucleotide linker $O-(CH_2)_8-OPO_2-O-(CH_2)_8OPO_2-O$. We have previously analyzed the thermodynamics and structures of the all-DNA analogue.[24,25] In addition, we have analyzed their conformations in solution using NMR, and we are able to correlate conformational differences with the observed order of stability.

Results

Thermodynamics. Figure 1A shows UV melting curves of the all-DNA triplex (DD·D) and the analogues containing either RNA (DD·R) or 2′-OMe (DD·R′) Hoogsteen strands at pH 7. Under these conditions, the melting curves are bimodal. The transition at the lower temperature corresponds to the melting of the triplex into duplex plus third strand, and the higher temperature transition corresponds to the melting of the duplex state into strands. As the pH is decreased to 5.5, the first transition moves to higher temperature and overlaps that of the duplex-strand transition. At neutral pH where the two transitions are well-resolved, the melting curve can be analyzed as a superposition of two sequential unfolding events as previously described,[25] from which the T_m values and van't Hoff enthalpy changes can be determined. As the latter transition is independent of pH (above pH 5), the melting curves at low pH can be analyzed by holding the duplex-strand transition fixed at its high pH values. The thermodynamic parameters for the three molecules are given in Table 1. The thermodynamic parameters for the triplex DD·D are similar to those previously determined.[25]

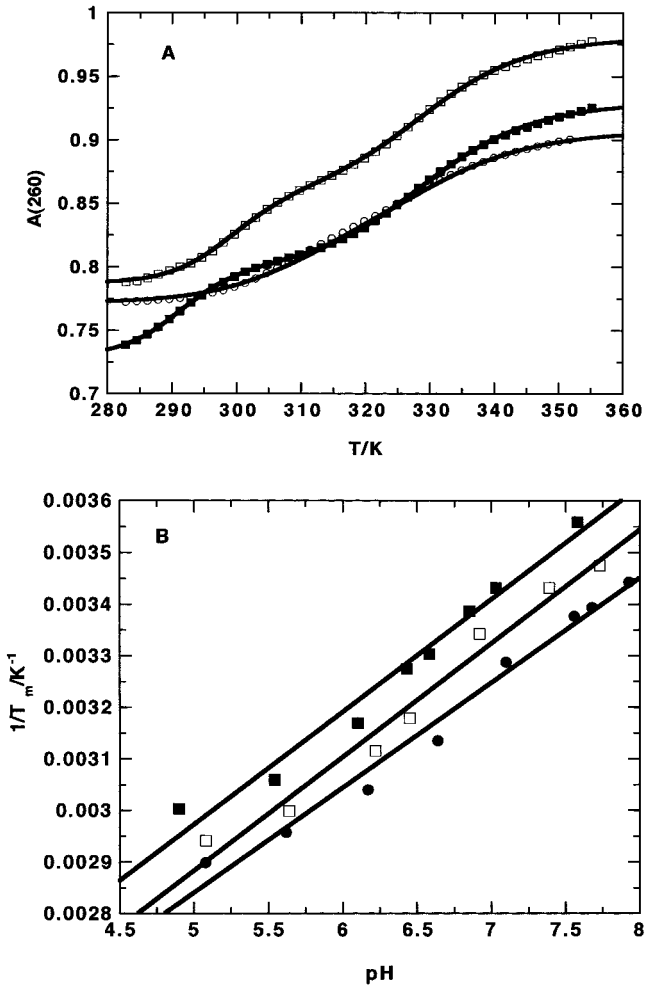


Figure 1. Thermodynamic stability of DD·D, DD·R′, and DD·R. (A) *A*(260) versus temperature. Continuous lines are best-fit regression lines according to eq (1,2). (■) DD·D at pH 7.03, (□) DD·R at pH 6.92, (○) DD·R′ at pH 7.10. For clarity, only every 10th datum is shown. (B) Dependence of $1/T_m$ versus pH. The slope of the lines is $(2.303R/\Delta H)\Delta p$ (see text). (■) DD·D, (□) DD·R, (●) DD·R′.

The value of $1/T_m$ decreases linearly with pH in the range 5.5−8 (Figure 1B), and the triplexes become too unstable to measure at higher pH values. As previously discussed[25] this is a consequence of the very high pK_a value of the protonated Hoogsteen cytosine residues in the triplex state, reflecting the large stabilization of the triplex by protonation. Below pH 5.5, the value of $1/T_m$ begins to tail off as the pK_a of the cytosines in the strand state (∼4.5) is approached. If Δp is assumed to

(16) Soliva, R.; Laughton, C. A.; Luque, F. J.; Orozco, M. J. *Am. Chem. Soc.* **1998**, *120*, 11226−11233.
(17) van Dongen, M. J. P.; Heus, H. A.; Wymenga, S. S, van der Marel, G. A.; van Boom, J. H.; Hilbers, C. W. *Biochemistry* **1996**, *35*, 1733−1739.
(18) Gotfredson, C. H.; Schultze, P.; Feigon. J. *J. Am. Chem. Soc.* **1998**, *120*, 4281−4289.
(19) Srinivasan, A. R.; Olson, W. K. *J. Am. Chem. Soc.* **1998**, *120*, 484−491.
(20) Inoue, H.; Hayase, Y.; Iwai, S.; Miura, K.; Ohtsuko, E. *Nucleic Acids. Res.* **1987**, *15*, 6131−6142.
(21) Lesnik, E. A.; Freier, S. M. *Biochemistry* **1998**, *37*, 6991−6997.
(22) Freier, S. M.; Altmann, K.-H. *Nucleic Acids Res.* **1997**, *25*, 4429−4443.
(23) Dagneaux, C.; Liquier, J.; Taillandier, E. *Biochemistry* **1995**, *34*, 16618−16623.
(24) Asensio, J. L.; Brown, T.; Lane, A. N. *Nucleic Acids Res.* **1998**, *26*, 3677−3686.
(25) Asensio, J. L.; Dhesai, J.; Bergqvist, S.; Brown, T.; Lane, A. N. *J. Mol. Biol.* **1998**, *275*, 811−822.

Table 1. Thermodynamics of the Three Triplexes[a]

	pH	T_m(T) (K)	T_m(D) (K)	ΔH_T[b]	ΔH_D[b]	ΔG_T^{298}[b]
DD·D	4.9	333	(329.5)	191	(110)	−16
	5.54	326.8	(329.5)	161	(110)	−13.4
	6.10	315.5	(329.5)	132	(110)	−8.4
	6.43	305.4	329	120	100	−3.7
	6.58	302.8	329.5	137	108	−2.4
	6.85	295.3	329.5	155	111	1.4
	7.03	291.4	329.2	150	114	3.4
	7.58	281	329	170	112	9.2
DD·R	5.08	340	(329.5)	158	(115)	−19.4
	5.64	333	(329)	133	(115)	−16.5
	6.22	321	(329)	183	(115)	−11.7
	6.45	314.6	(329)	125	(115)	−8.3
	6.92	299.2	328.6	150	113	−0.6
	7.39	291.4	328.8	152	110	3.6
	7.73	287.6	329.5	130	124	5.7
	8.39	<280	329	nd	121	>10
DD·R′	5.08	345	(329.5)	232	(110)	−23.3
	5.62	338.2	(329.5)	216	(110)	−20.3
	6.17	329.1	(329.5)	216	(110)	−16.2
	6.64	319	(329.5)	150	(110	−11.3
	7.10	306	329	133	110	−4.5
	7.56	296.4	329.5	160	100	0.9
	7.68	294.7	329.6	150	115	1.9
	7.93	290.5	329.3	133	113	4.4
	8.59	<280	329	nd	108	>11

	dG(298)/ dpH (kJ mol^{-1})	d(1/T_m)/ dpH[c] (K^{-1})	d(ln(K_{TH2})/ d(1/T) (K^{-1})	ΔH_{TH2}	Δp
DD·D	11.4	2.52×10^{-4}	-13.9×10^{-4}	116	2.00
DD·R	11.4	2.40×10^{-4}	-14.6×10^{-4}	121	1.97
DD·R′	11.3	2.215×10^{-4}	-16.3×10^{-4}	135	1.98

[a] T_m(T) is the melting temperature for the triplex−duplex, T_m(D) the melting temperature for the duplex, and ΔH_T, ΔH_D the van't Hoff enthalpies for the triplex and duplex transition, respectively. DD·D is the all-DNA triplex, DD·R′ is the 2′-OMe triplex, and DD·R is the RNA triplex. ΔG_T was calculated from ΔH and T_m as described in the text. Values in parentheses were fixed in the calculation (nearly coincident transitions). [b] kJ mol^{-1}. [c] d(1/T_m)/dpH, d(ln(K_{tH2})/d(1/T) calculated according to eqs 3 and 4 as described in the text. ΔH_{TH2} was calculated from the slope of the van't Hoff plot. Δp was calculated from eq 5 as (ΔH_{TH2} + 2ΔHion)d(1/T_m)/dpH)/2.303R).

be 2 (i.e., above pH 5.5, deprotonation of the strands is complete), then ΔH(app) can be obtained from d1/T_m/dpH (Table 1) using eq 4. ΔH(app) calculated by this method is 152, 173, and 160 kJ mol^{-1} for DD·D, DD·R′, and DD·R, respectively. These values are in reasonable agreement with the van't Hoff enthalpy determined from direct fitting (low pH values, and see below) (Table 1). The enthalpy of dissociation of the protonated triplex to the duplex plus protonated strand state can be calculated from the dependence of 1/T_m on pH by noting that at $T = T_m$, $K_{app} = 1$ and using K_{TH2} eqs 2 and 3. From the slope of the van't Hoff plots of ln(K_{TH2}) versus 1/T_m we obtain ΔH_{TH2}(DD·D) = 116 kJ mol^{-1}, ΔH_{TH2} (DD·R′) = 135 kJ mol^{-1} and ΔH_{TH2} (DD·R) = 121 kJ mol^{-1}. These values are lower than those obtained from direct fitting or from the slope of 1/T_m versus pH because the latter include the ionization enthalpy of the protonated cytosine in the strand state, which is ∼18 kJ mol^{-1}. Hence, for $\Delta p = 2$, the van't Hoff enthalpy determined from direct curve fitting should be ∼36 kJ mol^{-1} higher than the value estimated from the pH dependence of T_m. These corrected ΔH values are comparable to the average values obtained by curve fitting. It is noticeable, however, that the curve-fitted ΔH is apparently pH dependent, and decreases with increasing pH. At the higher pH values, the fitted ΔH are not consistent with the values obtained from evaluating the T_m data according to eq 3. The discrepancy at high pH arises from the incomplete UV curves at low temperature (minimum $T \approx 280$ K), so that lower baselines cannot be accurately defined. This gives rise to an erroneously low value of the directly fitted ΔH at high pH values (i.e., when $T_m < 290$ K). The enthalpy changes calculated from the pH dependence of T_m are likely to be more accurate. The enthalpy changes for the triplex−duplex transition are in the order DD·R′ > DD·R > DD·D (Table 1).

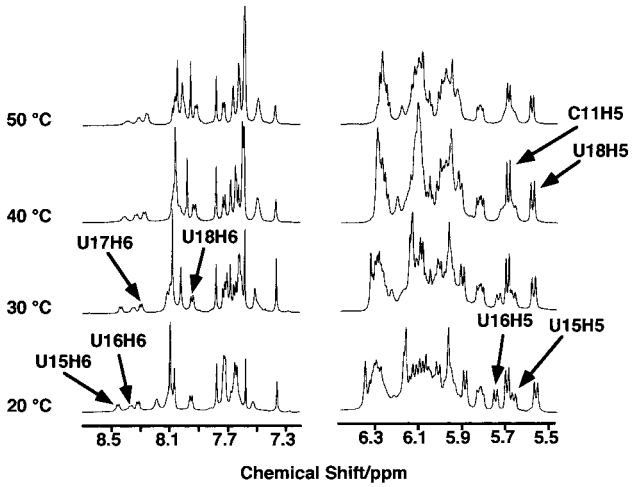


Figure 2. NMR spectra of DD·R′ showing chemical exchange in the 2′-OMe strand. 1D spectra were recorded at 600 MHz at the temperatures shown.

The O2′-Me derivative is substantially more stable ($\Delta T_m \approx$ 15 K) than the all-DNA triplex over the range 5 < pH < 8, which shows that the O2′-Me group has a substantial stabilizing influence on the triplex state. This stabilization is at least partly enthalpic. The enthalpy change is substantially larger for the O2′-Me strand than for the DNA strand as determined from the van't Hoff analysis (Table 1). The T_m values are in the order DD·R′ > DD·R > DD·D. However, ΔG is a thermodynamically more appropriate measure of stability. Using the experimental values of T_m and the derived values of ΔH, we have calculated ΔG at 298 K (Table 1); the ΔG(298) values are also in the order DD·R′ > DD·R > DD·D. The values of dΔG(298)/dpH are all similar indicating that the difference in stability remains constant at least up to pH 8. This is as expected if the pK_a is independent of the nature of the sugar, and pK_a′ ≫ pH (cf. eq 5). It is also possible to calculate ΔG(298) for the dissociation of TH_2 into D + SH_2 from the data in Table 1. We find ΔG(298) is −34, −30, and −26 kJ mol^{-1} for DD·R′, DD·R, and DD·D, respectively. This implies a stabilization by protonation of at least 34−40 kJ mol^{-1} (17−20 kJ mol^{-1} cytosine) compared with the fully unprotonated triplex. It is thought that the hydrogen-bonding capacity of the C2′-OH in RNA adds to the overall thermodynamic stability of RNA versus DNA.[26] Although the triplex containing the RNA third strand is more stable than the all-DNA triplex, it is less stable than the O2′-Me-containing triplex. Hence, the observed order of stability for these triplex is perhaps surprising. We have therefore used high-resolution NMR to determine the conformational properties of these triplexes and shed some light on their stability.

Conformation of the Three Triplexes. The NMR spectra of the all-DNA triplex DD·D have been analyzed previously.[24,25] Although the spectra of DD·D and DD·R show only a single set of resonances, the spectrum of DD·R′ shows additional peaks, as shown in Figure 2. In addition, at 30 °C the integrals of the U,C H6 and H5 of nonterminal residues in DD·R′ are

(26) Egli, M.; Portmann, S.; Usman, N. *Biochemistry* **1996**, *35*, 8489−8494.

only ~70% of those of the Watson−Crick CH6 resonances. Exchange cross-peaks involving these resonances were observed in ROESY spectra recorded at 40 °C (not shown), and the chemical shifts of the minor species do not correspond to those of the duplex + strand state. Hence, the observed exchange is due to two (or more) different triplex conformations. Furthermore H6 and H5 resonances become broader on increasing the temperature between 15 and 50 °C, showing increasing exchange.

To determine the sugar conformations, we have recorded DQF−COSY spectra of the three triplexes. The cross-peaks between H1′ and both H2′, H2″ indicated that the sugars of the DNA residues in the Watson−Crick strands are primarily in the "S" conformation in all three triplexes (not shown). This was confirmed by detailed analysis of the sums of coupling constants, which gives values of the pseudorotation phase angle P_s in the range 130−170° and fraction of the S state f_s of >0.9 in the purine strand and 0.6−0.8 in the pyrimidine strand. There is little difference in the sugar conformation and equilibria in the duplex moieties of the three triplexes, and they are all predominantly S. Furthermore, the sugar conformations determined for the isolated duplex DD are similar to those in the triplex states, indicating that only small changes in sugar conformations are induced by forming the triple helices. This is in contrast to an FT-IR study of triple helices containing an RNA third strand, where the purine strand was reported to contain "N"-type sugars.[23] In DD·D, the sugars in the Hoogsteen strand show a substantial fraction of the N state, especially the two protonated cytosine residues, though the thymines are predominantly S. In contrast, only the terminal Hoogsteen residues of DD·R and DD·R′ showed COSY H1′−H2′ cross-peaks. This indicates that the riboses in DD·R and DD·R′ are predominantly in the N conformation, except for the terminal residues which may exist as a mixture of N and S states. This conclusion is supported by the observation of strong H2′(i)−H6(i+1) NOE cross-peaks in the Hoogsteen strand of DD·R and DD·R′, which is characteristic of N-type sugars.

In the DD·D triplex, we have previously observed that the backbone angle γ for the two central A residues is t rather than the more common g^+.[14,24] TOCSY and ROESY spectra of the three triplexes were compared to determine γ for the residues in DD·R and DD·R′. The downfield shifted H4′ resonance, and the splitting of the H4′ resonance characteristic of the trans rotamer were evident for A3 in DD·R, but not for A4. All of the other H3′−H4′ cross-peaks for the deoxyriboses appeared as narrow singlets. In contrast, none of the deoxyribose H4′ of DD·R′ were unusually downfield-shifted, and these resonances appeared as narrow singlets (not shown). The apparent $\Sigma_{4'}$ values were estimated from the width at half-height for the singlets (which is an upper limit to $\Sigma_{4'}$), or from the splitting in NOESY or ROESY spectra (which gives a lower limit to $\Sigma_{4'}$) (Table 2). In addition to the coupling to H5′ and H5″, the observed $\Sigma_{4'}$ includes contributions from $^3J_{3'4'}$, the small $^4J_{4'-P}$, and the line width. As the deoxypurine residues are predominantly in the S state, $^3J_{3'4'}$ < 2−3 Hz. In the g^+ rotamer, both $^3J_{4'5'}$ and $^3J_{4'5''}$ are small (<3 Hz), whereas in either the t or g^- rotamers, one coupling is small (<3 Hz) and the other large (~12 Hz).[27] Hence, for $\Sigma_{4'}$ < 10 Hz, the dominant rotamer must be g^+. For A3H4′ and A4H4′ in DD·D, the observed half-width is ~14 Hz, which indicates that one of the two coupling constants must be large; the apparent splitting was >8 Hz. We have previously argued that this shows the presence of the t conformation for A3 and A4.[24] In DD·R, A3 showed a large half width (16 Hz), and a splitting of the H4′ resonance, indicating that this residue also may be trans about γ. However, the half-width for A4 is only 11 Hz, and did not show any splitting, which indicates that the contribution from the trans rotamer is small. This is different from DD·D. In DD·R′, the half width of A3H4′ is ~12 Hz, shows no splitting, and is not downfield-shifted, indicating that the contribution from $\gamma(t)$ is small. In A4, no significant contribution from $\gamma(t)$ is evident. Hence, there is a gradient of conformation effects in the purine strand on forming the triple helix that depends on the nature of the third strand, and the third strand containing the O2′Me modification does not induce the same conformation in the DNA purine strand as a DNA or RNA third strand. Finally, the width at half-height of the purine H4′ resonances in the duplex DD are all <10 Hz, indicating that all six purine residues are $\gamma(g^+)$.

Figure 3 shows a portion of a ROESY of DD·R′. Strong ROE cross-peaks are present between the 2′-methoxy protons and the H1′ of the same residue. This indicates that the methoxy methyl group is preferentially oriented toward its own H1′. In addition, there are significant ROEs between the methoxy protons and the H1′ of the next residue in the same strand, and with the H3′ of the purine residue at m-1 in the sequence. The observed NOEs can be explained by a single rotamer about the C2′−O2′ bond with the methyl group oriented toward the H1′ and away from the sugar ring.

Table 2. $\Sigma_{4'}$ Values for the Duplex Residues of the Three Triplexes $\Sigma_{4'}$[a]

	DD·D	DD·R	DD·R′	DD
A1	nd	<10	nd	nd
G2	<10	<10	<10	8
A3	14	16	12	10
A4	14	11	<10	9.5
G5	<10	<10	<10	9
A6	<10	<10	<10	8.5
T12	<10	<10	<10	<10
C11	nd	nd	<10	<10
T10	nd	nd	<10	nd
T9	nd	nd	<10	nd
C8	nd	nd	nd	nd
T7	<10	<10	<10	<10

[a] Values were estimated from the width at half height of the H3′−H4′ or H1−H4′ cross-peaks in NOESY spectra. This provides an upper limit to $\Sigma_{4'}$ (see text).

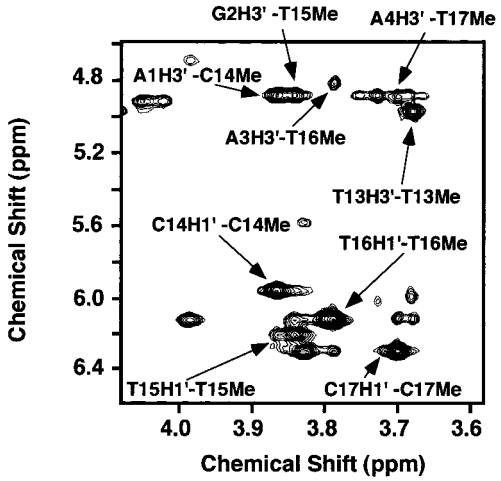


Figure 3. Interactions between O2′-Me protons and H1′ in DD·R′. The ROESY spectrum was recorded at 40 °C, pH 5 with a mixing time of 100 ms. Both intraresidue and cross-strand NOEs are shown.

(27) Wijmenga, S. S.; Mooren, M. N. W.; Hilbers, C. W. *NMR of Macromolecules. A Practical Approach*; Roberts, G. C. K., Ed.; IRL Press: Oxford, 1993. Chapter 8.

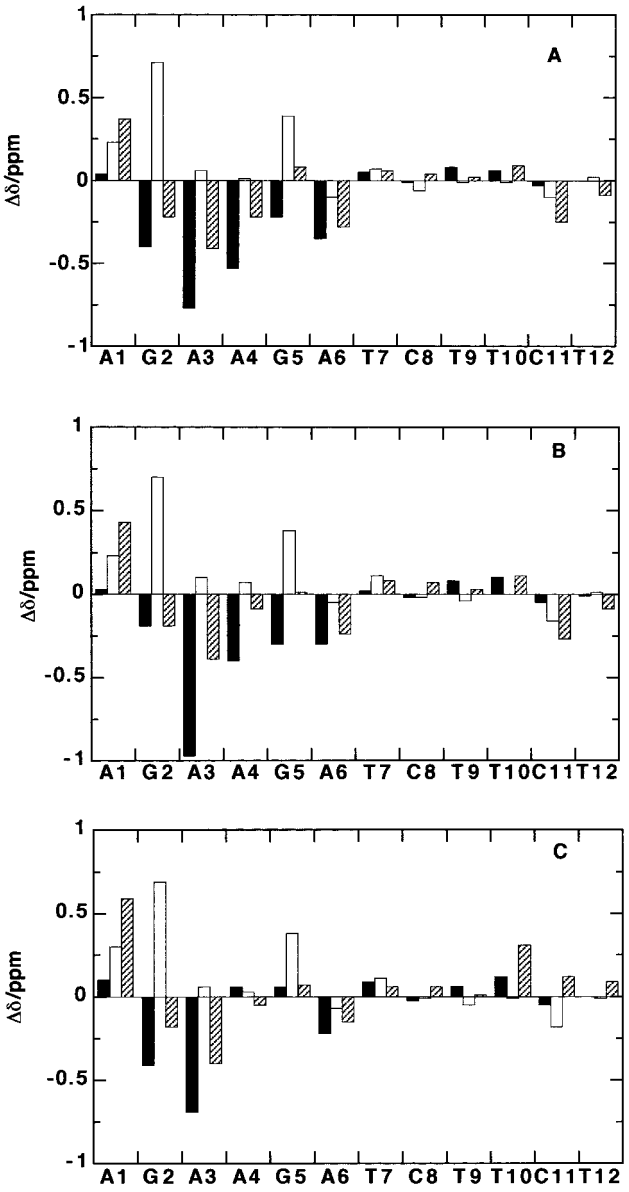


Figure 4. Changes in chemical shifts in the DNA duplexes on forming the triplexes. Chemical shift differences are reported as triplex−duplex. (■) H8/H6, (□) H1′, (□))H2′. (A) DD·D. (B) DD·R. (C) DD·R′.

The chemical shifts of the purine strand in DNA triplexes report on the conformation of the nucleotides. On forming the triplex, there are large changes in the shifts of the purine residues, especially H8 and H1′ [12,24] that reflect, among other things, changes in the glycosyl torsion angles and the torsion γ. Figure 4 compares the shifts of DD·D, DD·R, and DD·R′ with those of DD. The shift differences are much larger in the purine strand than in the pyrimidine strand, as expected for direct H-bonding of the Hoogsteen strand to the purine strand. Furthermore, in DD·D and DD·R the H8 resonances (major groove) show large negative changes for A3 and A4, and smaller ones for G2 and G5, whereas the H1′ resonances (minor groove) show large positive changes for G2 and G5, and small differences for A3 and A4. These differential changes in chemical shifts indicate conformational changes in the purine strand on forming the triplex. Interestingly, the pattern of shift changes for DD·D and DD·R are almost identical, whereas for DD·R′ there are large differences for resonances in A4 and G5. This suggests that the conformation of the duplex in DD·D and DD·R is very similar, whereas its conformation in DD·R′ is somewhat different. We have already shown that the conformations of the purine deoxyriboses are similar in all states (see above). However, the torsion γ(A4) varies from mainly *t* in DD·D to mainly g^+ in DD·R′. This is also correlated with the glycosyl torsion angle, such that χ(A4) is near −140 to −150° in DD·D to near −110 to −120° in DD·R′. These conformational differences presumably in part account differences in chemical shift. Indeed, in DD·R′, the conformation is more similar to that in DD, and the shifts for A4 and G5 are generally more similar to those in the free duplex. This conclusion is supported by the relative NOE intensities for A4H8−A4H2′ (intense in DD·R′ and DD, weak in DD·D, DD·R). Furthermore, the sequential H2″(i)−H8(i+1) and H2′(i)−H8(i+1) are of comparable intensity in DD·D and DD·R. In contrast, in DD·R′, the sequential H2″(i)−H8(i+1) NOEs are more intense than those of H2′(i)−H8(i+1) (not shown). Thus, the conformation of the purine strand of DD·R is similar to that in DD·D, and has undergone a conformational change on forming the triple helix. In contrast, there is less of a conformational change in the purine strand in DD·R′. Hence, there is a gradation of conformational changes in the order DD·D > DD·R > DD·R′.

Model Building. The solution structure of a related DNA duplex and triplex has been determined by NMR.[12,24] The DNA duplex was shown to be a normal B-like DNA duplex with fairly uniform nucleotide conformations along the sequence. NMR spectra of d(AGAAGA)·d(TCTTCT) are consistent with a uniform DNA structure, characterized by S type sugars, $\gamma(g^+)$ and strong H2′(i)-H8/H6(i) NOESY cross-peaks characteristic of $\chi \approx -110°$ to $-120°$. In addition the sequential NOEs H2″-(i)−H8/H6(i+1) > H2′(i)−H8/6(i+1).

The well-dispersed NMR spectra of the purine strands in all three of the triplexes made it possible to obtain a large number of distance restraints involving the purines. In addition, the analysis of the scalar coupling data allowed a complete specification of the nucleotide conformations of all of the purines. For the pyrimidine strands, the scalar data indicated the dominant sugar pucker, and sufficient NOEs were also observed to define the glycosyl torsion angles and provide some information about the helical structure. We have calculated model structures based on that of DD·D previously determined[24] with appropriate restraints on the nucleotide conformations as described above. For the purines the nucleotides were strongly constrained according to the analysis described above. The sugars of the two pyrimidine strands were also constrained tightly. Sequential and cross-strand NOE distance constraints were set loosely according to the classification strong, medium, and weak (cf. Methods). The coupling constants and observed sequential NOE intensities are consistent with a uniform B-DNA structure for the DNA duplex. In contrast, the DD·R triplex cannot be described as a standard B-like structure, both because of the changes in the nucleotide conformations in the purine strand and because the riboses in the RNA strand are C3′-endo. We have therefore calculated structures for this molecule using the experimental constraints on nucleotide torsion angles that correspond to the state where A3 has $\gamma = t$ and $\chi \approx -150°$.

Thirty structures were calculated for each triplex, and they were ordered according to potential energy and restraint energy. The best 10 chosen on these criteria were selected for further analysis. The superimposed structures are shown in Figure 5A. The pairwise rmsd values (all atoms) were 0.76 ± 0.09 Å (DD·D), 0.73 ± 0.24 Å (DD·R) and 0.89 ± 0.38 Å (DD·R′). In the pyrimidine strands of DD·R and DD·R′, the positions of the bases are reasonably well determined, but there is considerable variability in the backbone owing to the lack of restraints. This

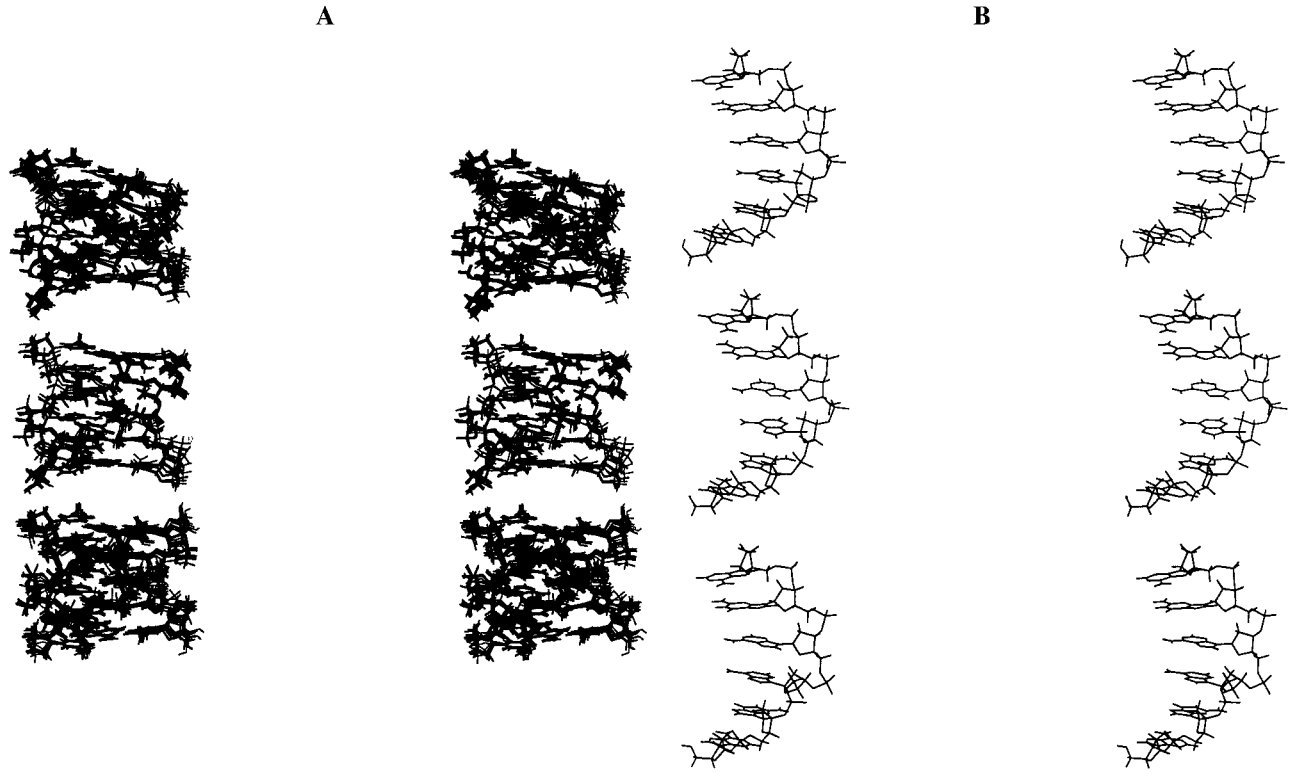


Figure 5. Comparison of solution conformations of DD·D, DD·R, and DD·R′. Models were calculated as described in the Experimental Section. Structures are shown as stereopairs. (A) Superimposition of 10 structures showing the lowest violations and potential energies. All three molecules are shown in the same alignment with respect to DD·R looking into the major groove of the DNA duplex: (Top) DD·D. (Middle) DD·R. (Bottom) DD·R′. (B) Comparison of the purine strands showing the variation in the phosphate backbone: (Top) DD·D. (Middle) DD·R. (Bottom) DD·R′.

Table 3. Helicoid Parameters in Refined Models[a]

parameter		DD·D	DD·R	DD·R′	B-DNA	A-DNA
twist(1,2)	deg	33.3	34.2	35.4	36	32.7
twist(1,3)	deg	34.9	39.7	41.6	–	–
rise	Å	3.16	3.12	3.13	3.4	2.6
Dx	Å	−3.3	−3.1	−3.7	−0.6	−5.4
inclination	deg	3.5	3.0	4.8	−6	19

[a] Local helicoid parameters were determined using CURVES 5.1.[28] Values have been averaged of the central four base-triplets.Values are given for the Watson−Crick duplex except for twists, which are given for both the duplex (1,2) and the Hoogsteen strand (1,3).

results in some backbone torsions adopting two families of conformations that are determined mainly by the force field, which shows that there was adequate conformational sampling in these calculations. However, because these features are not experimentally determined, we do not consider them further. The DNA purine strands are defined to a similar degree in each structure, and the rmsd values for each are: 0.51 ± 0.18 Å (DD·D), 0.32 ± 0.2 Å (DD·R), and 0.49 ± 0.178 Å (DD·R′). The rmsd between the purine strands of the three structures were 0.96 ± 0.1 Å for DD·R′/DD·D, 0.86 ± 0.08 Å DD·R/DD·D and 0.75 ± 0.12 Å for DD·R′/DD·R. These rmsd values reflect the differences in A3 and A4 of α and γ (see above), and demonstrate that the conformations of the chemically identical purine strands in the three molecules are significantly different from one another. As previously found, the glycosyl torsion angles are near −130° except for the nucleotides for which γ is trans, when $\chi \approx -140$ to −150°. For these nucleotides, there is a compensating change in α, from near −70° (g^-) to +155° (t). As the backbone angles were not constrained, these reflect stereochemical requirements for the g^+ to t transition for γ. The influence of this transition on the course of backbone can be seen in Figure 5B. As we have previously shown, the variation in the glycosyl torsion angles and the torsion γ in the purine strands of the three molecules leads to differences in the position of the phosphate backbone, and therefore in the widths of the grooves formed between the Watson−Crick strands and the Hoogsteen strand.[14]

Globally the three structures are similar in several key parameters. We have calculated helicoid parameters using the program CURVES[28] as shown in Table 3. In all three structures, the displacement of the helix axis (Dx) (Watson−Crick strands) is between −3 and −3.7 Å (i.e., into the minor groove) which lies between that found in B-DNA duplexes (0 to −1 Å) and A-RNA duplex (−4 to −5 Å), and is typical of parallel triplexes [9−14]. The inclination of the base pairs to the helical axis was found to be small (<5°), and the rise along the helix axis was 3.1−3.2 Å for all three structures. The main difference in helicoid parameters was the twist angle. For the Watson−Crick duplex, the twist angle was on average 33−35° for all three structures (i.e., 10.3−10.9 base-pairs per turn), but for strands 1,3, the twist angles differ significantly, with a twist of 35° for DD·D, 39.7° for DD·R and 41.6° for DD·R′. Thus, for DD·D, the helicoid parameters are most close to the B-form on average, except for the x-displacement, whereas for both DD·R and DD·R′, the twist of the third strand is substantially larger. It is notable that the helical rise in these duplexes is significantly lower than in a B-DNA duplex in solution, but larger than that found in an A-RNA duplex. For the third strand to remain in register with the DNA duplex, the twist angles and helical rise must match. This seems to be achieved by a combination of a smaller rise in the DNA, and an increased rise and decreased base inclination in the RNA strands, compared with the standard A conformation. The rmsd values calculated for the duplexes in the three triplexes compared with the free duplex are all between 1.5 and 1.6 Å.

(28) Lavery, R.; Sklenar, H. *Curves 5.1*, Manual; CNRS: Paris, 1996.

This is largely a result of the large displacement of the helix axis in the triplexes compared with the free duplex (cf. Table 3).

In DD·R′, the orientation of the O2′-methyl group is quite well determined, and makes numerous contacts, including H2′, H2″, and H3′, to the m-1 purine residue in the minor groove of the Watson−Crick duplex.

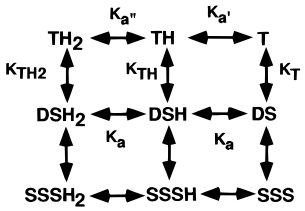
Discussion

The thermodynamic properties of the three triple helices are quite distinct from one another, which is reflected in their different conformations. The most detailed information about conformation is available for the purine strand, as a consequence of better spectral dispersion (and therefore more experimental restraints) than for the pyrimidine strands, but also because it interacts with two strands. The Hoogsteen strand makes direct contacts with the purine strand, but not with the Watson−Crick pyrimidine strand, so that changes in conformation in the purine strand can reflect direct effects of the formation of the triple helix. The most extensive conformational changes from the duplex states occur in DD·D where both A3 and A4 undergo a transition from $\gamma(g^+)$ to γ (t), and a change in the glycosyl torsion angle from the usual B-DNA value of −100 to −120° to −140 to −150°. This is achieved with no significant change in the sugar pucker. However, it is associated with a substantial variation in the position of the Hoogsteen strand and the width of the two grooves formed between the Hoogsteen strand and the two Watson−Crick strands.[14,24] A less extensive change in the conformation of the purine strand was observed in DD·R, where only A3 showed a significant amount of $\gamma(t)$. Furthermore, the chemical shifts of the duplex DNA strands were very similar in DD·D and DD·R, indicating rather similar overall conformations. In DD·R′, however, the amount of the $\gamma(t)$ conformer in the purine strands was small. Furthermore, the O2′-Me groups in DD·R′ provide additional cross-strand constraints, as they seem to populate predominantly one rotamer.

Significant changes in the conformation in the purine strand occur on forming the triplex in addition to the general changes in helicoid parameters common to all three-third strands (Table 3). This should cost some of the intrinsic binding energy. The conformational changes are in the order DD·D > DD·R > DD·R′ which is consistent with ranking of their thermodynamic stability. There are also differences in the conformations of the Hoogsteen strands; in DD·D, the third strand retains substantial C2′-endo character, whereas in both DD·R and DD·R′ the Hoogsteen strands are locked into the C3′-endo conformation. As the latter two molecules are more stable than DD·D, it is plausible that the optimum conformation of a triplex is in fact with C3′-endo sugars in the third strand, but for DNA this costs more energy to convert from mainly S to mainly N-type sugars. On this argument, DD·D should be the least stable triplex, as is indeed observed.

The observed thermodynamic stability of the triplexes seems to correlate with the degree of conformational differences between the triplex and duplex states. It is plausible that these conformational differences contribute substantially to the observed enthalpy changes of dissociation of the triplexes. The most stable triplex, DD·R′, is formed with minimal changes in conformation of the DNA purine strand, so that binding energy is not used to drive changes in conformation to optimize contacts. The O2′-methyl groups make additional van der Waals interactions that are not possible in the DNA triplex, and they inhibit the g^+/t transition in the purines which compresses the minor groove of the duplex. In contrast, the formation of the least stable triplex (DD·D) requires the most widespread conformational changes, both in the purine strand and in the sugars of the Hoogsteen strand, which suggests that some of the binding energy is lost in conformational rearrangement. DD·R is intermediate in both stability and the degree of conformational changes. This parallels the order of the enthalpy changes, although it is probable that there are effects other than different conformations. For example, the RNA hydroxyl protons are often thought to contribute to stability via their hydrogen bonding capacity.[26] The greater stability of the C2′-OMe in this case suggests that this contribution may be marginal. Indeed, the C2′-OMe groups make numerous van der Waals contacts in the major−minor groove and presumably affect the hydration on this groove compared with an RNA third strand.

Scheme 1



Experimental Section

Materials. d(AGAAGA-x-TCTTCT) (DD) and d(AGAAGA-x-TCTTCT-x-TCTTCT) (DD·D) were synthesized and purified as previously described.[12] x is the linker O−$(CH_2)_8$−OPO_2−O−$(CH_2)_8$−. Two analogues were prepared in a similar manner, using different phosphoramidites for the third (Hoogsteen) strand only, namely [d(AGAAGA-x-TCTTCT)-x-U′C′U′U′C′U] (DD·R′) where C′ and T′ are the nucleotides modified with C2′-OMe and [d(AGAAGA-x-TCTTCT)-x-r(UCUUCU)] (DD·R).

Methods. Thermodynamic Measurements. UV melting experiments were carried out and analyzed as previously described.[25] The buffer used for the melting studies was 10 mM sodium acetate, 10 mM sodium phosphate, 150 mM NaCl and 1 mM EDTA. This provides sufficient buffering power over the pH range 4.9−8, while maintaining the sodium concentration at 150 mM and the ionic strength between 176 and 198 mM. All UV melting curves were measured in triplicate.

Thermodynamic quantities were obtained from the melting curves at different pH values according to Scheme 1.

TH_2 represents the doubly protonated triplex, TH the two possible singly protonated triplexes, and T the fully deprotonated triplex. DSH_2, DSH, and DS represent the duplex+strand states, and $SSSH_2$, SSSH, and SSS represent the fully dissociated (all strand) state. The absorbance at 260 nm depends on the T_m values of the triplex−duplex and duplex-strand transitions, their van't Hoff enthalpies, and the specific absorbance of the triplex, duplex and strand states according to eq 1, as previously described.[25]

$$A(T) = (A_T + K_{tapp}A_D + K_{tapp}K_{dapp}A_S)/(1+K_{tapp}+K_{tapp}K_{dapp}) \quad (1a)$$

where $K_{tapp,}$ K_{dapp} are apparent dissociation constants. As the duplex stability is independent of pH in the range 5−9, K_{dapp} is also independent of pH. K_{tapp} is the apparent dissociation constant of the triplex states into the duplex plus strand states. The dissociation constants are temperature-dependent according to:

$$K_{t,d} = \exp[(\Delta H_{t,d}/R)(1/T_m - 1/T)] \quad (1b)$$

where ΔH is the apparent van't Hoff enthalpy change, R is the gas constant, and T_m is the melting temperature. From simulations and triplicate measurements, we estimate the precision of the fitted parameters as ∼±0.5 K for the T_m values, and ±10 kJ mol^{-1} for the van't Hoff enthalpies.

The enthalpy of dissociation of the protonated triplex to the duplex plus strand state can also be calculated from the dependence of T_m on

pH. Assuming that the CytN3H$^+$ ionize independently, then the apparent dissociation constant for triplex to duplex + strand states as given in Scheme 1 is:

$$K_{app} = K_{TH2}(h^2 + hK_a + K_a^2)/(h^2 + hK_a' + K_a'^2) \quad (2)$$

where $h = 10^{-pH}$. The two pK_a' values could be different, as is probably the case for adjacent CG.C+ triplets[16] However, for the molecules studied here, the separation between the positive charges is at least 10 Å. We have previously shown for similar sequences that these charges are delocalized and that the degree of interaction between the protonated cytosines must be small.[24,25]

As the value of the pK of CN3H (pK_a') in the triplex state is high[25] $h \gg K_a'$ up to pH 8. Furthermore, for an intramolecular triple helix the value of the apparent dissociation constant is unity at $T = T_m$. Hence, at $T = T_m$ eq 2 simplifies to

$$K_{TH2} = h^2/(h^2 + hK_a + K_a^2) \quad (3)$$

Assuming that p$K_a \approx 4.5$ at 298 K and $\Delta H_{ion} = 18$ kJ mol^{-1},[29] then the right-hand side of eq 3 can be evaluated at each pH, and the value of ΔH_{TH2} determined from a van't Hoff plot of ln(K_{TH2}) versus $1/T_m$.

Similarly, at pH values for which $K_a' \ll h \ll K_a$ (i.e., in the range 5.5 to 8)[25]

$$d(1/T_m)/dpH \approx 2.303R\Delta p/\Delta H \quad (4)$$

where Δp is the number of protons released on dissociation of the triplex.

ΔG(t) will also be pH dependent through eq 2. In the range p$K_a' \gg$ pH $\gg$ pK_a,

$$\Delta G = -RT \ln(K) \approx -RT\{\ln K_{TH2}K_a^2 + 2pH\ 2.303) \quad (5)$$

NMR Spectroscopy. NMR samples were prepared in 10 mM sodium phosphate, 100 mM KCl, pH 5.0. NMR spectra were recorded at 11.75 or 14.1 T on Varian UnityPlus and Unity spectrometers, respectively. Spectra recorded at 500 MHz, 30 °C: NOESY, DQF COSY, TOCSY, and ROESY, + 1D at T from 5 to 60 °C. Phase-sensitive[30] NMR spectra in H_2O were recorded at 5 °C using Watergate[31] for solvent suppression.

Modeling. Cross-peaks in NOESY and ROESY experiments were classified by visual inspection as strong (2−2.5 Å), medium (2−3.5 Å), and weak (3−5 Å) and used to create qualitative constraints. When a methyl group was involved the bounds were: Strong (2−2.5 Å), medium (2−4.5 Å), and weak (3.5−5.5 Å). Those proton pairs whose NOEs were absent in the 250 ms NOESY spectrum were constrained with a lower bound of 4.5 Å.

Sugar puckers were analyzed using sums of coupling constants derived from DQF-COSY as previously described.[24,32] The spectral overlap made it difficult, in many cases, to obtain a complete set of Σ values for the pyrimidine strands. Those residues in which $\Sigma_{1'} > 14.5$ Hz and r(H1′−H4′) > 2.8 Å were loosely constrained during the modeling to the S range (120° < P < 180°). All residues in the RNA strand were tightly constrained to C3′-endo conformation. Both NOESY and ROESY experiments indicate a low value for χ (i.e., ∼−160°) in the RNA strand, as expected for an A-like structure. These angles were constrained during the modeling to the range 180−220°.

Constraints on the torsion angle γ were obtained from the measurements of $\Sigma_{4'}$. The values obtained for the purine strand and for most residues in the third strand are consistent with a dominant rotamer in the g^+ region. In addition some experimental information about the backbone angle ϵ was obtained from the H3′ line widths. It is known that if the g^- rotamer is significantly populated, the large $^3J_{HP}$ (>20 Hz) make the $\Sigma_{3'}$ value much larger than 20 Hz.[27] In all cases $\Sigma_{3'}$ were in the 12−18 Hz range in the DNA strands and in the 23−24 Hz range in the RNA strand (the values are larger in the RNA due to the larger $J_{H3'-H4'}$ that correspond to a C3′-endo conformation) which rules out this possibility. These data were not used to create constraints in the structure refinement. Nevertheless, the final models were in perfect agreement with it.

The starting structure used for the modeling was the triplex structure of DD·D previously obtained by NMR.[24] In that structure, no constraints in the third strand sugar conformations were imposed. This gave S type sugars for all residues. However, the two protonated cytosine residues in the Hoogsteen strand showed a high fraction of the N state according to the coupling constants. We have calculated an alternative structure in which these two residues are constrained to N for further detailed comparison. Both of these structures should be significantly populated in solution.

Structures were refined and energy-minimized using Discover 95.0 (Molecular Simulations Inc., San Diego) with the Amber force field. No electrostatics were included, and no distance cutoffs were employed. Initial models were refined against the experimental data using a common protocol as follows. The structures were first energy-minimized with 500 cycles of conjugate gradients, followed by 1.5 ps of unrestrained molecular dynamics at 600 K to partially randomize the coordinates. The restraints were then applied, and energy-minimized (500 cycles of conjugate gradients), followed by 5.5 ps rMD at 600 K, 5 ps rMD at 450 K and 10.5 ps rMD at 300 K, and finally energy-minimized (2500 cycles conjugate gradients). The energies and structures were stored, along with the violations and values of the torsion angles.

For DD·D, a total of 328 nonredundant distances and 30 torsion restraints were used (19.9 per residue), 176 distances + 31 torsions for DD·R′ (11.5 per residue) and 165 distances + 50 torsions for DD·R (11.9 per residue). In addition, planarity restraints with a very low weight factor of 5 kcal mol^{-1} rad^{-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. The force constant for the hydrogen bond constraints was 25 kcal mol^{-1} A^{-2}, 50 kcal mol^{-1} A^{-2} for the NOE constraints and 60 kcal mol^{-1} rad^{-2} for the dihedral constraints.

Acknowledgment. This work was supported by the Medical Research Council of the UK and Oswel Research Products Ltd. J.L.A. gratefully acknowledges a Fellowship from the Spanish Ministry of Education. NMR spectra were recorded at the MRC Biomedical NMR Centre, Mill Hill. We thank Dr. J. O. Trent for valuable comments on the manuscript.

Supporting Information Available: Six tables (ASCII) of ^{1}H assignments, modeling parameters, sugar conformations, and coupling constants and torsion angles of the refined structures. Three figures of a ROESY spectrum of DD·R′, DQF-COSY spectra of DD·R and DD·R′, and ROESY spectra of DD·R and DD·R′ (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA991949S

(29) Fasman, G. D. *CRC Handbook of Biochemistry and Molecular Biology*, 3rd ed.; CRC Press: Boca Raton, FL, 1975; Vol. I, p 191.

(30) States, D. J.; Haberkorn, R. A.; Ruben, D. J. *J. Magn. Reson.* **1982**, *48*, 286−292.

(31) Piotto, M.; Saudek, V.; Sklenar, V. *J. Biomol. Struct.* **1992**, *2*, 661−665.

(32) Conte, M. R.; Bauer, C. J.; Lane, A. N. *J. Biomol. NMR.* **1996**, *7*, 190−206.