Dual Recognition of Double-Stranded DNA by 2′-Aminoethoxy-Modified Oligonucleotides: The Solution Structure of an Intramolecular Triplex Obtained by NMR Spectroscopy

Marcel J. J. Blommers,* Francois Natt, Wolfgang Jahnke, and Bernard Cuenoud*,‡

Core Technologies, Novartis Pharma AG, P.O. Box, CH-4002 Basel, Switzerland

Received July 9, 1998; Revised Manuscript Received October 16, 1998

ABSTRACT: The solution structure of an intramolecular triple helical oligonucleotide has been solved by NMR. The third strand of the pyrimidine·purine·pyrimidine triplex is composed of 2′-aminoethoxy-modified riboses, whereas the remaining part of the nucleic acid is DNA. The structure around the aminoethoxy modification was obtained with the help of selective isotope labeling in conjunction with isotope-editing experiments. Dinucleotide steps and interstrand connectivities, as well as the complete backbone conformation of the triplex, were derived from *J*-couplings, NOEs, and ^{31}P chemical shifts. The structure of this triplex, solved by distance geometry, explains the extraordinary stability and increase in rate of triplex formation induced by 2′-aminoethoxy-modified oligonucleotides: apart from the formation of seven base triples, a well-defined hydrogen-bonding network is formed across the Crick−Hoogsteen groove involving the amino protons of the aminoethoxy moieties and the phosphates of the purine strand of the DNA. The modified strand adopts a conformation which is close to an *A-type* helix, whereas the DNA duplex conformation is best described as an unwound *B-type* helix. The groove dimensions and helical parameters of the 2′-aminoethoxy-modified rY·dRdY triplex are surprisingly well conserved in comparison with DNA triplexes.

Molecules that can sequence-specifically recognize DNA represent a new important category of compounds that have the potential to interfere with the regulation of gene expression, as they can compete with transcription factor activators or repressors for binding at specific gene promoters. Over the past decade, two major classes of molecules have been described (*1*). The first class is polyamide hairpins that bind DNA in the minor groove (*2*). Recently, using a combination of polyamide-containing imidazole, pyrole, and hydroxypyrole, a general code for the recognition of the four Watson−Crick base pairs has been identified (*3*). The second approach involves the binding of an oligonucleotide in the major groove of DNA to form a triple helical complex (*4, 5*). A triple helix can be formed over more than 20 base pairs but requires the sequence of the DNA target to be a polypurine/polypyrimidine stretch. Moreover, this approach is limited by the slow association rate for triplex formation, the weak binding affinity under physiological pH, and the lack of nuclease resistance of unmodified oligonucleotide. Much effort has been invested in the design of oligonucleotides with increased nuclease stability and higher affinity for double-stranded DNA (*6*).

Interestingly, these two classes of molecules derived their binding affinity and specificity from their interactions with the bases of DNA, whereas in protein−DNA complexes, many important contacts involve the bases as well as the phosphodiester backbone of DNA (*7, 8*). Therefore, molecules that could interact specifically and simultaneously with both the bases and the phosphodiester backbone should compete more efficiently with proteins for DNA binding. Recently, we reported that 2′-aminoethoxy-modified oligonucleotides bind double-stranded DNA with a very high affinity (*9*). It is the best modification of a series selected 2′-substituents. The 2′-aminoethoxy moiety provides a +3.5 °C increase in thermal stability per modification, compared to an unmodified DNA oligonucleotide control of the same sequence and length. Using a fully 2′-aminoethoxy-modified 15-mer oligonucleotide, we also found that the high affinity for DNA binding was associated with an association rate of 3.6×10^4 M^{-1} s^{-1}, which is a factor 10^3 faster than the rate measured for a DNA control. Importantly, structure−affinity relationships and molecular dynamics simulations suggested that these oligonucleotides can contact both the bases and the phosphodiester backbone of DNA at each base pair of the duplex. Moreover, molecular dynamics simulations also suggested that the aminoethoxy side chain adopts a well-defined gauche$^+$ conformation in the complex, enabling an optimum interaction between the positively charge amino residue and a pro-R oxygen of a negatively charged phosphate of the DNA second strand.

Whereas triplex structures in the past have only been characterized at low resolution using fiber diffraction, in recent years DNA triplexes have been characterized by nuclear magnetic resonance at high resolution (*10*). Structures have been reported for Y·RY triplexes (*11*−*13*) and for structures of alternative base triplexes positioned into Y·RY triplexes (*14*−*16*). Very recently, the structure of a Y·RY triplex containing an RNA third strand has been solved (*17*). Structures have been reported for R·RY triplexes as well

* Authors to whom correspondence should be addressed.
‡ Present address: Novartis Horsham Research Center, Wimblehurst Road, Horsham, West Sussex, RH12 4AB, England.

10.1021/bi9816352 CCC: $15.00
Published on Web 12/01/1998

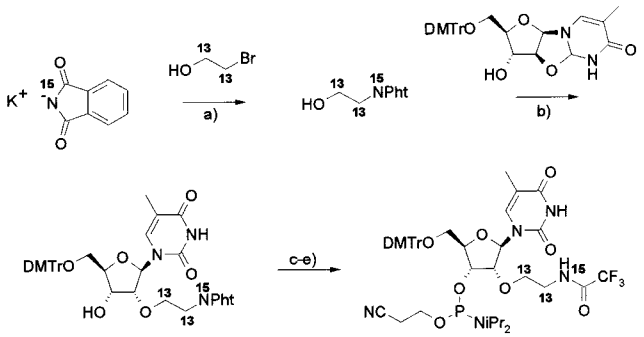
(*18*, *19*). Furthermore, several triplexes with unnatural bases in the third strand have been characterized by NMR[1] (*20*−*22*).

Here, we report the NMR solution structure of an intramolecular triple helix complex containing a 2′-aminoethoxy-modified third strand that unambiguously confirmed our previously proposed interaction between the amino group of the substituent and one phosphate of the DNA second strand. This study provides also the detailed triple helical structure whose third strand is RNA like.

MATERIALS AND METHODS

2′-Aminoethoxy-Modified Phosphoramidites. Protected 2′-aminoethoxy-modified phosphoramidites thymidine and C5 methyl cytidine were synthesized as previously reported (*9*). A new, shorter route was developed for the synthesis of the isotopically labeled $^{13}C-^{13}C-^{15}N$ protected 2′-aminoethoxy thymidine phosphoramidite **4** (Scheme 1). First, the isotopically labeled aminoethoxy side chain was assembled by reaction of ^{15}N-potassium phthalimide **1** (Fluka) with 2-bromoethanol (1,2-$^{13}C2$), (Cambridge Isotope Laboratories) to provide the 2-phthalimidoethanol **2** in good yield (*23*). Using a modified procedure (*24*), subsequent reaction of **2** with 5′-DMTr protected 2,2′-anhydrothymidine afforded intermediate **3**, which was further transformed into the desired protected phosphoramidite **4** by standard procedures. The identity of **4** was confirmed by proton, fluor, and phosphorus NMR spectroscopy and high-resolution ES-MS (observed ion, [M−H+2Na]+; calculated mass, 925.3417; observed mass, 925.3421).

Scheme 1: Synthesis of the Protected Phosphoramidite **4** Containing the $^{13}C-^{13}C-^{15}N$ Isotopically Labeled 2′-Aminoethoxy Side Chain[a]



[a] (a) DMF, 24 h, 135 °C, 72%; (b) 2 (5.8 equiv), $Ti(O^iPr)_4$ (2.4 equiv), $NaHCO_3$ (cat), THF, 140 °C, 2 days, 32%; (c) hydrazine (5 equiv), EtOH:H_2O (95:1), 2 h, 22 °C; (d) CF_3COOEt (10 equiv), Et_3N (7 equiv), 1 h, 22 °C, 73% for two steps; $[(^iPr)_2N]_2POCH_2CH_2CN$ (2.2 equiv), $(^iPr)_2NH_2^+$tetrazole$^-$ (2.4 equiv), CH_2Cl_2, 3 h, 22 °C, 69 %.

[1] Abbreviations: NMR, nuclear magnetic resonance; COSY, correlation spectroscopy; HSQC, heteronuclear single quantum correlation; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; TOCSY, total correlated spectroscopy; Maldi-TOF, matrix-assisted laser desorption-time-of-flight; ES-MS, electrospray mass spectroscopy; DMTr, dimethoxytrityl; DSS, 2,2-dimethyl-2-silapentane-5-sulfonic acid; DNA, deoxyribonucleic acid; RNA, ribonucleic acid; T, thymine; C, cytosine; C^+, N3-protonated cytosine; A, adenine; G, guanine; Y·RY, triplex with pyrimidines in strands 1 and 3, and purines in strand 2; rY·dRdY, Y·RY triplex with riboses in strand 3 and deoxyriboses in strands 1 and 2; R·RY, triplex with pyrimidines in strand 1 and purines in strands 2 and 3.

Oligonucleotide Synthesis, Purification and Characterization. Nucleoside phosphoramidites were oligomerized using an ABI 394 RNA/DNA synthesizer. The resulting oligonucleotide was deprotected in 0.9 mL of concentrated aqueous ammonia and 0.1 mL of piperidine at 55 °C for 18 h, purified by semipreparative ion-exchange chromatography (ResourceQ 1 mL, Pharmacia; buffer A, 10 mM NaOH; buffer B, 10 mM NaOH and 1.5 M NaCl, 0−100% B in 30 column volume, 1.5 mL/min), and desalted (Sep-Pak cartridges, Waters). Two oligonucleotides were synthesized.

The unlabeled oligonucleotide was 5′-cttttccTTTTTCCTTTTCTTTTTGAAAAGG and the isotopically labeled oligonucleotide was 5′-ctt***t***tccTTTTTCCTTTTCTTTTTGAAAAGG [t, 2′-aminoethoxy-T; c, 2′-aminoethoxy-C5-Me-C, ***t***, 2′-aminoethoxy($^{13}C-^{13}C-^{15}N$)−T]. The purity and identity of each oligonucleotide was determined by ion-exchange chromatography (same conditions than purification but on analytical scale, 0.5 OD_{260} injected) and by mass spectroscopy [expected mass (unlabeled oligonucleotide), 9844.7; measured (Maldi-TOF), 9850; expected mass (isotopically labeled oligonucleotide), 9847.7; measured (ES-MS, 9847.3). Concentration of the oligonucleotides was determined at 260 nm using a calculated molar extinction coefficient of 276 640 (M^{-1} cm^{-1}).

NMR Experiments. The oligonucleotides were dissolved in a 50 mM phosphate buffer containing 100 mM NaCl. The pH was adjusted to 5.7 using DCl. D_2O (5%) was added to the H_2O samples. The D_2O samples were made by lyophylization of the H_2O samples, dissolving them in D_2O and adjusting the pH. Each sample was heated to 60 °C and slowly cooled directly after the preparation. The concentrations of the unlabeled and selectively labeled oligonucleotides were 2.2 and 1.2 mM, respectively. The samples were carried over in susceptibility-matched 5 mm NMR tubes (Shigemi Co., BMS-005V) using a sample volume of 200 μL. The use of small sample volumes in conjunction with susceptibility-matched NMR tubes excluded the inhomogeneous region of the radio frequency field γB_1, which increased the quality of the spectra (*25*). All NMR experiments were performed on a Varian Unity*plus* 600 spectrometer, operating at a ^1H-frequency of 600 MHz, using a ^1H,^{13}C, and ^{31}P triple resonance probe with *z*-gradients.

One dimensional ^1H spectra of the sample in H_2O were recorded using the watergate solvent suppression method (*26*). NOESY experiments (*27*) of the sample in H_2O were recorded at 0 and 20° C with a mixing time of 100 ms using the watergate method. A relaxation delay of 2 s was used. All experiments of the sample in D_2O were recorded at 35 °C. NOESY experiments of the sample in D_2O were recorded using a mixing time of 50, 100, and 200 ms, respectively, and a relaxation delay of 5 s. Also, a double quantum-filtered COSY (*28*) was recorded. Clean TOCSY experiments (*29*, *30*) were recorded with mixing times of 20, 40, and 60 ms. These phase-sensitive 2-D spectra were recorded with 512 experiments and 4096 complex data points using quadrature detection (*31*) in both dimensions.

^{13}C-edited spectra were recorded with the selectively labeled triplex oligonucleotide. A sensitivity-enhanced ^1H-^{13}C HSQC, optimized for the detection of CH_2 groups (*32*), was collected using 128 increments with 8 transients/increment and a spectral width of 8000 Hz. 2-D ^{13}C-edited NOESY spectra were also recorded sensitivity-enhanced and

optimized for the detection of CH_2 groups. Mixing times of 50, 100, 200, and 300 ms were used. The recycle delay was 2 s. The spectral width in the indirect ^{13}C dimension was 5000 Hz, and 384 complex increments were recorded using 32 transients per increment.

All data were transferred to a Silicon Graphics workstation and processed using VNMR 5.3 software. The 2-dimensional data were apodized using shifted sinebell and Gaussian filters and zero-filled to 2048 data points in the t_1 dimension. Peak volumes from the NOE spectra were obtained using integration using VNMR 5.3. Cross-peak multiplet patterns in the double quantum-filtered COSY experiment were analyzed with help of simulated cross-peaks obtained using the programs SPHINX and LINSHA (*33*).

Structure Elucidation. Distance geometry calculations were carried out using the program DGII (*34*) which is integrated into the molecular modeling package InsightII (MSI), version 97.0, running on a Silicon Graphics Octane with two R10000 processors. The backbone torsion angles γ, β, and ϵ were defined using upper and lower bounds which vary 15° around the rotamers usually found; i.e., γ^+ is 38−68°, β^+ is 180−211°, and ϵ^t is 177−207°. The torsion angles α and ζ were restraint between 285 and 315°. The torsion angle N−C−C−O in the aminoethoxy was restraint 50−70°. Hydrogen bonds in the classical T·AT and C^+·GC triples were constraint to a distance of 1.8 Å. Sugar conformations were restraint to an S-type conformation by restraining H4′−H2″ (3.8−4.2 Å), H3′−N1/9 (4.3−4.7 Å), and H2′−C5′ (3.1−3.5 Å) and by using dihedral restraints ν_o (0.5−10.5°) (around O4′−C1′), ν_1 (−29.8 to −19.8°) (around C1′−C2′), ν_2 (29.6−39.6°), ν_3 (−36.2 to −26.2°), ν_4 (10.9−20.9°) and to an N-type conformation by restraining H4′−O2′ (2.5−2.9 Å), H3′−N1 (2.8−3.2 Å), H2′−C5′ (4.4−4.8 Å) and by using the dihedral restraints ν_o (−20.9 to −10.9°), ν_1 (26.2−36.2°), ν_2 (−39.6 to −29.6°), ν_3 (19.7−29.7°), ν_4 (−10.5 to −0.5°). Combination of these three distance restraints with dihedral restraints resulted in a better convergence than using dihedral restraints alone. Distance restraints were estimated with the help of the build-up curves and the two-spin approach as well as using a relaxation matrix approach (*35*). The latter approach was exclusively used for the NOEs involving sugar residues with rigid sugar conformations. Upper and lower distance restraints were obtained by adding and subtracting 15 and 10%, respectively. Upper and lower bounds were smoothed using triangle smoothing. A total of 100 structures were embedded and optimized using a four-dimensional Cartesian space. The embedded structures were further refined with 20 000 steps of simulated annealing. Finally, the resulting structures were subjected to 2000 steps of conjugated gradient energy minimization using DISCOVER and the AMBER force field (*36*). For the aminoethoxy group, we used parameters of protonated lysine. Electrostatic interactions were scaled to $1/4r_{ij}$. The best 20 structures, qualified on the basis of energy and an analysis of restraint violation, were used to calculate groove dimensions and helical parameters. The latter were calculated using InsightII with a module based on Dickerson's NEWHEL91 program.

RESULTS

Triplex Formation. To study the triplex formation of a 2′-aminoethoxy-modified oligonucleotide with double-stranded DNA, we designed a 31-mer oligonucleotide, for which the sequence and assumed secondary structure is depicted in Figure 1. It is designed to form an intramolecular triple helical structure. The 5′-end of the oligonucleotide contains seven 2′-aminoethoxy-substituted ribose sugars, whereas all other residues have regular deoxyribose sugars. The cytosines at positions 1, 6, and 7 are methylated at C5 (*37*). The triplex belongs to the class of pyrimidine·purine·pyrimidine (Y·RY) triplexes. We anticipated that it might be difficult to obtain detailed and unambiguous information about the structure of the aminoethoxy moiety. For this reason, we synthesized also the 31-mer in which we selectively labeled the aminoethoxy substituent attached to residue T_4 with ^{13}C and ^{15}N and used isotope editing to simplify the NOE spectra.

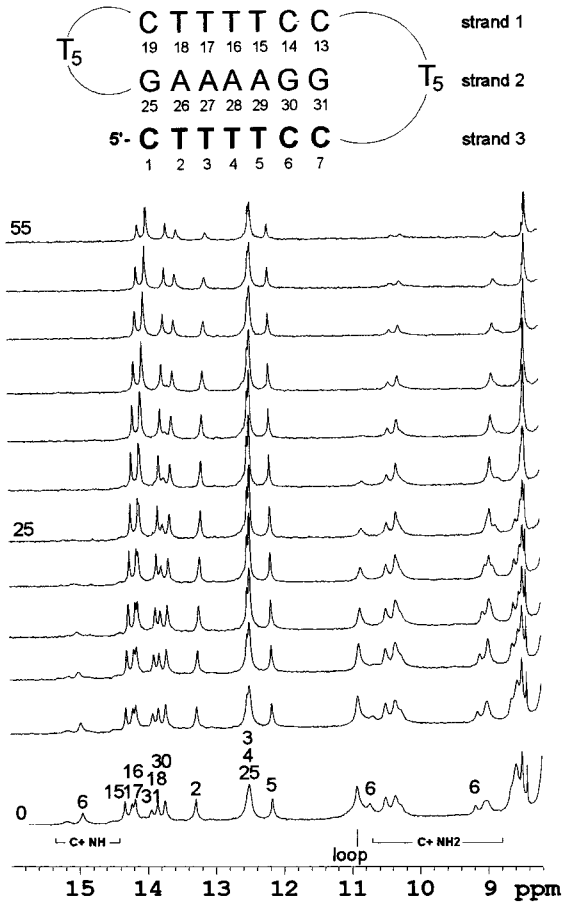


FIGURE 1: Low field part of the 1-dimensional spectra of the triplex dissolved in H_2O at pH 5.7 recorded at different temperatures (0−55 °C). The secondary structure of the triplex, the residue and the strand numbering is given on top. The third strand is composed of 2′-aminoethoxy modified riboses. The resonance assignment is indicated in the bottom spectrum which is recorded at 0 °C.

Figure 1 shows the downfield part of the ^1H spectrum of the 31-mer oligonucleotide dissolved in water at pH 5.7 as a function of the temperature. The spectrum could be assigned using watergate NOESY experiments. It is clearly visible that, at lower temperatures, the position of the resonances confirms the presence of the seven base triples as indicated in the secondary structure sketched on top, i.e., the bases of strands 1 and 2 interact via a Watson−Crick hydrogen-bonding scheme, whereas the bases of strands 2 and 3 form Hoogsteen pairs (*38*−*40*). The signals corresponding to the unpaired loop thymines around 11 ppm

Table 1: Chemical Shifts for Protons in the Triplex at 35 °C, pH 5.7, Relative to DSS

	NH	NH2[a]	NH2	H2,5	H8,6	H1′	H2′	H2″	H3′	H6′	H6″	H7′	H7″
C1	ex	8.93	10.44	2.08	8.47	6.03	4.56				4.32	3.42	3.42
T2	13.71			1.75	8.06	5.91	4.57			4.26	4.38	3.54	3.44
T3	12.51			1.68	7.95	5.83	4.52		4.60	4.28	4.47	3.53	3.44
T4	12.48			1.67	8.00	5.86	4.47		4.53	4.28	4.45	3.52	3.43
T5	12.19			1.68	8.04	5.88	4.56		4.44	4.20	4.40	3.51	3.43
C6	ex	8.93	10.32	1.83	8.10	5.94	4.49		4.31	4.15	4.29	3.47	3.40
C7	ex	8.93	10.32	1.66	8.08	5.96	4.07		4.49		4.25	3.39	3.39
C13		8.54	7.27	6.03	7.81	6.16	2.72	2.42	4.75				
C14		8.46	7.28	5.76	7.63	6.05	2.30	2.26					
T15	14.19			1.71	7.55	6.11	2.27	2.27					
T16	14.07			1.77	7.48	6.13	2.27	2.27					
T17	14.07			1.70	7.49	6.09	2.27	2.27					
T18	13.78			1.70	7.49	6.09	2.27	2.27					
C19		8.11	6.74	5.47	7.50	6.01	2.43	2.34	4.74				
G25	12.48				7.37	5.97	2.64	3.01	4.78				
A26		7.65	7.46	7.22	7.15	6.12	2.65	2.96	4.82				
A27		7.72	7.52	7.16	7.07	6.05	2.63	2.92	4.81				
A28		7.72	7.52	7.16	7.00	5.96	2.51	2.83	4.80				
A29		8.02	7.78	7.48	7.19	6.01	2.32	2.78	4.80				
G30	13.62				7.12	5.98	2.40	2.67	4.73				
G31	13.71				7.75	5.84	2.57	2.40	4.54				

[a] The hydrogen-bonded proton resonates downfield relative to the other proton in each pair.

broaden at elevated temperatures. The signals of the first and last triple broaden at higher temperatures due to exchange with the solvent. The signals around 9 and 10.5 ppm correspond to amino proton resonances of N3-protonated cytosine residues, which are characteristic for Y·RY triplexes. Interestingly, the presence and reduced exchange of the thymine imino proton signals of the third strand at elevated temperatures indicates that the third strand remains bound to the duplex within the whole temperature range measured (0−55 °C). This is consistent with the high unimolecular UV-melting temperature (81 °C) of this complex observed under similar buffer and pH conditions (100 mM NaCl and 10 mM phosphate, pH 5.5). It is noteworthy that this triplex which requires protonation of three C5-methylcytosines, is remarkably stable at pH 5.7, in contrast to other DNA triplexes for which the duplex−triplex transition is only reached at pH 4.8 (*15*).

Resonance Assignment. The assignment of resonances was based on standard procedures using watergate NOESY experiments of the sample in H_2O and NOESY and TOCSY experiments of the sample in D_2O (*41, 42*). The results of the resonance assignment are collected in Table 1. The resonance assignment of the aminoethoxy moieties will be described in one of the following sections since it is connected to the conformational analysis. As expected from our previous interpretation of the imino proton spectra in Figure 1, the NOE connectivities correspond to a triplex structure. Thus, on the basis of interstrand NOEs involving aromatic protons, imino protons, and amino protons, the presence of the regular C^+·GC and T·AT triples could be confirmed. On the basis of the sequential NOEs in each strand and the typical interstrand NOEs found for triplex structures, the stacking of the seven triples could be demonstrated. The major striking difference with respect to the observed NOEs in this triplex in comparison with previously studied DNA triplexes is the presence of significantly stronger interstrand H1′−H8 NOEs involving the protons of the third and second strands, respectively. Although the NOEs indicate stacking of one or more bases of the loops on the triples at each end of the triplex, the loop imino protons exchange very rapidly with the solvent and a relatively open or flexible structure of the loops is assumed.

We note that, in analogy to DNA triplexes (*11*), the chemical shifts of the aromatic protons are shifted with respect to duplexes (cf. Table 1). For the following chemical shift data, only the data for the central five triples are considered: the H8 resonances (second strand) resonate between 7.00 and 7.19 ppm, whereas in B-DNA duplexes, they are found between 7.4 and 8.4 ppm (*43*); the H6 resonances of bases of the first strand resonate between 7.95 and 8.10 ppm, whereas they are found in B-DNA between 7.0 and 7.8 ppm. Thus, the H8 resonances shift upfield, and the H6 resonances shift downfield compared to B-DNA. When the chemical shifts for B-DNA are estimated according to Wijmenga and co-workers (*44*) and compared with the shifts for the triplex, the average shift is 0.73 and −0.37, respectively. The most plausible explanation of the upfield and downfield shifts are the ring current effects of the aromatic systems in the base triples, e.g., the H8 proton (second strand) is located just above the 5′-neighboring purine base in the triplex, but is moved aside the purine base in the duplex. This is evident from the published DNA triplex structures as well as from the one currently studied (vide infra). Thus, in the case of the triplex, the H8 spin experiences an upfield shift. The H6 proton in first strand of the triplex structure is located above the carbonyl of the pyrimidine base of the 3′-neighboring residue and experiences a downfield shift. Overall, the chemical shift positions of 2′-aminoethoxy modified rY·dRdY triplex with respect to the Y·RY triplexes are well conserved and suggest that the stacking of the base-triples is similar.

Characterization of the Triplex Structure. The structural characterization of this triplex has two important implications. In the first place, it is one of the first triple helices in which the third strand is not DNA, but rather should be classified among RNA because the 2′-position carries an oxygen. It is thus very interesting to obtain a full structure. Second, the conformational details of the aminoethoxy-modified sugars and its possible interaction with other parts of the triplex

should provide significant structural insights in the origin of the favorable kinetic and thermodynamic properties induced by the presence of this substituent (*9*).

The structure of the triplex can be obtained from the NOE and *J*-coupling data collected for the various proton pairs. A more systematic analysis in terms of the local conformation of mono- and dinucleotide units, the sugar−phosphate backbone, and the aminoethoxy moieties is a very useful intermediate step in the full structure determination and will be presented below. This will also include dynamical aspects, which do not follow from the structure calculations per se.

Local Conformation of the Residues. The conformation of the furanose sugars is often described as an equilibrium between the *N-type*, i.e., close to C3′-endo, and *S-type*, i.e., close to C2′-endo, conformation. In this study, the cross-peak patterns in the double quantum-filtered COSY spectrum have been analyzed in conjunction with simulations using the programs SPHINX and LINSHA to obtain information of the sugar conformation. Unfortunately, because of overlap of especially the pyrimidine sugar resonances, the conformational analysis could not be performed for each sugar, but for each strand, a significant number of cross-peaks could be analyzed. Based on this analysis it is concluded that the sugar conformation of the first and second strand is best described by an equilibrium involving the *N-type* and *S-type* conformations. The *S-type* conformation is dominant (>70%). For the third strand, the chemical shifts of the H2′ resonances fall in the region for riboses and are well-separated from the chemical shifts of the H2′ and H2″ resonances of the deoxyribose sugars. The seven cross-peaks to be expected for the H1′−H2′ connectivies in the double quantum-filtered COSY and TOCSY spectra in the ribose region are absent, even in the TOCSY spectra recorded with a mixing time up to 60 ms. Therefore, it is concluded that the seven aminoethoxy-substituted sugars adopt predominately an *N-type* conformation. On the basis of NOE intensities of the NOEs involving the aromatic proton and sugar proton resonances, it followed that all bases adopt a position in the anti region.

Conformation of the Backbone. There are a few features which give qualitative information on the conformation of the sugar−phosphate backbone. Because the bases are found in the anti conformation, one would expect NOEs involving the aromatic proton and the H5′ or H5″ resonances at short mixing times in case the torsion angle γ deviates from gauche$^+$ (*35*). These potential NOEs fall in an isolated region of the NOE spectrum. In the case of the triplex under investigation, we do not observe such NOEs for the residues in the three strands.

When the conformation of the torsion angle β deviates from trans, the H5′ or H5″ resonance has a large ^3J-coupling to the phosphorus nucleus ($J \approx 23$ Hz) which is easily observed in the double quantum-filtered COSY and NOESY spectra (*35*). Although several cross-peaks involving H5′ or H5″ resonances of residues in the three strands could be observed, none of them were split by a large coupling.

In case the torsion angles α and ζ around the phosphorus nucleus change to a conformation where one of these torsion angles is trans, the ^{31}P resonance usually experiences a downfield shift of about 2.5 ppm (*45*). The ^{31}P spectrum measured for the triplex is a poorly dispersed spectrum. The ^{31}P resonances are found in a region between 1.23 and −0.19 ppm relative to cAMP.

Finally, it should be remarked that no *E.COSY-type* splitting patterns are observed for the H2′−H3′ cross-peaks of the deoxyribose sugars in the double quantum-filtered COSY spectrum, which indicates that ϵ is in the trans domain (*46*), at least for the first two strands. This could only be checked for T5 and C7 in the third strand, because the cross-peaks in this case are located too close to the diagonal. It should be noted, however, that ϵ is never found in a gauche$^-$ conformation in combination with an N-pucker of the ribose, because of steric reasons. The gauche$^+$ conformation is forbidden.

Taking these observations together, it is concluded that the backbone of the three strands adopts a γ^+-β^t-α^--ζ^--ϵ^t conformation and thus experiences only minor alterations compared to regular A-DNA or B-DNA conformations.

Long-Range Distance Restraints. Apart from short-range NOE data, which are available within residues and between adjacent residues, NOEs are also observed involving resonances of the third and the second strand. Remarkably, NOEs involving the H8 resonances of the second strand and the H1′ resonances of the third strand (in the residue corresponding to the subsequent triple) are quite strong. The NOE intensity corresponds to a calculated distance in the range 3.4−3.8 Å, which is much shorter than observed in DNA triplexes. This is probably because the sugar conformation in the third strand has switched to an *N-type* conformation, which moves the H1′ atom into the direction of the second strand (vide infra).

Analysis of the Aminoethoxy Substituent. Having described the NMR data which are related to the conformational features of the triplex, it is now interesting to discuss, in detail, the information which can be obtained from the aminoethoxy substituent present at the 2′-position of the riboses in the third strand. Because the stereospecific assignment of the prochiral protons of the aminoethoxy moieties is intimately connected with conformation, they are discussed hand-in-hand with the conformational analysis. The definition of the atom names is given in Figure 3 (left panel).

The assignment of the aminoethoxy ^{1}H resonances was based on ^{13}C chemical shifts and followed from a ^{1}H-^{13}C HSQC spectrum recorded of the selectively labeled oligonucleotide. The C−CH_2−N protons resonate around 3.5 ppm (^{13}C = 42.3 ppm) and the O−CH_2−C protons resonate around 4.4 ppm (^{13}C = 60.0 ppm). These resonances are shifted relative to their random coil shifts measured for a single-stranded oligonucleotide with an aminoethoxy-substituted (3.8 and 3.9 ppm). Thus, the O−CH_2−C protons shift 0.5 ppm downfield upon triplex formation, which suggests that their local environment differs with respect to the assumed random conformation. This downfield shift is dominated by ring current shifts of the attached base which becomes apparent after the structure determination (vide infra). Unfortunately, due to the fast exchange of the NH_3 protons with the solvent, we cannot not take advantage of the incorporation of ^{15}N. This is not totally unexpected; the exchange rates are likely similar or worse than those of lysine ($>3 \times 10^3$) (*47*).

Because the resonances of the aliphatic protons are not degenerate, those for residue T_4 can be easily picked up in the ^{13}C-isotope-edited NOE spectrum displayed in Figure 2B. This spectrum is a subspectrum of the nonedited NOE spectrum in Figure 2A in that only NOE cross-peaks

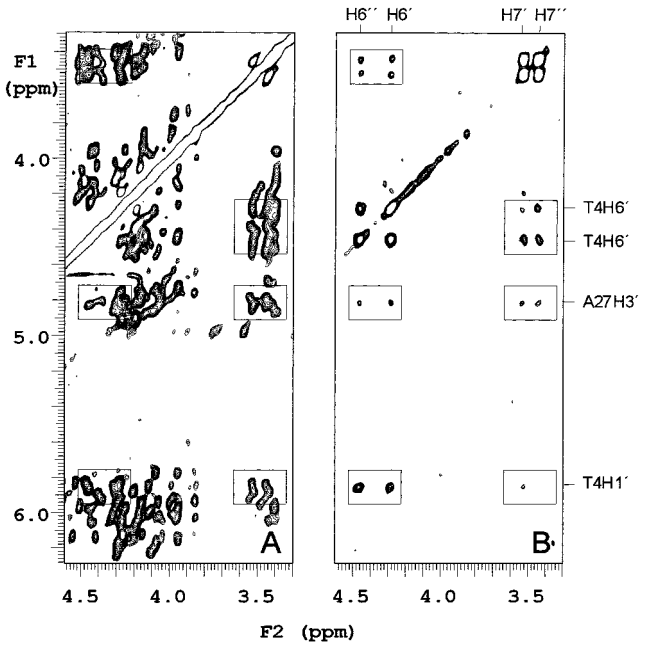


FIGURE 2: Conventional NOESY spectrum (A) and isotope edited NOESY spectrum (B) recorded at 35 °C. The assignment of the NOEs between aminoethoxy protons (bound to ^{13}C) and sugar protons (bound to ^{12}C) is indicated. The boxes in the isotope edited spectrum are drawn at the same position in the NOESY spectrum. Both spectra are recorded with a mixing time of 200 ms.

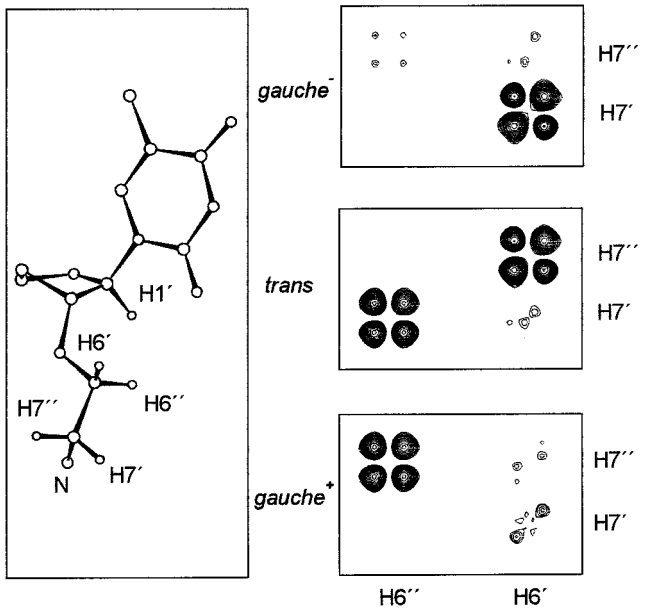


FIGURE 3: Simulation of the cross-peaks in the double-quantum filtered COSY spectrum for the gauche$^-$, trans, and gauche$^+$ conformation of the aminoethoxy substituent. The nomenclature of the protons is given in the ball-and-stick model in the left panel. The model shown has the O−C−C−N atoms of the aminoethoxy in gauche$^+$ conformation. For clarity only the relevant atoms are shown.

involving ^{13}C-bound protons are visible. The isotope editing simplifies the spectrum significantly and allows the selective observation of NOEs to the aliphatic protons of T_4. The most relevant signals are boxed. For comparison, the same boxes are drawn in the NOE spectrum shown in Figure 2A. The ^{13}C-isotope-edited NOE spectrum allows the unambiguous assignment of the NOEs involving the H6′/H6″ proton resonances and the H1′ resonance. Particularly interesting is the NOE observed involving $A_{27}H3'$ and the aminoethoxy aliphatic resonances. From the ^{13}C-edited NOE spectrum, it is obvious that the magnetization is originating from ^{13}C-bound protons, i.e., from the aminoethoxy moiety of residue T_4. In the regular NOE spectrum, the corresponding region is too crowded, and this particularly important long-range NOE would not have been assigned. It is interesting to note that, at shorter mixing times (not shown), the cross-peaks involving H3′ protons and the downfield-shifted C7′-bound protons are of medium intensity, whereas the NOEs involving the H3′ protons and the upfield shifted C7′-bound protons are very weak or absent. It is thus important to translate these NOEs into distances as well as stereospecific assignments. This is elucidated below.

For the conformation of the aminoethoxy moiety, the three staggered states around the O−C6−C7−N atoms, i.e., gauche$^+$ (60°), trans (180°), and gauche$^-$ (−60°), are considered. For these three rotamers, the coupling patterns to be expected in double quantum-filtered COSY are different (cf. Figure 3). According to the Karplus equation (*48*), for the trans case, there are two proton pairs which experience a large ^3J-coupling around 10 Hz, whereas in the other two conformations, only one large ^3J-coupling is calculated. When the experimental double quantum-filtered COSY in this region is inspected (Figure 4C), seven antiphase multiplet signals are observed for the seven aminoethoxy groups. They are all found in the region corresponding to the upfield shifted C7-bound proton resonances and the downfield-shifted C6-bound proton resonances. On the basis of this observation, one must conclude that the O−C6−C7−N atoms are in gauche conformation, but we cannot yet discriminate between gauche$^+$ and gauche$^-$ on these data alone.

In Figure 4B, part of the NOE spectrum recorded with a short mixing time (50 ms) is shown in which the H1′−H6′/H6″ cross-peaks are expected. Very strong NOEs are observed involving the H1′ resonance and the downfield-shifted H6′/H6″ resonance. The NOE intensity is roughly the same as observed for the H5−H6 pair of the cytidine residues. The upfield shifted resonance is observed in NOE spectra recorded with longer mixing times, as shown, e.g., in Figure 2. On the basis of NOE build-up curves, it was estimated that the corresponding distances are 2.6 and 3.2 Å, respectively. A grid search analysis, where the torsion angles C1′−C2′−O2′−C6′ and C2′−O2′−C6′−C7′ were systematically varied, revealed that these distances are only possible when the short distance is assigned to the H6″−H1′ pair. In the reverse assignment, the aliphatic protons clash with other sugar protons. As a consequence of the assignment of H6′ and H6″ resonances, the C7 protons can now also be assigned stereospecifically, based on double quantum-filtered COSY (cf. Figure 3). In conclusion, the O−C6−C7−N atoms are in gauche$^+$ orientation, as drawn in the left panel of Figure 3, and the prochiral aliphatic protons could be unambiguously and stereospecifically assigned (cf. Figure 4). On the basis of this analysis, we can assign the previously mentioned important long-range NOE involving the H3′ of the purine residue in the second strand and the H7′ of the aminoethoxy moiety, which is of medium intensity at short mixing times. This information will enable us to derive a very accurate structure of the triplex around the aminoethoxy moiety.

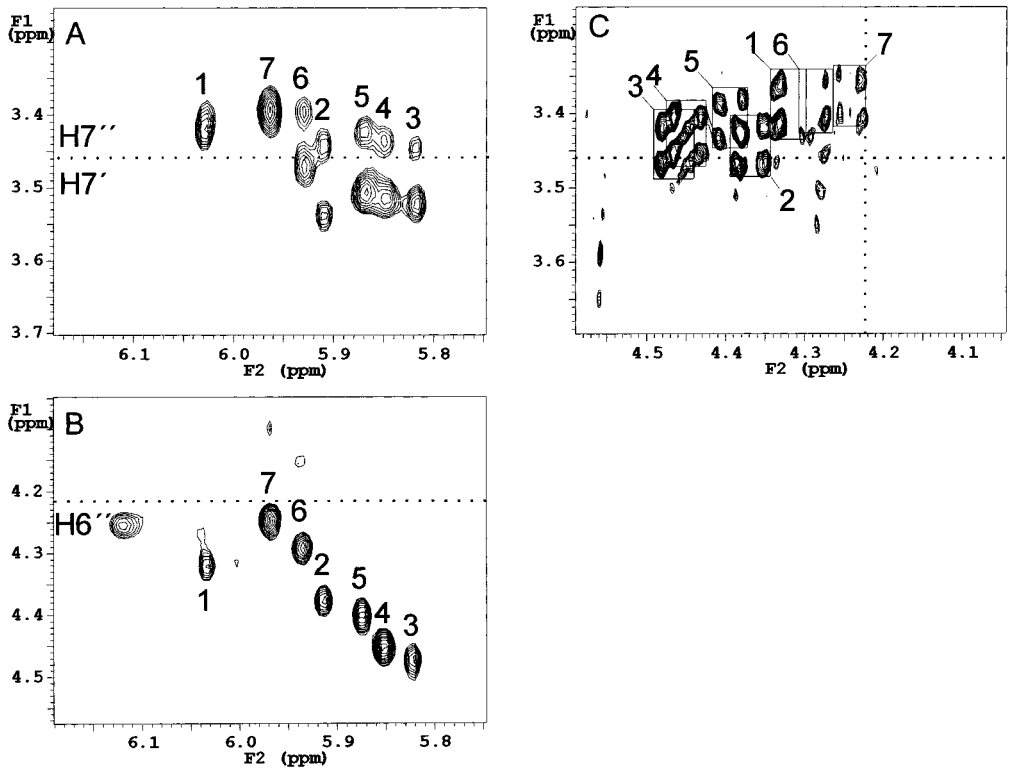


FIGURE 4: Part of the NOESY and double-quantum filtered COSY spectrum recorded of the triplex at 35 °C. Combination of these spectra with the simulation displayed in Figure 3 results in the stereospecific resonance assignment of the aminoethoxy protons. (A) NOEs involving H7′ and H7″ resonances and H1′ resonances obtained at a mixing time of 200 ms. (B) Very strong NOEs are found between H6″ and H1′ at a mixing time of 50 ms. (C) Seven antiphase quartets are indicated with boxes in the double-quantum filtered COSY. Dotted lines drawn in the plots indicate that these cross-peaks origin exclusively from downfield shifted C6′-bound protons and upfield shifted C7′-bound protons. The assignment which follows is indicated (see text).

High-Resolution Structure of the Triplex. After collecting all NMR restraints for this aminoethoxy-modified triplex, for which the most relevant ones are described above, its structure was calculated using distance geometry. To this end, 270 NOE derived restraints, 49 hydrogen bond restraints, and 114 dihedral restraints were used (on average 21 experimental restraints per residue). No attempts were made to derive the structure of the two loops. Details of the structure calculation protocol can be found in the Materials and Methods.

Figure 5 shows a superposition of 20 structures obtained by distance geometry calculations. The structure of the triplex is well-defined. The average pairwise RMSD for the triplex amounts 1.35 ± 0.38. The average pairwise RMSD of the heavy atoms in the central three triples amounts 0.60 ± 0.18. The triples are stacked almost perpendicular to the helical axis, and the *A-type* phosphate backbone of the third strand including *N-type* sugars is obviously well equipped for the formation of seven stable triples which are stacked contiguously in a triple helix. The aminoethoxy side chains are well-defined, and it clearly can be seen that each amino group is pointing to the phosphate of the previous residue in the second strand. In fact, the amino protons are within hydrogen-bonding distance across the Crick−Hoogsteen groove with the O1 oxygens (or Pro-R oxygens) of the phosphates in the purine strand. The two free oxygens at phosphorus are defined O1 and O2 by a clockwise counting looking along O3′−P−O5′ (*49*).

Dynamical Aspects. It is worthwhile to investigate possible dynamics at the aminoethoxy with the current spectral and structural data at hand. When the torsion angle O2′−C6−C7−N is switched to a gauche$^-$ conformation, the amino group sticks into solution, but when it is changed to a trans conformation, interestingly, the amino group from residue *i* comes close to the O1 oxygen of the phosphate of the neighboring residue, i.e., that belonging to triple *i* (the triples are numbered consequently with the 5′-3′ direction of the second and third strand). It is thus justified to speculate the presence of an equilibrium involving the trans and gauche$^+$ conformation of above-mentioned torsion angle resulting in the occurrence of alternate interactions of the amino group and the two phosphates. This has been shown to occur occasionally during molecular dynamics simulations of a triplex model (F. Casset and R. M. Wolf, unpublished results). When the cross-peaks in the double quantum-filtered COSY spectrum are inspected (cf. Figures 3 and 4C), it is obvious that the gauche$^+$ conformation is dominant for all aminoethoxy chains. The presence of a certain fraction of the trans conformation would be visible by the detection of two additional antiphase multiplets per aminoethoxy which are expected at H6″−H7′ and H6′−H7″. In the case of residues 2−6, the C6′ and C7′ protons are completely assigned and these additional cross-peaks are not observed. On the basis of simulations using the programs SPHINX and LINSHA, it is estimated that the population of the gauche$^+$ conformation is at least 80%. Within the limitation with respect to sensitivity, it can be concluded that the amino group interacts exclusively with the phosphate of the previous base triple $i - 1$ in the second strand (vide infra, Figure 7). Casual interaction of this amino group with the phosphate at position *i* cannot be detected. We note that, for residues 1 and 7, the C7′-bound proton resonances are degenerate.

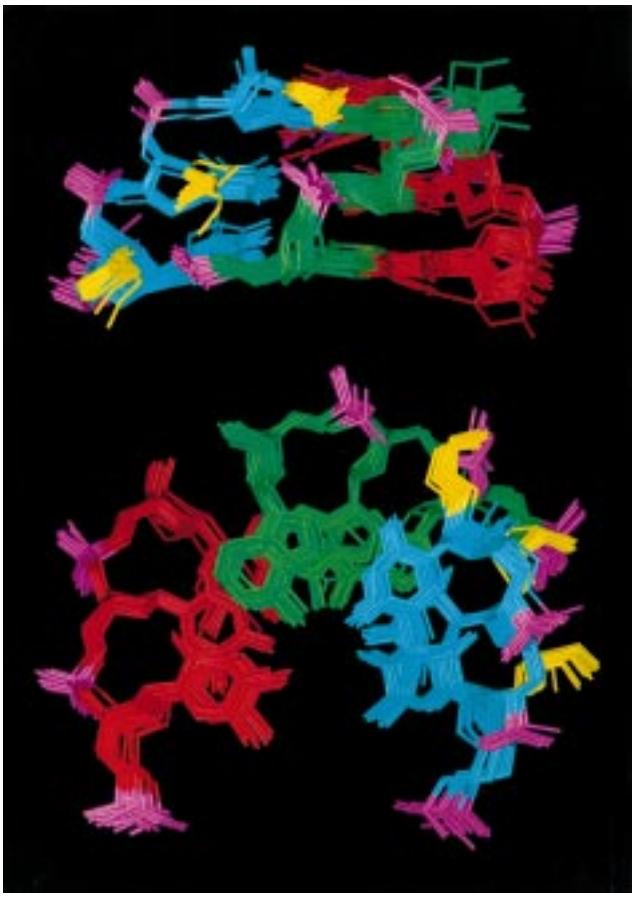


FIGURE 5: Superposition of 20 structures obtained by distance geometry calculations displayed perpendicular and along the helical axis. For clarity, only the residues involved in the three central triples are shown. Protons are not displayed. The strands are colored red, green, and cyan for strands 1, 2 and 3, respectively. The phosphates are colored purple, and the aminoethoxy substituent is colored yellow.

For residue 1, the NOE involving H1′ and H6″ is relative weak. Thus, the aminoethoxy moieties at the 5′-end and 3′-end of the modified strand are less rigid.

In conclusion, the interaction of aminoethoxy-modified oligonucleotides with double-stranded DNA forms a triple helical complex involving the formation of Hoogsteen base pair interactions in the major groove and, in addition, well-defined charge interactions between the aminoethoxy side chains and the phosphates of the purine strand.

DISCUSSION

Homology with Other Triplexes. Comparison between published Y·RY DNA triplexes and our 2′-aminoethoxy-modified rY·dRdY is particularly interesting, since this structure is one of the first studied which can be classified in the family of ribo·deoxyribo·deoxyribo triple helices. All the sugars in the third strand carry an oxygen atom at their 2′-position, and therefore, should adopt an *N-type* sugar conformation. It has been reported that this class of helices has an enhanced stability in comparison with DNA triplexes (*17, 50*−*52, 56*) and that the third strand has little influence on the structure of the Watson−Crick part of the triplex. Comparative analysis of four global helical parameters derived from our structure with those obtained from published triplexes, as well as with the values for *A-type* and *B-type* DNA (cf. Table 2), reveals that the helical parameters for the 2′-aminoethoxy modified rY·dRdY triplex fall surprisingly well within the same range of values obtained for DNA triplexes. Thus, the change of sugar conformation in the third strand has little effect on the relative position of the base triples in the triplex. The conformation of the Watson−Crick duplex is somewhat unwound, resulting in a reduced relative twist and in an increase in displacement from the helical axis, reaching an intermediate value between *A-* and *B-type* DNA, similar to the one reported for other triplexes.

Table 2: Helix Parameters of Current and Published Triplex Structures

triplex	axial rise (deg)	helical twist (deg)	residues/ turn	*x*-displacement (Å)
rY·dRdY triplex (current work)	3.5	30	12.1	−2.5
rY·dRdY triplex[a]	3.2	31	11.6	−2.5
dY·dRdY triplex[b]	3.2	35	10.3	−3.0
dY·dRdY triplex[c]	3.6	31	11.7	−1.9
dY·dRdY triplex[d]	3.4	31	11.5	−2.4
dY·dRdY triplex with T.CG[e]	3.4	32	11.3	−2.1
dY·dRdY triplex with G.TA[f]	3.4	31	11.6	−1.9
dY·dRdY triplex with G7.GC[g]	3.1	29	12.4	−2.9
dY·dRdY triplex with D3.TA[h]	3.2	31	11.6	−1.4
dY·dRdY triplex with PU.TA[i]	3.3	30	12.1	−1.6
dY·dRdY triplex[j]	3.3	28	12.9	na[l]
dY·dRdY triplex[j]	3.2	30	12.1	na
A-DNA[k]	2.6	33	10.9	−5.4
B-DNA[k]	3.4	36	10.0	−0.7

[a] Ref 17. [b] Ref 12. [c] Ref 54. [d] Ref 13. [e] Ref 16. [f] Ref 15. [g] Ref 20. [h] Ref 22. [i] Ref 21. [j] Ref 53. [k] Ref 49. [l] Not available.

Table 3: Groove Widths of Triplexes

triplex	Watson−Crick groove[a] (Å)	Crick−Hoogsteen groove (Å)	Watson−Hoogsteen groove (Å)	Watson−Crick major groove (Å)
rY·dRdY triplex	12[b]	8−9	16	21
dY·dRdY triplex[b]	13	8−9	13	20
dY·dRdY triplex[c]	16	7.5	9	16
A-DNA[d]	17			8
B-DNA[d]	12			17

[a] The Watson−Crick groove denotes the groove between strands 1 and 2. It resembles the minor groove in DNA; the Crick−Hoogsteen groove is the groove between strands 2 and 3; the Watson−Hoogsteen groove is formed by strands 1 and 3.The groove dimension is the average value of the shortest distance between phosphorus atoms of two different strands. [b] Ref *15* (NMR). [c] Ref *53* (fiber diffraction).[d] Ref *49*.

Another global parameter which may be considered is the size of the various grooves. These values, which are the average of the shortest interstrand phosphate−phosphate distances, are collected in Table 3. Comparison of these values with those measured for a Y·RY triplex shows that the groove dimensions are well conserved. It is especially interesting to note that the dimension of the narrowest groove, the Crick−Hoogsteen groove, is conserved. In this groove, the structure could have been perturbed due to the presence of the aminoethoxy moiety and its interaction with the second strand phosphates. Obviously, no perturbation is observed and the dimension of this groove seems to be optimal to accommodate the aminoethoxy moiety, with respect to both bonded and nonbonded interactions. The groove dimensions have also been compared with the dimensions measured for

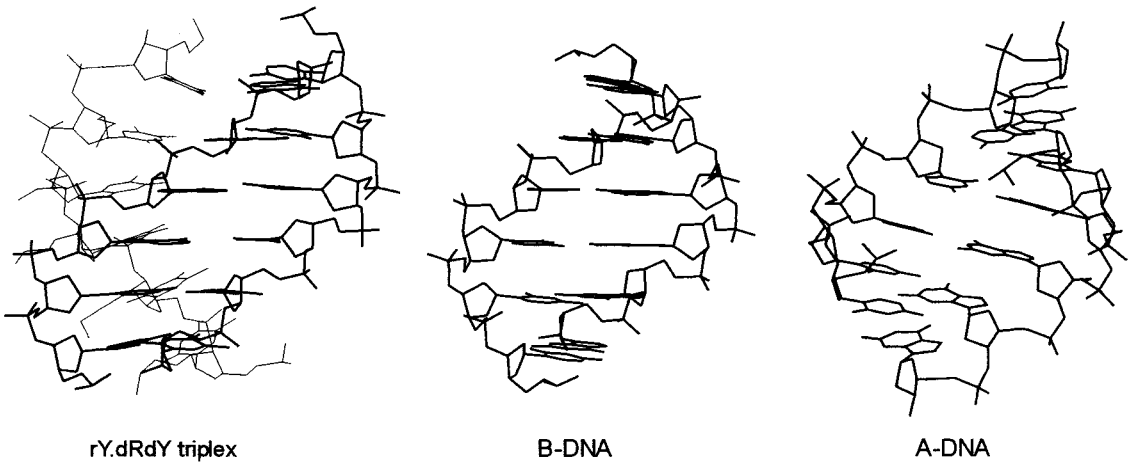


FIGURE 6: Visual comparison of the structure of the 2′-aminoethoxy modified rY·dRdY triplex and the structure of regular B-DNA and A-DNA. The third strand drawn with thin lines is close to an *A-type* helix. The Watson−Crick part resembles much more a *B-type* helix. It is, however, unwound with respect to B-DNA in order to accommodate the third strand.

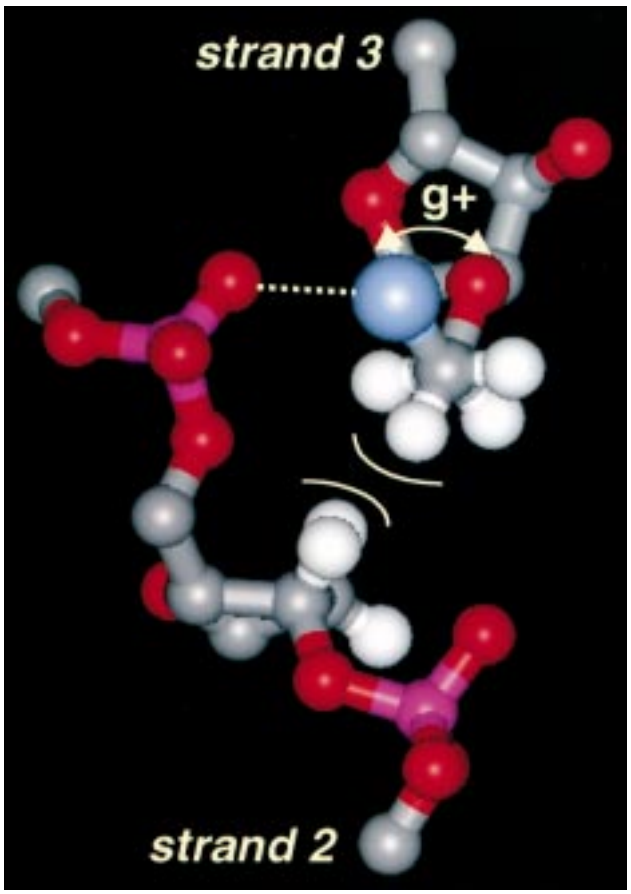


FIGURE 7: Newman projection of the 2′-aminoethoxy side chain in a gauche$^+$ conformation. Only parts of the second and third strands are shown for clarity. The charged amino moiety (blue) interacts with one pro-R oxygen of the second strand DNA phosphodiester backbone [dotted line, 2.8 (Å)]. The H6″ and H7′ alkyl protons of the aminoethoxy side chain (white) are within van der Waals contact [2.3−2.5 (Å)] of the H2′ and H3′ protons of the second strand ribose (white), respectively.

the *A-type* triplex and *A-type* and *B-type* DNA duplexes (in Table 3). The dimension of the Watson−Crick minor groove is unchanged with respect to B-DNA. When we compare the Watson−Crick major groove width, one can clearly see that the DNA duplex has increased its major groove from 17 to 21 Å in order to accommodate the third strand, but remains certainly closer to B-DNA than to A-DNA. The groove dimensions differ significantly from the *A-type* triple helical structure determined by fiber diffraction.

Finally, when the Watson−Crick part of the triplex is viewed along the helical axis (Figure 6), it resembles much more to the structure of unwound B-DNA than that of A-DNA, even though the third strand is close to an A-type helix.

A triplex in which the third strand is RNA has been studied by Hilbers and co-workers (*56*). These investigators observed significant interstrand NOEs involving the H1′ and H8 resonance, corresponding to a distance of 3.2−3.5 Å (C. W. Hilbers, personal communication). As noted before, we also observed similar strong interstrand NOEs (H1′−H8) resonance that corresponds to a distance of 3.4−3.8 Å, whereas this distance is 4.5−5.0 Å in DNA triplexes. Therefore, these results support our classification of the currently studied triplex structure to the rY·dRdY family. Because 2′-aminoethoxy side chains make such specific interactions in the Crick−Hoogsteen groove, it is interesting to compare in detail this structure with a true rY·dRdY triplex structure. Very recently, such a structure became available (*17*), and as can be appreciated from the comparison of the helical parameters (Table 2), these match quite well with those of the aminoethoxy-modified triplex. The sugar conformations in the third strand of our structure are more restricted to N-type, which is explained by the presence of the aminoethoxy substituent. The minor groove width in the modified triplex is also reduced with respect to Y·RY triplexes but not to the same extent (12 vs 11 Å).

Differences with Other Triplexes. In all DNA triplexes which have been characterized by NMR so far, the sugar moieties adopt predominately an *S-type* conformation. In the currently investigated modified nucleotides, the electronegative substituent at the 2′ position shifts the equilibrium to *N-type*. Indeed, when the aminoethoxy-modified residue is incorporated in a dinucleotide (M.J.J.B. and F.N., unpublished results), the equilibrium of *N-type* and *S-type* conformation shifts to *N-type* (60% *N-type*) relative to the natural dinucleotide (80% *S-type*). In the triplex, these modified sugars, incorporated into the third strand, shift completely to a pure *N-type* conformation (100%).

The change of sugar pucker in the third strand to *N-type* has important structural consequences for the third strand,

although the other two strands do not seem to be perturbed (see previous section). One of these changes is the intrastrand phosphate distance. When the sugar pucker changes from *S-type* to *N-type* in, e.g., *B-type* and *A-type* polynucleotide helices, the distance between subsequent phosphorus atoms decreases from 7.0 to 5.9 Å. (*57*). Since repulsion between the negatively charged phosphates contributes to the conformational energy, ribose sugars adopt sometimes *S-type* sugar conformations, e.g., in hairpin loops (*58, 59*). But in the case of the 2′-aminoethoxy modified rY·dRdY triplex, the distance between subsequent phosphates in the third strand is small (6.25 Å) compared to the value measured for the third stand of a DNA triplex (7.0 Å), or to the values of the first (6.7 Å) and second strand (6.85 Å) of the DNA duplex.

Among other differences intimately connected with the sugar conformation is the position of the base relative to the sugar, e.g., in *A-* and *B-type* polynucleotide helices, the glycosidic torsion angle is 206 and 260°, respectively. In the 2′-aminoethoxy modified rY·dRdY triplex, an average value of 205° is measured in the third strand. This value is in contrast with 235° measured in Y·RY triplexes. Moreover, the topology of the grooves is somewhat different, although the groove dimensions between the two classes of triplexes are equal. The *N-type* sugar residues protrude more into the Crick−Hoogsteen groove than is the case for *S-type* sugars.

In conclusion, the structures of the 2′-aminoethoxy-modified rY·dRdY triplex and Y·RY triplex have a similar topology of base triples and groove dimensions. The structure of the first and second strand is, in both cases, best described as an unwound *B-type* double helix for which the major groove dimension is increased 3−4 Å. Most of the differences are found in the third strand conformation: the major groove accommodates an *A-type* rY strand in the 2′-aminoethoxy-modified rY·dRdY triplex, whereas it accommodates a *B-type* Y strand in the Y·RY triplex.

Conformational Aspects of the Aminoethoxy Modification. To determine accurately the conformation of the aminoethoxy side chain, we introduced selectively two ^{13}C labels in one particular aminoethoxy side chain of the triple helical complex. This label allows for the recording of NMR spectra of the triplex with a significantly reduced complexity (cf. Figure 2). Only in these ^{13}C-edited spectra was the identification of crucial NOEs such as the $T_4H7'-A_{27}H3'$ NOE possible. The introduction of selective ^{13}C labels is relatively straightforward if the oligonucleotide can be chemically synthesized, although a similar approach has recently been described using enzymatic synthesis (*60*).

It is likely that the so-called *gauche effect* (*61*) plays a significant role in determining the preferred conformation of the aminoethoxy side chain. This effect directs the torsion angles of the type X−C−C−Y, in which X and Y are electronegative substituents, into a gauche$^+$ or gauche$^-$ conformation. Using the molecular mechanics program Discover in conjunction with the AMBER force field (*36*), we calculated that the 2′-aminoethoxy in a trans conformation is less favorable by about 1.5 kcal/mol, compared to the gauche conformations. Indeed, in the triplex structure, the torsion angle O2′−C6−C7−N is gauche$^+$. The conformation of the aminoethoxy which fulfills the gauche rule brings the amino group close to the phosphate of the second strand (Figure 7). This enables a hydrogen bond interaction between the amino proton of the aminoethoxy side chain and the O1 oxygen of the phosphate. Thus, the aminoethoxy group links the residues of triple (i) in the third strand crossing the Hoogsteen−Crick groove to the phosphate belonging to the $i - 1$ triple. As mentioned previously, the size of the Crick−Hoogsteen groove is unaltered with respect to the dY·dRdY triplex. This groove size is obviously perfect to accommodate the aminoethoxy group in a gauche$^+$ conformation and is optimal for electrostatic interaction of the amino group with the phosphate across the narrow groove. In addition, interesting van der Waals interactions between the aminoethoxy protons of C6−C7 and those of the C2′−C3′ second strand ribose are present (Figure 7) and could also contribute to the stability of the complex.

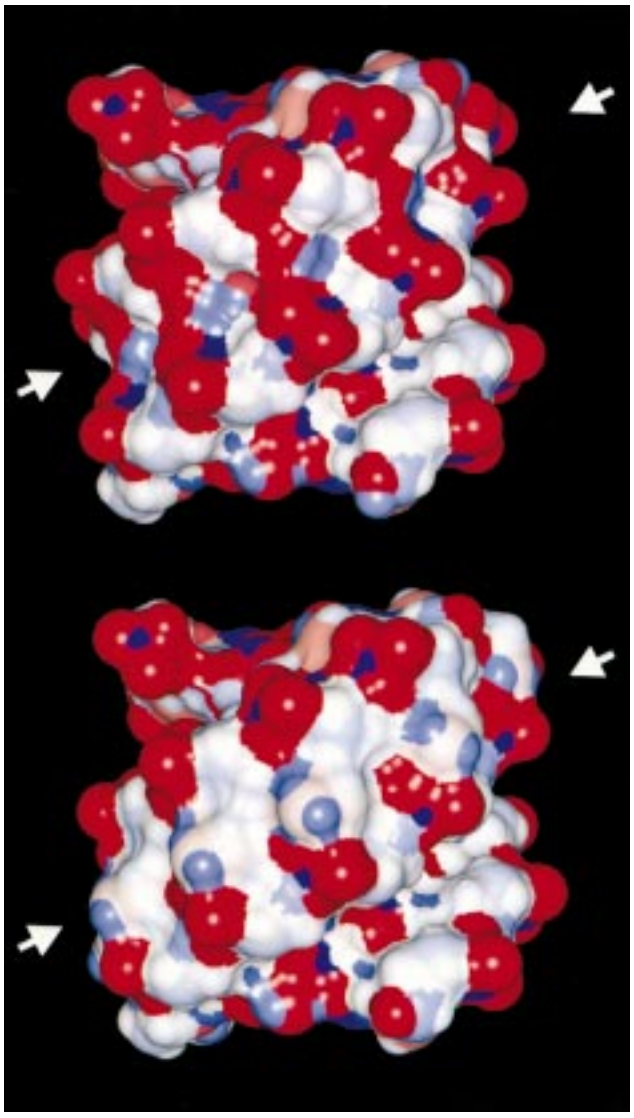


FIGURE 8: Surface rendered model of the 2′-aminoethoxy modified rY.dRdY triplex viewed into the Hoogsteen−Crick groove. The position of this groove is indicated with arrows. The distribution of the partial charges is shown via color coding: intense blue ≥ +0.5, intense red ≤ −0.5, white ≈ 0. A model of the triplex in which the aminoethoxy substituents have been replaced by hydroxyl groups is displayed on top. The experimental model of the triplex carrying aminoethoxy substituents is displayed below.

Figure 8 shows a comparison of the Crick−Hoogsteen groove of the currently studied triplex with that of a model of the rY.dYdR triplex, which is obtained after replacing the aminoethoxy substituents by hydroxyl groups. It is very

clear from this representation that the modification fills most of the space in the groove and that the positively charged amino group compensates for the negative charge of the second strand. Moreover, it is surrounded by the negative charge density of the bottom and the wells of the groove. It has been found by NMR that long residence water molecules are present in the Crick−Hoogsteen and Watson−Hoogsteen grooves of DNA triplexes (*62*). These water molecules or hydrated positively charged ions have possibly a similar role to the charged aminoethoxy side chain and screen unfavorable electrostatic interaction between the phosphates of strands 2 and 3.

Correlation with Biophysical Results. Under physiological conditions, the binding of unmodified oligodeoxyribonucleotide to DNA duplex is about 1000-fold slower compared to the formation of a Watson−Crick duplex (*63, 64*). This effect has been attributed in part to the strong repulsion between the negatively charged phosphates of the third strand and those of the DNA duplex and possibly to a conformational change of the DNA upon triplex formation. Nevertheless, the binding mechanism for triple strand formation (*62, 63*) appears to be similar to the nucleation-zipping mechanism developed for double strand formation (*64, 65*). A nucleation step involves the formation of three to five Hoogsteen base pairs and is followed by a zipping step, where the rest of the oligonucleotide wraps around the DNA duplex. 2′-Aminoethoxy-modified oligonucleotides have dramatic effects on the kinetic and thermodynamic binding parameters (*9*). A 15-mer 2′-aminoethoxy-modified oligonucleotide has been shown to have a 1000-fold increase in association rate constant for its target DNA when compared to an unmodified oligonucleotide of the same sequence and provides an overall gain in free energy of 6.3 kcal/mol.

Consistent with the solution structure, the origin of the faster association rate for triplex formation is likely to arise from two effects. Because the zwitterionic 15-mer oligonucleotide reported in our study (*9*) has an overall charge of zero (14 phosphates and 14 protonated aminoethoxy side chains), it should result in a global decrease in phosphate−phosphate electrostatic repulsion. Supporting this view, we found that several other chemical modifications that decrease the overall negative charge of the third strand also provide some increase in association rate, but to a lesser extent (B. Cuenoud, unpublished data). More importantly, the unique ability of the aminoethoxy side chains to contact specifically the phosphates of the DNA target upon binding is likely to influence directly both the nucleation and the zipping process, resulting in an extremely large enhancement in association rate constant.

The solution structure provides also a rational for the structure-affinity relationships observed with different substituents at the 2′ position (*9*). The orientation of the aminoethoxy side chain, in a preorganized low-energy gauche conformation, enables an optimal interaction between the protonated amino group and the phosphate. 2′-Aminopropoxy-modified oligonucleotide displays a lower thermal stability (ΔT_m/modification = +2.1°), a result that can be explained by both an increase in side-chain flexibility, due to the lack of a gauche effect and an unfavorable spatial geometry for optimum ion-pair interaction. One interesting finding is the ability of the 2′-hydroxyethoxy side chain to provide a strong thermal stability for triplex binding (ΔT_m/modification = +1.9°). In this case, a decrease in charge repulsion cannot be invoked. The 2′-hydroxyethoxy side chain is likely to adopt a gauche$^+$ conformation, with the hydroxyl interacting with the same phosphate as the 2′-aminoethoxy. Most of the stability difference provided by the two side chains is likely to arise from the nature of this contact, an ion-pair interaction being stronger than an uncharged hydrogen bond donor contacting a negatively charge phosphate. In protein−DNA complexes, uncharged contacts between hydrogen bond donors and DNA phosphates play also an important role for site specific recognition, and often involve polar side chains or the NH amides of the protein backbone (*7, 8*).

CONCLUSION

This work unambiguously demonstrates that, as previously predicted, 2′-aminoethoxy-modified oligonucleotides contact both the phosphodiester backbone and the bases at each base-pair steps of the DNA duplex. In the triple helical complex, the 2′-aminoethoxy side chain lies in a perfect conformation for optimum interaction with the pro-R oxygen of the phosphodiester belonging the DNA second strand. Moreover, this work represents one of the first detailed structural studies of an rY·dRdY type triplex. Remarkably, the A-type conformation of the third strand does not influence significantly the unwounded B-type conformation of the DNA duplex, also found in Y·RY triplexes. The unique biophysical properties of 2′-aminoethoxy-modified oligonucleotides combined with a very large nuclease resistance provide promising compounds for the regulation of gene expression.

ACKNOWLEDGMENT

We thank Arlette Garnier and Veronique Drephal for excellent technical assistance, and Christian Guenat for mass spectroscopy measurements.

REFERENCES

1. Nielsen P. E. (1997) *Chem. Eur. 3*, 505−508.
2. Trauger, J. W., Baird, E. E., and Dervan, P. B. (1996) *Nature 382*, 559−561.
3. White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., and Dervan, P. B. (1998) *Nature 391*, 468−471.
4. Neidel. S. (1997) *Anticancer Drug Res. 12*, 433−442.
5. Soyfwer V. N., and Potaman V. N. (1995) *Triple-Helical Nucleic Acids*, 1st ed., Springer, New York.
6. Giovannangeli C., and Hélène C. (1997) *Antisense Nucleic Acid Drug Dev. 7*, 413−421.
7. Pabo, C. O., and Sauer R. T. (1992) *Annu. Rev. Biochem. 61*, 1053−1095;
8. Rhodes, D., Schwabe, J. W. R., Chapman, L., and Fairall, L. (1996) *Philos. Trans. R. Soc. London B 351*, 501−509.
9. Cuenoud, B., Casset, F., Huesken, D., Natt, F., Wolf, R. M., Altman, K.-H., Martin, P., and Moser, H. E. (1998) *Angew. Chem. Int. Ed. 37*, 1288−1291.
10. Radhakrishnan, I., and Patel D. J. (1994) *Biochemistry 33*, 11405−11416.
11. Macaya, R., Wang, E., Schultze, P., Sklenar, V., and Feigon, J. (1992) *J. Mol. Biol. 225*, 755−773.
12. Bartley, J. P., Brown, T., and Lane, A. N. (1997) *Biochemistry 36*, 14502− 14511.
13. Tarköy, M., Phipps, A. K., Schultze, P., and Feigon, J. (1998) *Biochemistry 37*, 5810−5819.
14. Radhakrishnan, I., Patel, D. J., Veal, J. M., and Gao, X. (1992)*J. Am. Chem. Soc. 114*, 6913−6915.
15. Radhakrishnan, I., and Patel, D. J. (1994) *Structure 2*, 17−32.

16. Radhakrishnan, I. and Patel, D. J. (1994) *J. Mol. Biol. 241*, 600−619.
17. Gotfredsen, C. H., Schultze, P., and Feigon, J. (1998) *J. Am. Chem. Soc. 120*, 4281−4289.
18. Radhakrishnan, I., and Patel, D. J. (1993) *J. Am. Chem. Soc. 115*, 1615−1617.
19. Radhakrishnan, I., and Patel, D. J. (1993) *Structure 1*, 135−152.
20. Koshlap, K. M., Schultze, P., Brunar, H., Dervan, P. B., and Feigon, J. (1997) *Biochemistry 36*, 2659−2668.
21. Phipps, A. K., Tarköy, M., Schultze, P., and Feigon, J. (1998) *Biochemistry 37*, 5820−5830.
22. Wang, E., Koshlap, K. M., Gillespie, P., Dervan, P. B., and Feigon, J. (1996) *J. Mol. Biol. 257*, 1052−1069.
23. Chu, S. H., Chen, Z. H., Weng, Z. Y., Rowe, E. C., Chu, E., and Chu, M.-Y. W. (1987) *J. Heterocycl. Chem. 24*, 989−995.
24. Ross, B. S., Springer, R. H., Tortorici, Z., and Dimock, S. (1997) *Nucleosides Nucleotides 16*, 1641−1643.
25. Jahnke, W. (1996) *J. Magn. Reson., Ser. B 113*, 262−266.
26. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
27. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546−4553.
28. Piantini, U. Sørenson, O. W., and Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800−6801.
29. Braunschweiler, L., and Ernst, R. R. (1983) *J. Magn. Res. 53*, 521−558.
30. Griesinger, C., Otting, G., Wüthrich, K., and Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870−7872.
31. States, D. J., Haberkorn, R. A., and Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286−292.
32. Schleucher, J., Schwendinger, M., Sattler, M., Schmidt, P., Glaser, S. J., Sørensen, O. W., and Griesinger, C. (1994) *J. Biomol. NMR 4*, 301−306.
33. Widmer, H., and Wüthrich, K. (1986) *J. Magn. Reson. 70*, 270−279.
34. Havel, T. F. (1991) *Prog. Biophys. Mol. Biol. 56*, 43−78.
35. Blommers, M. J. J., van de Ven, F. J. M., Van der Marel, G. A., van Boom, J. H., and Hilbers, C. W. (1991) *Eur. J. Biochem. 201*, 33−51.
36. Weiner, S. J., Kollman, P. A., Ngygen, D. T., and Case, D. A. (1986) *J. Comput. Chem. 7*, 230−235.
37. Xodo, L. E., Manzini, G., Quadrifoglio, F., van der Marel, G. A., and van Boom, J. H. (1991) *Nucleic Acids Res. 19*, 5625−5631.
38. De los Santos, C., Rosen, M., and Patel, D. J. (1989) *Biochemistry 28*, 7282−7289.
39. Rajagopal, P., and. Feigon, J. (1989) *Nature 339*, 637−640.
40. Mooren, M. M. W., Pulleyblank, D. E., Wijmenga, S. S., Blommers, M. J. J., and Hilbers, C. W. (1990) *Nucleic Acids Res. 18*, 6523−6529.
41. Wijmenga, S. S., Mooren, M. M. W., and Hilbers, C. W. (1993) in *NMR of Macromolecules. A practical Approach* (Roberts, G. C. K., Ed.) pp 217−288, IRL Press, Oxford.
42. Feigon, J., Koshlap, K. M., and Smith, F. W. (1995) *Methods Enzymol. 261*, 225−255.
43. Van de Ven, F. J. M., and Hilbers, C. W. (1988) *Nucleic Acids Res. 16*, 5713−5726.
44. Wijmenga, S. S., Kruithof, M., and Hilbers, C. W. (1998) *J. Biomol. NMR 10*, 337−350.
45. Roongta, V. A., Jones, C. R., and Gorenstein, D. G. (1990) *Biochemisty 29*, 5245−5258.
46. Blommers, M. J. J., Nanz, D., and Zerbe, O. (1994) *J. Biomol. NMR 4*, 595−601.
47. Wüthrich, K. (1986) *NMR of proteins and nucleic acids*, 1st ed., John Wiley & Sons, New York.
48. Haasnoot, C. A. G., de Leeuw, F. A. A. M., and Altona, C. (1983) *Tetrahedron 36*, 2783−2792.
49. Sänger, W. (1984) *Principles of nucleic acid structure*, 1st ed., Springer Verlag, New York.
50. Roberts, R. W., and Crothers, D. M. (1992) *Science 258*, 1463−1465.
51. Han, H., and Dervan, P. B. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 3806−3810.
52. Escudé, C., François, J.-H., Sun, J.-S., Ott, G., Sprinzl, M., Garestier, T., and Hélène, C. (1993) *Nucleic Acids Res. 21*, 5547−5553.
53. Arnott, S., and Selsing, E. (1974) *J. Mol. Biol. 88*, 509−521.
54. van Dongen, M. J. P. (1998) Thesis Nijmegen, The Netherlands.
55. Ji, J., Hogan, M. E., and Gao, X. (1996) *Structure 4*, 425−435.
56. van Dongen, M. J. P., Heus, H. A., Wijmega, S. S., van der Marel, G. A., van Boom, J. H., and Hilbers, C. W. (1996) *Biochemistry 35*, 1733−1739.
57. Sundaralingam, M. (1974) *Structure and Conformation of Nucleic Acids and Protein-Nucleic Acid Interactions* (Sundaralingam, M., and Rao, S. T., Eds.) pp 487−524, University Park Press, Baltimore.
58. Heus, H., and Pardi, A. (1991) *Science 253*, 191−193.
59. Varani, G., Cheong, C., and Tinoco, I. (1991) *Biochemistry 30*, 3280−3289.
60. Mer, G., and Chazin, W. J. (1998) *J. Am. Chem. Soc. 120*, 607−608.
61. Wiberg, K. B. (1996) *Acc. Chem. Res. 29*, 229.
62. Radhakrishnan, I., and Patel, D. J. (1994) *Structure 2*, 395−405.
63. Rougée, M., Faucon, B., Mergny, J.-L., Barcelo, F., Giovannangeli, C., Garestier, T.; Hélène C. (1992) *Biochemistry 31*, 9269−9278.
64. Xodo, L. E. (1995) *Eur. J. Biochem. 228*, 918−926.
65. Pörschke, D. and Eigen, M. (1971) *J. Mol. Biol. 62*, 361−381.
66. Craig, M. E., Crothers, D. M., and Doty P. (1971) *J. Mol. Biol. 62*, 383−401.

BI9816352