Solution Structure of an Intramolecular DNA Triplex Containing 5-(1-Propynyl)-2′-deoxyuridine Residues in the Third Strand[†,‡]

A. Kathryn Phipps, Markus Tarköy, Peter Schultze, and Juli Feigon*

Department of Chemistry and Biochemistry and Molecular Biology Institute, University of California, Los Angeles, California 90095-1569

Received November 17, 1997; Revised Manuscript Received January 30, 1998

ABSTRACT: Incorporation of the modified base 5-(1-propynl)-2′-deoxyuridine (propynylU) in the third strand of a triplex leads to enhanced triplex stabilization. To investigate effects of the propyne nucleotide on triplex structure and the factors underlying the increased stability, we have determined the solution structure of the intramolecular DNA pyrmidine-purine-pyrimdine d(AGAGAGAA-$(EG)_6$-TTCTCTCT-$(EG)_6$-PCPCPCPP) (PDD-EG), which contains 5-(1-propynl)-2′-deoxyuridine (P) in the third strand and hexakis(ethylene glycol) linkers [$(EG)_6$]. The structure was calculated using X-PLOR with distance and dihedral angle restraints obtained from two-dimensional NMR experiments and refined with the direct relaxation matrix method. The structures show that the extended aromatic electron cloud of the propynylU nucleotide stacks well over the 5′-neighboring nucleotides, resulting in increased stabilization. The propynylU nucleotides also affect the overall structure of the triple helix. A comparison of the structure to that of the nonmodified intramolecular DNA triplex of the same sequence, d(AGAGAGAA-$(EG)_6$-TTCTCTCT-$(EG)_6$-TCTCTCTT) (DDD-EG), shows that PDD-EG has a more A-DNA like X displacement and inclination than DDD-EG yet still maintains predominantly S-type sugar puckers as found in DDD-EG and other DNA triplexes.

Since the initial discovery of triple helix formation in 1957 (*1*), many advancements in the understanding of triplex structure and in the technology involving oligonucleotide delivery into the cell have been made, allowing for the plausible use of antigene therapeutics. This type of "antigene strategy" would have an advantage over similar "antisense strategies" by being able to target one gene with a small amount of oligonucleotide, rather than a large, continuous supply for the production of multiple copies of the mRNA gene transcript (*2*−*5*). Both of these strategies require sequence specific oligonucleotide strands that are stable enough to form under cellular conditions and are able to reach the target sequences before being degraded. For parallel motif triplexes (YRY) (*6*), one of the two types of triplexes that form with the addition of a third strand to homopyrimidine-homopurine tracts, many factors that help to stabilize triplex formation have been found. These triplexes, which are generally pH dependent due to the requirement of protonated cytosines (C^+s) in the homopyrimidine third strand, can be stabilized by reduction of interstrand repulsion with addition of cations, with low pH, and by hydrophobic substituents in the third strand, such as 5-methyldeoxycytidine (m^5C) (*7*, *8*). Recent studies have taken advantage of new advancements in enhancing triplex formation by using modified oligonucleotides, which increase the stability as well as allow for greater possibilities in sequence recognition, triplex-stabilizing molecules, such as the intercalator coralyne (*9*), and the use of conjugated molecules for cross-linking (*10*). Triplexes have been used for probing both single- and double-stranded oligonucleotides (*11*, *12*) and for sequence specific cleavage (*13*, *14*). Other current triplex applications using the antigene approach include the suppression of gene expression for the insulin-like growth factor type I receptor in rat C6 glioblastoma cells (*15*) and the detection of triplex formation through the use of cross-linking, in both human cells and cells infected with HIV (*16*, *17*).

Stabilization of potential therapeutic oligonucleotides can be achieved by the incorporation of the modified base 5-(1-propynyl)-2′-deoxyuridine (propynylU), which stabilizes both triplex and duplex formation and also enhances antisense potency (*18*, *19*). This modification is one of many C5-substituted uracil nucleosides that were synthesized as nucleoside analogue drugs and tested for possible antiviral and antitumor activities more than 10 years ago (*20*, *21*). More recent studies by Froehler et al. (*22*) have demonstrated that propynylU as well as m^5C substitutions in the third (target) strand stabilize triple helix formation compared to T and C, respectively, as well as duplex formation with single-strand RNA. The enhanced stabilization from the incorporation of the modified base propynylU has been proposed to be due to entropic stabilization and improved base stacking (*22*). Recent antisense research has used oligonucleotides containing the C5 propyne substitutions, as well as a phosphorothioate backbone, to resist nuclease attack and recruit and/or activate RNAse H cleavage, to inhibit gene expression by its hybridization to mRNA and subsequent prevention of translation (*23*, *24*). The C5 propyne-

[†] This work was supported by NIH Grant GM 37254 to J.F.

[‡] Coordinates for these structures have been deposited in the Brookhaven Protein Data Base (entry 1P3X).

* To whom correspondence should be addressed.

S0006-2960(97)02811-0 CCC: $15.00
Published on Web 04/07/1998

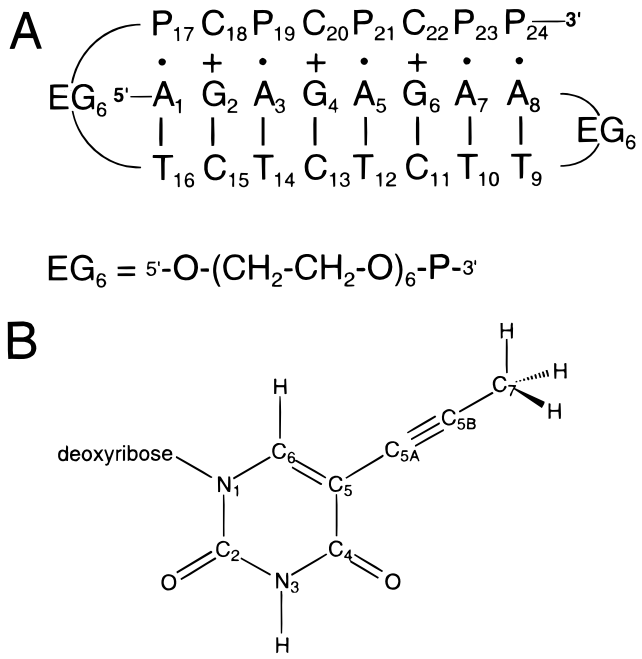

FIGURE 1: (A) Sequence and base pairing scheme of the intramolecular triplex with propynyl substituents in the third strand and ethylene glycol loops [$(EG)_6$ = 5′-O-$(CH_2CH_2O)_6$-P-3′]. Watson−Crick hydrogen bonds are represented by |, and Hoogsteen hydrogen bonds are represented by • or by + for the protonated C. (B) Chemical structure of 5-(1-propynyl)-2′-deoxyuridine.

substituted pyrimidines combined with other 2′ modifications and phosphorothioate backbones create the most potent antisense inhibitors. Furthermore, it has been shown that they can be delivered efficiently via cationic lipids and, therefore, have a possible use as human therapeutics (*25*). The C5 propynyl pyrimidine phosphorothioate oligonucleotides (propyne-S-ONs) also have antiviral activities and have been shown to disrupt the Rev−RRE complex (*26*). Results of antisense studies with propynyl substitutions have shown the propyne-S-ONs to be the most potent and gene specific antisense agents (*27*) even when compared to C5-substituted analogues that are thermally slightly more stable (*24*). Recent triplex work has taken advantage of using the enhanced triplex-forming ability of 5-(1-propynyl)-2′-deoxyuridine in targeting duplex DNA with an oligonucleotide made up entirely of propynylUs and m^5Cs tethered to a molecule which can recruit topoisomerase I to the duplex DNA (*28*).

To investigate more fully the unique properties of oligonucleotides containing the propynyl substitutions, we have determined the solution structure of the intramolecular DNA pyrimidine-purine-pyrimidine triplex d(AGAGAGAA-$(EG)_6$-TTCTCTCT-$(EG)_6$-PCPCPCPP) (PDD-EG), which contains 5-(1-propynyl)-2′-deoxyuridine (P) in the third strand and hexakis(ethylene glycol) linkers [$(EG)_6$] (Figure 1A). The 1-propynyl substituent contains an extra single and triple bond (two additional carbon atoms, C5A and C5B), which linearly extends the methyl group from the rest of the base as compared to that of the thymidine (Figure 1B). The refined structure is presented and compared to the nonmodified intramolecular DNA triplex of the same sequence, d(AGAGAGAA-$(EG)_6$-TTCTCTCT-$(EG)_6$-TCTCTCTT) (DDD-EG) (reported in ref *29*). These two triplexes differ by only five nucleotides, where the thymidines in the third strand have been replaced by 5-(1-propynyl)-2′-deoxyuridines.

MATERIALS AND METHODS

Synthesis of 5-(1-Propynyl)-2′-deoxyuridine Phosphoramidite. The sugar hydroxyl groups of 5-iodo-2′-deoxyuridine were protected by treatment with *p*-toluyl chloride as described (*30*). The protected 5-iodo nucleoside was coupled with propyne by Pd catalysis in a triethylamine solution and purified as described (*20*). Subsequent tritylation and phosphitylation steps were carried out following standard protocols (*31*).

Sample Preparation. Synthesis of the oligonucleotide d(AGAGAGAA-$(EG)_6$-TTCTCTCT-$(EG)_6$-PCPCPCPP) was carried out on an ABI 392 DNA synthesizer using standard phosphoramidite chemistry and protocols, except for the coupling time for the hexakis(ethylene glycol) linkers which was extended to 10 min. The natural deoxyribonucleotide phosphoramidites were obtained from Pharmacia, and the hexakis(ethylene glycol) phosphoramidite (Spacer 18) was obtained from Glen Research. The oligonucleotide was deprotected with concentrated aqueous ammonia, incubated at 55 °C for 14 h, and then evaporated to dryness in vacuo. The sample was purified by FPLC using a strong anion exchange column (Mono Q, Pharmacia) with 10 mM NaOH and 10 mM NaOH/1 M NaCl (30−100%) solutions. Peak fractions were combined and adjusted to pH 7. To remove the remaining salts, the product was then precipitated with ethanol and chromatographed on a Sephadex G15 desalting column. The fractions containing the oligonucleotide were combined and lyophilized to dryness. The NMR sample was prepared by dissolving the oligonucleotide in 90% H_2O/10% D_2O (H_2O experiments) or 99.999% D_2O (D_2O experiments) in 100 mM NaCl (pH 5.2) and 5 mM $MgCl_2$. The sample was exchanged between H_2O and D_2O by repeated drying of the sample in the NMR tube under a stream of N_2(g). The final oligonucleotide concentration of the NMR sample was 1.7 mM in a volume of 200 μL in a Shigemi NMR tube.

NMR Spectroscopy. The NMR experiments were performed on Bruker DRX500 and AMX500 spectrometers. A NOESY spectrum (*32*) of the sample in H_2O was acquired at 274 K with a τ_m of 150 ms and 64 acquisitions per t_1 value using a $1\bar{1}$ spin echo read pulse sequence (*33*) to suppress the water signal. A NOESY spectrum in D_2O was acquired with a mixing time of 250 ms and 64 scans at 298 K. Water suppression of the residual water peak in D_2O spectra was accomplished by presaturation during the recycle delay. All NOESY results were obtained in the States−TPPI mode (*34*). A DQF-COSY spectrum (*35*), a TOCSY spectrum (*36*) with a 100 ms MLEV17 mixing pulse, and a heteronuclear natural abundance $^1H-^{13}C$ HSQC spectrum (*37*) were also acquired at 298 K. In addition to the above spectra used for assignments and initial distance and dihedral angle restraints, a series of NOESY spectra were acquired for the relaxation matrix refinement with mixing times of 40, 80, 140, and 200 ms and 48 acquisitions per t_1 value. All spectra were collected with 2048 points in t_2. For the H_2O NOESY, 250 ms NOESY, DQF-COSY, HSQC, and mixing time NOESY experiments, 524, 555, 900, 400, and 512 points, respectively, were collected in t_1. The recycle

delays for the H_2O NOESY and the 250 ms NOESY experiments were 1.8 s and for the other D_2O NOESY experiments were 2.0 s. The data were processed with XWINNMR (Bruker Instruments, Inc.).

UV Melting Spectroscopy. Optical melting studies of PDD-EG and DDD-EG were performed on a Varian Cary 1E spectrophotometer equipped with a temperature probe. The heating rate was 0.5 °C/min, and data were acquired between 5 and 90 °C. The concentration of oligonucleotide was adjusted to give an absorbance in the range of around $0.2-0.5OD_{260}$ in a volume of 1 mL. Buffer conditions were 50 mM NaOAc, 0.1 M NaCl, and 5 mM $MgCl_2$ (pH 5.20). The T_m was calculated from the maximum of the first derivative of the melting curve.

Torsion Angle Restraints. The coupling constants in deoxyriboses 2−24 were determined from the simulated spectra of the H1′−H2′ and H1′−H2″ cross-peaks of a DQF-COSY spectrum using the program CHEOPS (*38*) as described (*29*). The endocyclic sugar torsion angles ν_1 and ν_2 were then determined from these coupling constants using the program PSEUROT (*39*, *40*) which fits them to a two-state model for the sugar pucker (Supporting Information Table 1). The torsion angles from the major conformation were used in the structure calculations, except in the cases of the C^+ sugars. In these cases, although the PSEUROT analysis gave approximately a 50:50 ratio of the N and S types, additional information from NOE intensities and carbon chemical shifts was consistent with S-type conformations. Therefore, these sugars were restrained with the calculated S-type torsion angles.

Structure Refinement. The distance restraints between nonexchangeable protons used for the initial structure calculation were acquired from the 250 ms NOESY spectrum. The AURELIA software package (*41*) was used to pick and integrate cross-peaks. The NOESY peak lists were generated using a semiautomatic assignment procedure which combined the information from the picked peaks with a chemical shift list as described (*29*). Interproton distances obtained from the D_2O NOESY spectra and upper and lower bounds were determined in the same way as described in the previous paper (*29*) with the exception of cross-peaks that had to be manually added to the peak list because they failed to be accurately picked and integrated by the AURELIA integration algorithm. This problem usually arose due to cross-peak overlap, and these cross-peaks were therefore given an upper bound of 6.5 Å. Distances for the exchangeable protons were given a lower bound of 2 Å and an upper bound of 5.5 Å as described previously (*29*).

Thirty refined structures were calculated by distance geometry and simulated annealing using X-PLOR Version 3.1 (*42*), followed by full relaxation matrix refinement of the 10 lowest-energy structures utilizing the same protocols that were described in the preceding paper (*29*). Helical parameters were calculated using the Watson−Crick duplex portion of the triplex structures with CURVES 5.1 (*43*) and were averaged over the 10 structures. Averaged parameters for the base pairs are reported as calculated for both a nonlinear axis and a linear axis.

RESULTS

Exchangeable Proton Resonances and Assignments. The two-dimensional (2D) NOESY spectrum of PDD-EG in H_2O at 1 °C is similar to spectra obtained for DDD-EG (*29*) and other YRY intramolecular triplexes (*6*, *44*). Sequential connectivities can be followed through the Hoogsteen paired third strand from C18 to P24 (Figure 2A). Unfortunately, C20 and C22 are overlapped in the imino region along with T9. The imino for T16 was assigned from the cross-peaks of its methyl protons, and P17 could not be assigned due to overlap. One cross-peak between T9 and T10 is observed in this region. Interstrand NOEs between the Watson−Crick paired strands and weak NOEs between the Hoogsteen paired strands are also present in the spectrum (Figure 2A) and help to confirm the folding of the triplex. In the imino−amino/aromatic region of the spectrum (Figure 2B), the typical downfield shift for the amino proton resonances of the third strand C^+ residues is seen (*44*).

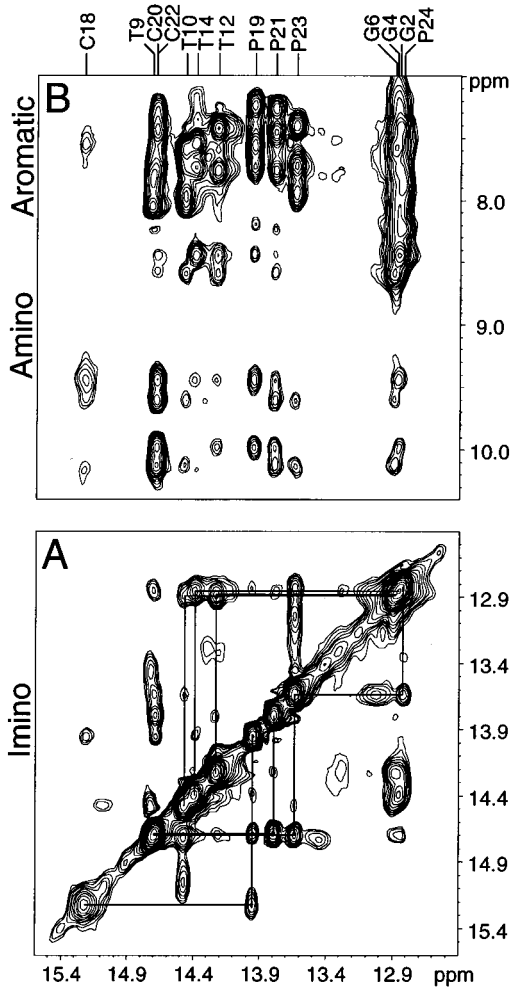

FIGURE 2: Portions of a NOESY spectrum of PDD-EG in 90% H_2O/10% D_2O at 1 °C and a mixing time of 150 ms showing (A) the imino−imino and (B) imino−amino/aromatic cross-peaks. In panel A, sequential connectivities for C18−P24 of the Hoogsteen paired third strand are shown by the lines below the diagonal. Cross-peaks between the Watson−Crick strands (T10−G6, T12−G4, and T14−G2) are shown by the lines above the diagonal. Weak cross-peaks between the Hoogsteen strands (T10−P23, T12−P21, and T14−P19) are marked with an asterisk above the diagonal. Assignments of the imino resonances are given at the top of panel B.

Sequential Assignment of Nonexchangeable Protons. Sequential assignments of triplex proton resonances were made following previously published protocols (*45*, *46*) using information from the NOESY, DQF-COSY, and TOCSY

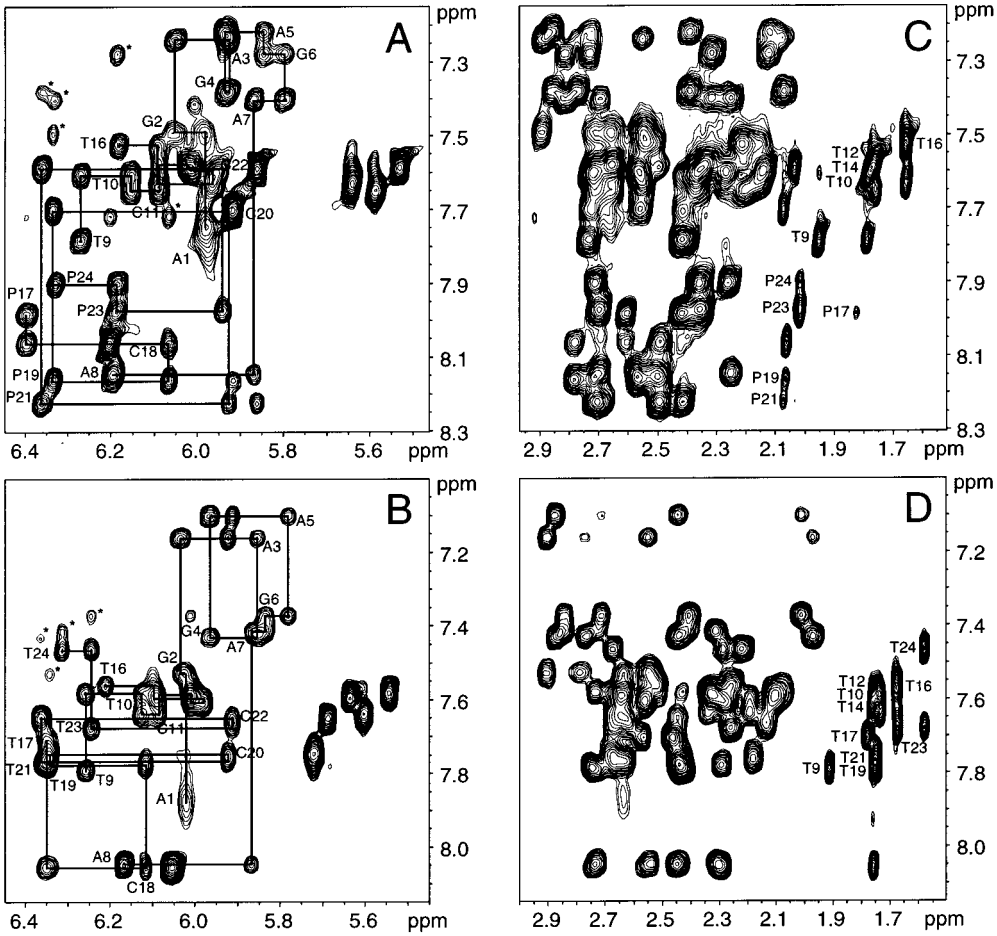

FIGURE 3: Portions of NOESY spectra in D_2O at 25 °C with a mixing time of 200 ms showing the aromatic−H1′/H5 cross-peaks for (A) PDD-EG and (B) DDD-EG and the aromatic−H2′/H2″/Me cross-peaks for (C) PDD-EG and (D) DDD-EG. In panels A and B, sequential connectivities are indicated by the lines for each strand. Cross-peaks denoted with an asterisk are interstrand NOEs between the Hoogsteen strand H1′ and the purine strand H8 protons in the 5′-neighboring triplet, e.g. P24−A7 in panel A. The peaks with the highest intensity are the H6−H5 cross-peaks from the cytosines. In panels C and D, the methyl cross-peaks for both the thymidines and 5-(1-propynyl)-2′-deoxyuridines are labeled.

spectra as well as a natural abundance $^1H-^{13}C$ HSQC spectrum. Sequential connectivities of the nucleotide strands of PDD-EG are shown for the aromatic−H1′ region of a D_2O NOESY spectrum at 25 °C in Figure 3A. The same region of DDD-EG is shown in Figure 3B for comparison. For PDD-EG, the first (purine) and third (Hoogsteen paired pyrimidine) strand connectivities can be easily traced because of good chemical shift dispersion for both the aromatic and H1′ hydrogens. The second (Watson−Crick paired pyrimidine) strand, however, has a fairly narrow range of chemical shifts, and assignments required the additional information from the H1′−H2′/H2″ region. Interstrand NOEs between H1′ protons of the Hoogsteen paired pyrimidine strand and the purine strand's H8 protons in the 5′-neighboring triplet are also observed. Interstrand peaks are important in determining the structure since they provide distance restraints between the Hoogsteen paired (first and third) strands. Interstrand contacts between the Watson−Crick paired strands are also present, such as the AH2 protons of the first strand to the H1′ protons of the second strand in the 5′-neighboring triplet, e.g. C13−A5. Sequential connectivities were also observed in the aromatic−H3′ and aromatic−H2′/H2″ regions. A comparison of the aromatic−H2′/H2″/Me region for PDD-EG (Figure 3C) and DDD-EG (Figure 3D) clearly shows the increased chemical shift dispersion gained through the propynyl substitution.

A portion of the $^1H-^{13}C$ HSQC spectrum showing the H1′, H3′, H4′, H5′/H5″, and ethylene glycol protons of PDD-EG is shown in Figure 4. Terminal residues of PDD-EG are displaced from the rest of the residues in their ^{13}C shift due to the terminal OH. The ethylene glycols that are attached to the oligonucleotide strands have chemical shifts close to H5′/H5″. The ^{13}C chemical shifts are nearly identical between corresponding residues in PDD-EG and DDD-EG (not shown). Resonance assignments for all of the base and sugar protons are listed in Table 1.

Ethylene Glycol Linker Resonances. As discussed for DDD-EG (*29*), NOEs are observed between residues at the ends of the nucleotide strands and some ethylene glycol resonances. Some of the methyl groups from both the propynylU and thymidine (T9, T16, and P17) have a set of cross-peaks to various protons in the ethylene glycol linker. There are also strong NOE cross-peaks to the H6 of the 5′-terminal residues (T9 and P17), as well as weak cross-peaks to the 3′-terminal H1′ protons (A8 and T16). We were unable to specifically assign the ethylene glycol residues, and therefore, the ethylene glycol loops were not included in the structure calculations.

Deoxyribose Conformation. Pseudorotation phase angles were calculated for nucleotides 2−24 on the basis of coupling constants obtained from simulation of the DQF-COSY at 25 °C. All of the simulated cross-peaks exhibited reasonable

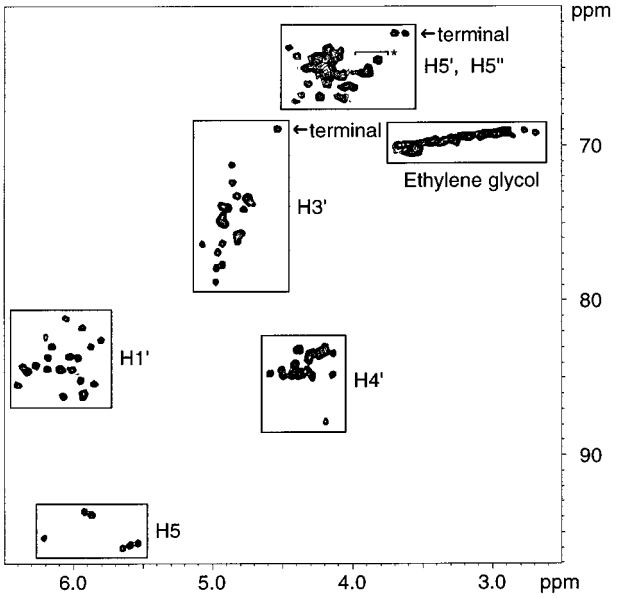

FIGURE 4: Portion of a $^{1}H-^{13}C$ HSQC spectrum showing the H1′, H3′, H4′, H5′, H5″, and ethylene glycol proton resonances of PDD-EG. Resonances from terminal residues (A1H5′/H5″ and P24H3′) of the oligonucleotide are indicated by the arrows. The asterisk denotes the ethylene glycols that are attached to the oligonucleotide strands.

correlation coefficients in CHEOPS, ranging from 91 to 96% except for two (A7 and P19) which were 88%. On the basis of this analysis, all of the sugars except for the C^+ residues adopt a predominantly S-type (C2′-endo) pucker at equilibrium, according to the two-state model used by PSEUROT. The C^+ residues were found to have about equal populations of S-type and N-type sugar conformations (discussed below).

The pseudorotation phase angles in the final calculated structures were all between 103 and 181° (values derived from CURVES) for restrained nucleotides 2−24. Nucleotide 1, which was left unrestrained due to very weak COSY peaks and spectral overlap, had an average phase angle of 203.6 ± 7.0° from the 10 structures. These pseudorotation angles are in good agreement with those obtained for other intramolecular DNA triplexes, which also have mostly S-type sugars (*6*, *29*, *38*, *47*, *48*).

Thermal Stability. The relative stability of PDD-EG and DDD-EG was determined by monitoring the change of absorbance at 260 nm versus temperature. PDD-EG had a T_m of 77.5 °C which was higher than DDD-EG's T_m of 67.0 °C (data not shown). This stabilization of 10.5 °C from the propynylU substitutions (or 2.1 °C per substitution) is similar to the stabilization observed in the initial studies done by Froehler et al. (*22*) (2.4 °C per substitution under different conditions), using a bimolecular triplex with a different sequence.

Structure Calculations. The 10 lowest-energy structures of PDD-EG after full relaxation refinement are presented as a stereopair in Figure 5. The three oligodeoxynucleotide strands form a well-defined, right-handed triple helix. The average pairwise rmsd for 10 structures is 1.03 ± 0.13 Å for all heavy atoms of the triplex. Structure statistics for the refinement are summarized in Table 2. The 542 nonexchangeable restraints do not include distances of conformationally immobile protons, such as CH5−CH6, and the 138 exchangeable restraints do not include any intrabase distances. There are 36 restraints that involve the propyne methyls, 30 of which are internucleotide. The helical parameters determined by Curves 5.1 are plotted in Figure 6, along with the average values of standard A- and B-DNA.

DISCUSSION

Comparison of Spectra of PDD-EG and DDD-EG. Both PDD-EG and DDD-EG share analogous characteristics in their spectra, which are also common to other triplexes with deoxyribose sugars (*6*). NMR spectroscopic and thermodynamic data provide evidence for the folding of the oligodeoxynucleotide into a triplex as designed. The observed pattern and intensities of the NOE cross-peaks, e.g.

Table 1: Proton Chemical Shifts of PDD-EG (Parts per Million) at 25 °C

residue	H6, H8	H2, H5, Me	H1′	H2′	H2″	H3′	H4′	imino[a]	amino[a]
A1	7.71		5.98	2.67	2.71	4.98	4.49		na
G2	7.49		6.05	2.55	2.89	5.07	4.32	12.85	na
A3	7.23	7.54	5.92	2.08	2.77	4.82	4.59		7.73, 7.56
G4	7.38		5.93	2.39	2.86	4.92	4.35	12.87	na
A5	7.21	7.41	5.84	2.12	2.73	4.80	4.50		7.43, 7.35
G6	7.27		5.80	2.32	2.82	4.81	4.38	12.88	na
A7	7.40	7.61	5.87	2.25	2.70	4.89	4.31		7.94, 7.72
A8	8.14	8.04	6.20	2.51	2.57	4.91	4.38		7.78,7.81
T9	7.78	1.95	6.27	2.42	2.74	4.93	4.28	14.71	
T10	7.60	1.79	6.15	2.37	2.70	4.94	4.24	14.45	
C11	7.64	5.58	6.09	2.21	2.66	4.78	4.29		7.58, 7.61[b]
T12	7.57	1.76	6.02	2.33	2.65	4.93	4.23	14.23	
C13	7.58	5.53	6.01	2.14	2.56	4.75	4.19		8.44, 7.34[b]
T14	7.59	1.78	5.97	2.27	2.55	4.89	4.19	14.37	
C15	7.63	5.64	6.09	2.18	2.54	4.72	4.20		8.43, 7.20[b]
T16	7.52	1.66	6.18	2.22	2.51	4.82	4.13	12.86	
P17	7.98	1.83	6.40	2.42	2.61	4.97	4.49	na[c]	
C18	8.06	6.20	6.07	2.50	2.79	4.75	4.40	15.20	10.17, 9.44[b]
P19	8.16	2.07	6.33	2.56	2.73	4.93	4.41	13.94	
C20	7.70	5.92	5.93	2.41	2.73	4.86	4.38	14.68	9.97, 9.43[b]
P21	8.22	2.08	6.36	2.50	2.71	4.96	4.43	13.78	
C22	7.59	5.86	5.94	2.35	2.71	4.85	4.38	14.67	10.11, 9.60[b]
P23	7.97	2.04	6.19	2.36	2.72	4.93	4.30	13.63	
P24	7.90	2.02	6.33	2.27	2.38	4.53	4.14	12.82	

[a] Chemical shifts at 1 °C. [b] Lower field resonance is the hydrogen-bonded amino proton N4(1). [c] na, not assigned.

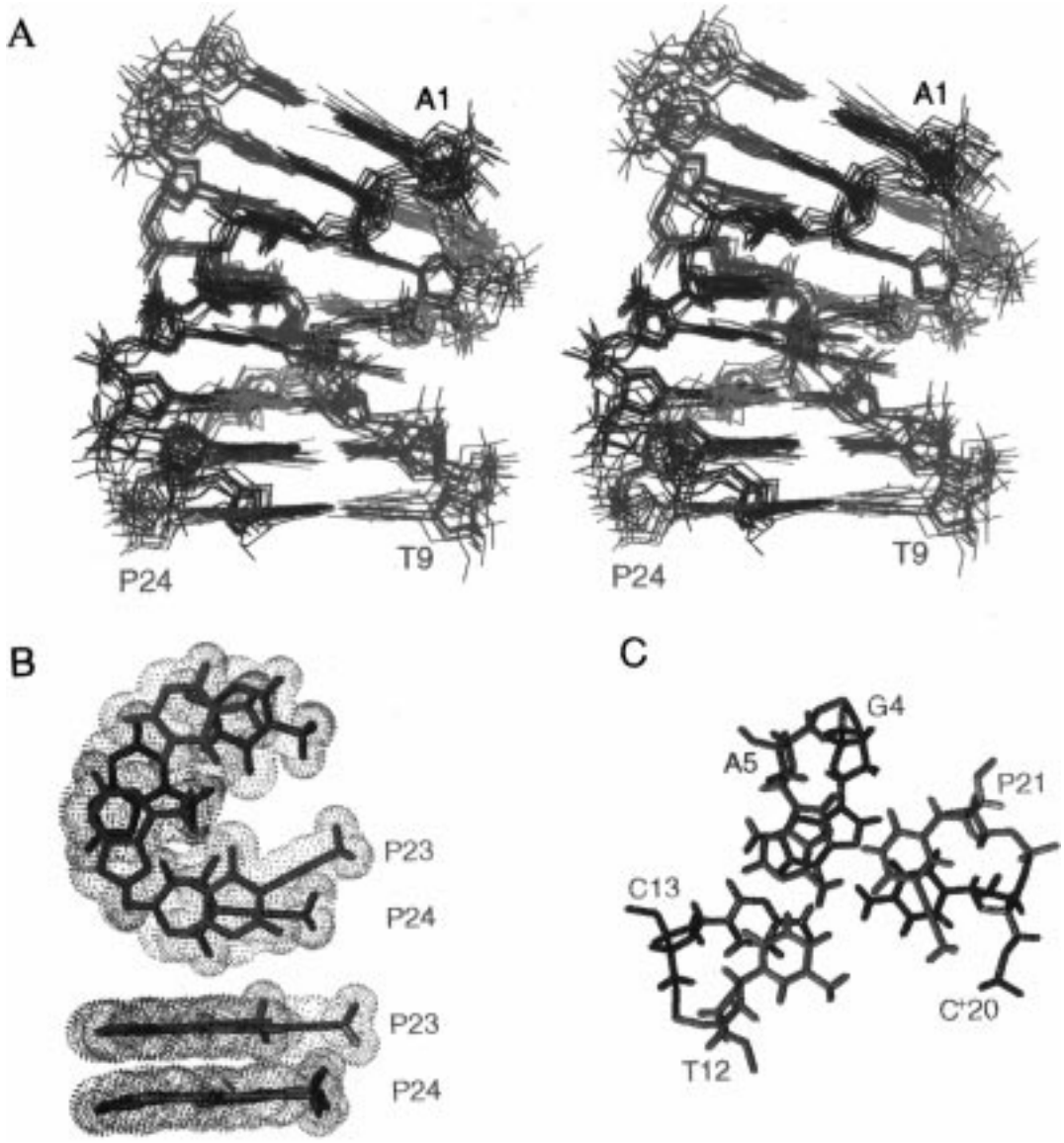

FIGURE 5: (A) Stereoview of a superposition of all heavy atoms of the 10 lowest-energy structures of PDD-EG. The view is into the major groove. The Watson−Crick paired pyrimidine strand is green (T9−T16) and the purine strand blue (A1−A8); the Hoogsteen paired cytidines (C18, C20, and C22) are cyan and the Hoogsteen paired propynylU nucleotides (P17, P19, P21, P23, and P24) violet. Only the nucleotides are shown, since the ethylene glycol linkers were not included in the structure calculation. (B) Top and side view of the two triplets (P23·A7·T10 and P24·A8·T9) containing the consecutive propynylUs. The van der Waals radii for the bases are shown as a dotted surface. (C) Stacking of the 5-(1-propynyl)-2′-deoxyuridine-containing triplet (P21·A5·T12) on the neighboring triplet (C^+20·G4·C13).

H2″−H6/H8 > H2′−H6/H8, are characteristic of a B-DNA helix. Intrastrand NOE cross-peaks such as third strand H1′−purine strand H8 in the 5′-neighboring triplet (e.g. P24−A7 in Figure 3A) are indicative of an intramolecular DNA triplex structure. In the D_2O NOESY spectra, PDD-EG has a pattern of sequential base sugar connectivities similar to that of DDD-EG (*29*), but there is more chemical shift dispersion in the resonances from the third strand which contains 5-(1-propynyl)-2′-deoxyuridines instead of thymidines (Figure 3). The NOE cross-peak intensities involving the propynylU methyl protons are smaller than those of the thymidine methyl protons due to their longer distances from protons on the aromatic ring of the base. Intrasugar NOE intensities are roughly the same for both triplexes, consistent with no dramatic changes in the sugar conformations, and the protonated cytidine cross-peaks involving the deoxyriboses have similar intensities compared to the corresponding cross-peaks of the other nucleosides. A nearly identical C1′ region is evident upon comparison of the ^{1}H−^{13}C HSQC spectra for both triplexes (DDD-EG HSQC not shown). It has been observed that the ^{13}C chemical shifts of the C1′ are somewhat correlated to the sugar conformation (*49*). Spectra of a variety of DNA and RNA duplexes and triplexes indicate that S-type sugars tend to have higher field C1′ chemical shifts than N-type sugars. In the related intramolecular triplex RDD-EG, which differs by having a third strand comprised of RNA (*50*), the protonated rCs have much higher C1′ chemical shifts than any of the dC^+ C1′ chemical shifts in either DDD-EG or PDD-EG.

PDD-EG Triplex Structure. The refined structures for PDD-EG are seen in Figure 5, and the comparison of the first two strands of the triplex (Watson−Crick duplex strands) with A-DNA and B-DNA is shown in Figure 7. Visually, the helical structure of the first two strands of PDD-EG appears more similar to A- than to B-DNA. However, all of the restrained sugars for PDD-EG are within the range of

Table 2: Input and Results of Structure Refinement

NOE restraints	680
intranucleotide	376
sequential	195
nonsequential internucleotide	109
involving exchangeable protons	138
torsion angle restraints	36
hydrogen bond restraints	70
relaxation matrix refinement	
number of peak integrals used at each mixing time	368
average $R^{1/6}$ factor before	0.1188 ± 0.0029
average $R^{1/6}$ factor after	0.0695 ± 0.0010
refinement statistics (10 final lowest-energy structures)	
NOE violations >0.5 Å	0
dihedral angle violations >5°	0
average pairwise rmsd (all heavy atoms) (Å)	1.03 ± 0.13
average rmsd from ideal covalent geometry	
bond lengths (Å)	0.022
bond angles (deg)	6.34

C2′-endo sugar puckers. Dihedral angle restraints derived from the DQF-COSY and information from NOESY and HSQC spectra, along with the lack of any NOE violations greater than 0.5 Å or dihedral angle violations greater than 5° (Table 2), suggest that the sugars are present in an S-type conformation for the majority of the time at equilibrium. As is the case for other triplex structures solved so far, the duplex has a lower twist than A-DNA and the minor groove is widened as a result of the unwinding of the helix to accommodate the third strand. However, in contrast to DDD-EG, the bases are displaced from the helical axis enough to create an axial hole, as is characteristic of A-DNA (Figure 7). The triplex has a bent appearance that is most likely a result of having the hydrophobic propynylUs in the third strand. Visually, all but the last two triplets have an inclination similar to that of A-DNA (discussed further below). The last two base triplets, which contain the two consecutive propynylUs, have a relatively small inclination, due to the steric clash of the two propynyl substituents (Figure 5B). Thus, the top part of the triplex fits best to A-DNA, while the bottom is forced to have more parallel triplets due to steric clash, thus decreasing the inclination and resulting in a more B-DNA-like helix.

The T_m of PDD-EG obtained from the UV melting studies shows the expected stabilization of this triplex versus DDD-EG due to the propynylU substitutions in the third strand (*22*). From the structure of the triplex, it can be seen that the propynyl group, which is coplanar with the heterocyclic base, extends the electron cloud of the propynyl base over the aromatic ring of the base below (Figure 5C). The increase in stacking interactions from the propynyl group explains some of the enhanced stability observed for the propynyl-substituted triplex. However, since both methyl and propynyl groups stabilize triplex formation, it is likely that the hydrophobicity of the propyne also plays an important role in the stability and conformation of the overall triplex structure. Therefore, the propynyl substitution might be responsible for a localized hydrophobic environment, which may contribute to the structural differences seen between the helices of this triplex and DDD-EG (*29*).

Helical Parameters of PDD-EG. The helical parameters for the Watson−Crick duplex part of the triplex were calculated using the program CURVES for a more quantitative comparison to duplex A-DNA and B-DNA and other triplexes. Two sets of values are reported, obtained by calculating the values using a local axis for each base pair (nonlinear axis) and one common axis for all base pairs (linear axis) (Figure 6). The default method in CURVES uses the nonlinear axis (*43*). However, the X displacement and inclination values calculated using the nonlinear axis do not reflect the visual characteristics of the structures as well as the values calculated with the linear axis do. Furthermore, for a direct comparison of two different structures, helical parameters should be calculated using the same axis for the two molecules being compared. This is most closely approximated when comparing to ideal A-DNA and B-DNA, which inherently have a linear axis, by using the linear axis calculated by CURVES, as discussed below.

Differences in the macroscopic features of A-DNA and B-DNA have been traditionally identified by their X displacement and inclination (*51*). However, as can be seen in Figure 6, different values for these parameters are obtained for PDD-EG depending on whether the nonlinear or linear axis is used, although the trends are the same in both cases. For PDD-EG, the nonlinear axis X displacement of the Watson−Crick base pairs varies from about −1.3 to −1.8 Å with an average of -1.61 ± 0.54 Å (Figure 6A). The linear axis calculation (-3.07 ± 0.57 Å) gives an average value about halfway between the values for A-DNA and B-DNA. An axial hole with a diameter smaller than that of A-DNA is observed in the structure of PDD-EG when the axis is superimposed on an A-DNA helical axis (Figure 7). The problem in comparing the helical parameters obtained is that, because the structures do not have the same calculated axis, they will have different results for their X displacements. This is especially noticeable when the helix is not strictly either A-form or B-form, but rather something between, which may contain a bend or kink and make it difficult to fit a single axis. This is the case for PDD-EG where the top fits well to A-DNA, but the axes for PDD-EG versus A-DNA are clearly different. For PDD-EG, the linear axis is closer to the placement of the A-DNA axis than is the nonlinear axis, and therefore gives parameters closer to those of A-DNA which makes it the better choice for this comparison. The inclination is also very dependent on the definition of the helical axis, as seen in Figure 6B, and varies from about −8.2 to 11.9° for the nonlinear axis and from 2.5 to 25.3° for the linear axis. The inclinations for standard A-DNA and B-DNA are about 19 and −5.9°, respectively. Again, the top half of the triplex has an inclination closer to that of A-DNA while the bottom half is closer to that of B-DNA. As can be seen in Figure 6C,D, the axis definition does not have as much effect on the rise and twist values. The rise per residue ranges from 2.6 to 4.1 Å, with an average of about 3.26 ± 0.23 Å (2.95 ± 0.32 Å for linear). The helical twist varies from 19 to 37°, with an average of $28.9 \pm 3.1°$ ($29.8 \pm 3.3°$ for linear). The helical twist between the first four Watson−Crick base pairs alternates between higher and lower values for ApG and GpA steps, respectively. This may reflect the alternation in the propynylUs in the Hoogsteen paired strand at this end of the molecule. Although the helical twist values are measured for the Watson−Crick base pairs, the position of the base pairs is affected by stacking interactions and steric constraints

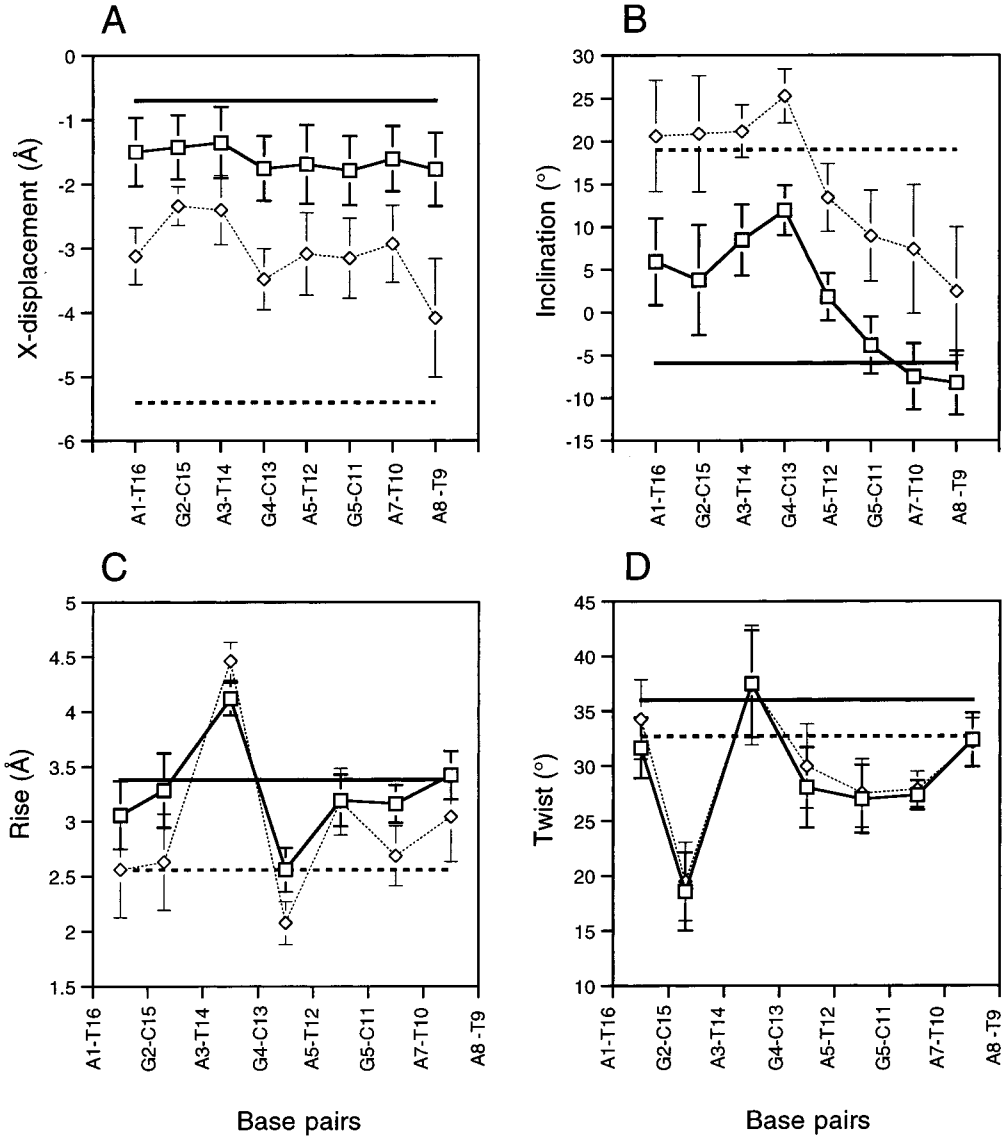

FIGURE 6: Helical parameters for PDD-EG: (A) X displacement, (B) inclination, (C) rise, and (D) twist. Error bars indicate the standard deviation of the measurement from the 10 lowest-energy structures. Values using both nonlinear (squares and solid line) and linear (diamonds and dotted line) axes are plotted. The values from fiber diffraction for standard A-DNA and B-DNA (*54*) are indicated by dashed and solid lines, respectively.

(as discussed above) for all three bases in the triplets.

Comparison of the PDD-EG and DDD-EG. The basic structural similarities of PDD-EG and DDD-EG are also common to other DNA triplexes (*6*). Both of these triplexes have predominantly S-type sugars in all three strands, with the exception of the C^+ sugars, which have a more even ratio of N/S conformations. The shift in N/S equilibrium for the C^+ can be explained by the gauche versus anomeric effects on the pseudorotational equilibrium (*52*, *53*), which was briefly discussed in the previous paper (*29*). The pseudorotational equilibrium is usually dominated by the gauche effect; however, protonation of the base enhances the normally weaker anomeric effect and drives the equilibrium more toward N in this case, leading to a fairly even population of N- and S-type puckers. Additionally, these triplexes have a widened minor groove from the untwisting of the duplex to accommodate the third strand, which is a typically observed feature of deoxyribose triplexes (*6*). This is reflected in both triplexes with average linear axis twist values of 31.4 ± 2.9° for DDD-EG and 29.8 ± 3.3° for PDD-EG. The rise for DDD-EG (3.3 ± 0.4 Å) and PDD-EG (3.3 ± 0.2 Å for nonlinear) is close to the average of 3.4 Å for B-DNA, with the exception of the linear axis values for PDD-EG (2.9 ± 0.3 Å).

There are several differences between the two structures, however. PDD-EG has an axial hole suggesting an X displacement larger than that of DDD-EG, where no axial hole is seen (Figure 7). There is also a difference in the inclination of the base pairs in the two structures. In DDD-EG, all of the base pairs have a small inclination, giving a straight appearance to the helix. In contrast, for PDD-EG, the inclination of the base pairs changes over the length of the helix (Figure 6B), which gives PDD-EG a slightly bent appearance when viewed from the side (Figure 7). Although the bend is not defined by any long-range NOEs, it is consistently defined locally by the short-range NOEs. Direct causal relationships between the differences observed between the two structures and particular NOE cross-peaks are difficult to establish. No small specific set of NOEs that leads to the differences in the structures could be identified.

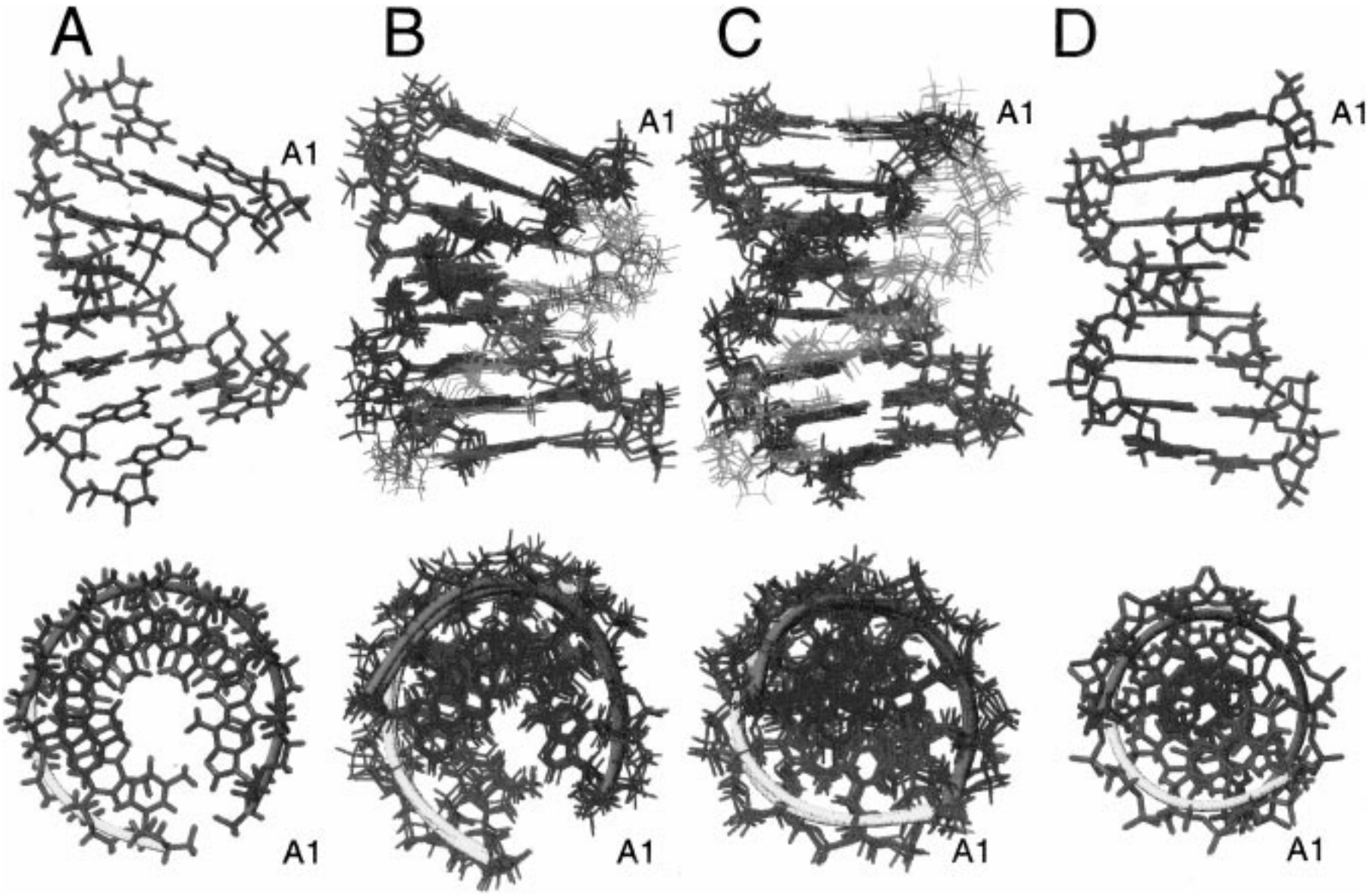

FIGURE 7: Side views (top) and axial views (bottom) of (A) A-DNA, (B) superpositions of the five lowest-energy structures of PDD-EG, (C) superpositions of the five lowest-energy structures of DDD-EG, and (D) B-form DNA. The standard A- and B-DNA helices were constructed from fiber diffraction-derived coordinates (*54*) and the same nucleotide sequence as PDD-EG using Insight II software (Biosym Technologies). The A-DNA and B-DNA helices are in the same orientations as the triplexes. The coloring scheme is the same as that in Figure 5 (both C and T nucleotides of the third strand for DDD-EG are cyan).

For example, test calculations in which the distance restraints involving the propyne methyl groups were removed still gave a similarly bent looking structure. The differences in the structural features of the two molecules are clearly greater than the variability within the ensemble of conformations calculated for each molecule, as shown by the comparison of the superpositions of the two structures (Figure 7B,C). The observed differences in the overall structures appear to be due to the effect of the five propynyl substituents. Therefore, the presence of the propynylUs in the third strand not only increases the stability of PDD-EG but also has an effect on the overall structure of the triplex.

Conclusions. The enhanced stability of triplexes and duplexes, achieved by incorporation of 5-(1-propynyl)-2′-deoxyuridines, most likely arises from the increases in both base-stacking interactions and local hydrophobicity of the propynyl substituent. The difference in the hydrophobicity of these modified nucleotides may play an important role in the overall structure of this triplex and other molecules containing propyne nucleotides. Factors other than increased stability induced by the propynylU play a role in making it a more effective antisense inhibitor versus other C5 aryl-substituted nucleotides. Perhaps it is the specific tertiary structure induced by the propynyl substituents that makes it a more potent antisense inhibitor, even over the more thermally stable thiazole substitutions. Further structural studies of duplexes and triplexes containing 5-(1-propynyl)-2′-deoxyuridines or other C5 aryl-substituted 2′-deoxyuridines may provide useful information about the effects on structure arising from the incorporation of different hydrophobic substituents into these oligonucleotides.

SUPPORTING INFORMATION AVAILABLE

Coupling constants and correlation coefficients for PDD-EG obtained using CHEOPS and sugar conformations obtained using PSEUROT (1 page). 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. Uhlmann, E., and Peyman, A. (1990) *Chem. Rev. 90*, 543−584.
3. Crooke, S. T. (1992) *Annu. Rev. Pharmacol. Toxicol. 32*, 329−376.
4. Maher, L. J., III (1996) *Cancer Invest. 14*, 66−82.
5. Neidle, S. (1997) *Anti-Cancer Drug Des. 12*, 433−442.
6. Wang, E., and Feigon, J. (1998) *Structures of Nucleic Acid Triplexes*, Oxford University Press, Oxford (in press).
7. Soyfer, V. N., and Potaman, V. N. (1996) *Triple-Helical Nucleic Acids*, Springer-Verlag, New York.
8. Sun, J.-S., Garestier, T., and Hélène, C. (1996) *Curr. Opin. Struct. Biol. 6*, 327−333.
9. Moraru-Allen, A. A., Cassidy, S., Alvarez, J. L. A., Fox, K. R., Brown, T., and Lane, A. N. (1997) *Nucleic Acids Res. 25*, 1890−1896.
10. Chan, P. P., and Glazer, P. M. (1997) *J. Mol. Med. 75*, 267−282.
11. Moses, A. C., and Schepartz, A. (1996) *J. Am. Chem. Soc. 118*, 10896−10897.
12. Ihara, T., Maruo, Y., Takenaka, S., and Takagi, M. (1996) *Nucleic Acids Res. 24*, 4273−4280.
13. Grant, K. B., and Dervan, P. B. (1996) *Biochemistry 35*, 12313−12319.
14. Moser, H. E., and Dervan, P. B. (1987) *Science 238*, 645−650.
15. Rininsland, F., Johnson, T. R., Chernicky, C. L., Schulze, E., Burfeind, P., and Ilan, J. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 5854−5859.
16. Guieysse, A. L., Praseuth, D., Grigoriev, M., Harel-Bellan, A., and Hélène, C. (1996) *Nucleic Acids Res. 24*, 4210−4216.
17. Giovannangeli, C., Diviacco, S., Labrousse, V., Gryaznov, S., Charneau, P., and Helene, C. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 79−84.
18. Wagner, R. W., Matteucci, M. D., Lewis, J. G., Gutierrez, A. J., Moulds, C., and Froehler, B. C. (1993) *Science 260*, 1510−1513.
19. Wagner, R. W. (1994) *Nature 372*, 333−335.
20. De Clercq, E., Descamps, J., Balzarini, J., Giziewicz, J., Barr, P. J., and Robins, M. J. (1983) *J. Med. Chem. 26*, 661−666.
21. Goodchild, J., Porter, R. A., Raper, R. H., Sim, I. S., Upton, R. M., Viney, J., and Wadsworth, H. J. (1983) *J. Med. Chem. 26*, 1252−1257.
22. Froehler, B. C., Wadwani, S., Terhorst, T. J., and Gerrard, S. R. (1992) *Tetrahedron Lett. 33*, 5307−5310.
23. Moulds, C., Lewis, J. G., Froehler, B. C., Grant, D., Huang, T., Milligan, J. F., Matteucci, M. D., and Wagner, R. W. (1995) *Biochemistry 34*, 5044−5053.
24. Gutierrez, A. J., Matteucci, M. D., Grant, D., Matsumura, S., Wagner, R. W., and Froehler, B. C. (1997) *Biochemistry 36*, 743−748.
25. Flanagan, W. M., Kothavale, A., and Wagner, R. W. (1996) *Nucleic Acids Res. 24*, 2936−2941.
26. Fenster, S. D., Wagner, R. W., Froehler, B. C., and Chin, D. J. (1994) *Biochemistry 33*, 8391−8398.
27. Bonham, M. A., Brown, S., Boyd, A. L., Brown, P. H., Bruckenstein, D. A., Hanvey, J. C., Thomson, S. A., Pipe, A., Hassman, F., Bisi, J. E., Froehler, B. C., Matteucci, M. D., Wagner, R. W., Noble, S. A., and Babiss, L. E. (1995) *Nucleic Acids Res. 23*, 1197−1203.
28. Matteucci, M., Lin, K. Y., Huang, T., Wagner, R., Sternbach, D. D., Mehrotra, M., and Besterman, J. M. (1997) *J. Am. Chem. Soc. 119*, 6939−6940.
29. Tarköy, M., Phipps, A. K., Schultze, P., and Feigon, J. (1998) *Biochemistry 37*, 5810−5819.
30. Robins, M. J., Barr, P. J., and Giziewicz, J. (1982) *Can. J. Chem. 60*, 554−557.
31. Gait, M. J. (1984) in *Oligonucleotide synthesis: a practical approach* (Rickwood, D., and Hames, B. D., Eds.) IRL Press, Oxford and Washington, DC.
32. Kumar, A., Ernst, R. R., and Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1−6.
33. Sklenář, V., and Bax, A. (1987) *J. Magn. Reson. 75*, 378−383.
34. Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) *J. Magn. Reson. 85*, 393−399.
35. Piantini, U., Sørensen, O. W., and Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800−6801.
36. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−360.
37. Santoro, J., and King, G. (1992) *J. Magn. Reson. 97*, 202−207.
38. Macaya, R. F., Schultze, P., and Feigon, J. (1992) *J. Am. Chem. Soc. 114*, 781−783.
39. De Leeuw, F. A. A. M., and Altona, C. (1983) *J. Comput. Chem. 4*, 428−437.
40. van Wijk, J., Huckriede, B. D., Ippel, J. H., and Altona, C. J. (1992) *Methods Enzymol. 211*, 286−306.
41. Neidig, K. P., Geyer, M., Gorler, A., Antz, C., Saffrich, R., Beneicke, W., and Kalbitzer, H. R. (1995) *J. Biomol. NMR 6*, 255−270.
42. Brünger, A. T. (1992) *X-PLOR (Version 3.1) Manual*, Yale University Press, New Haven, CT.
43. Lavery, R., and Sklenar, H. (1988) *J. Biomol. Struct. Dyn. 6*, 63−91.
44. Sklenář, V., and Feigon, J. (1990) *Nature 345*, 836−838.
45. Macaya, R. F., Wang, E., Schultze, P., Sklenář, V., and Feigon, J. (1992) *J. Mol. Biol. 225*, 755−773.

46. Feigon, J., Koshlap, K. M., and Smith, F. W. (1995) in *Methods in Enzymology* (James, T. L., Ed.) pp 225−255, Academic Press, San Diego.

47. Koshlap, K. M., Schultze, P., Brunar, H., Dervan, P. B., and Feigon, J. (1997) *Biochemistry 36*, 2659−2668.

48. Radhakrishnan, I., and Patel, D. J. (1994) *Biochemistry 33*, 11405−11416.

49. Lankhorst, P. P., Erkelens, C., Haasnoot, C. A., and Altona, C. (1983) *Nucleic Acids Res. 11*, 7215−7230.

50. Gotfredsen, C. H., Schultze, P., and Feigon, J. (1998) *J. Am. Chem. Soc.* (in press).

51. Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.

52. Plavec, J., Tong, W. M., and Chattopadhyaya, J. (1993) *J. Am. Chem. Soc. 115*, 9734−9746.

53. Thibaudeau, C., Plavec, J., and Chattopadhyaya, J. (1996) *J. Org. Chem. 61*, 266−286.

54. Arnott, S., and Hukins, D. W. (1972) *Biochem. Biophys. Res. Commun. 47*, 1504−1509.

BI972811U