Structural Study of a DNA·RNA Hybrid Duplex with a Chiral Phosphorothioate Moiety by NMR: Extraction of Distance and Torsion Angle Constraints and Imino Proton Exchange Rates†

Carlos González,‡ Wojciech Stec,§ Anna Kobylanska,§ Richard I. Hogrefe,∥ Mark Reynolds,∥ and Thomas L. James*,‡

Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143-0446

Received May 10, 1994; Revised Manuscript Received July 18, 1994®

ABSTRACT: The solution structure of the thiophosphate-modified DNA·RNA hybrid duplex d(GCTATAA$_{ps}$-TGG)·r(CCAUUAUAGC) has been studied by NMR. Two samples with pure stereochemistry in the modified phosphate have been investigated. Two-dimensional NMR (2D NMR) methods have been applied to assign nearly all the resonances in both duplexes. Scalar coupling constants have been determined by comparing quantitative simulations with experimental double-quantum filtered COSY (DQF-COSY) cross-peaks. More than 300 distance constraints have been obtained from two-dimensional nuclear Overhauser spectroscopy (2D NOE) spectra recorded in D_2O and H_2O by using a complete relaxation matrix analysis as implemented in the program MARDIGRAS. This hybrid duplex presents a heteronomous structure. Riboses in the RNA strand are found in a N-type conformation typical of the A-form family as shown by the lack of H1′–H2′ cross-peaks in DQF-COSY spectra and confirmed by the measured interproton distances. In contrast, the DNA strand adopts a different conformation with sugar puckers partially in the *S*-type domain, which is not in agreement with the A-family of structures. Coupling constants in deoxyriboses are not consistent with any single sugar conformation. Therefore, sugar pucker pseudorotation parameters are calculated according to a two-state dynamic equilibrium between N- and *S*-type conformers. In general, the population of major *S* conformer is lower than in double-stranded DNA duplexes, indicating that hybrid duplexes may be more flexible than pure DNA or RNA. The only differences observed in the spectra between the two stereoisomers studied originate from resonances of protons located near the modified phosphate. No significant differences in interproton distance have been detected, and only a slight difference of sugar pucker in the 5′ neighbor has been found. The sulfur atom appears to be well-accommodated without further changes in the structure of the hybrid.

DNA·RNA hybrid duplexes are found in many biological processes such as intermediate steps in transcription or primers for DNA replication via Okazaki fragments. They also play a role in reverse transcription from RNA into DNA during retroviral replication. Hybrid duplexes are substrates for ribonuclease H, which cleaves the RNA strand in DNA·RNA duplexes, but it is inactive toward double-stranded RNA. This enzyme is also part of the reverse transcriptase and is involved in the elimination of the viral RNA from the DNA·RNA duplex during the reverse transcription process (Varmus, 1988). The crystal structure of ribonuclease H is known (Katayanagi et al., 1990; Yang et al., 1990) and has been used to model the enzyme/hybrid duplex complex. The first modeling studies were carried out assuming an A-form structure for the hybrid duplex (Yang et al., 1990; Nakamura et al., 1992). Only very recently, a solution structure of a hybrid has been used to explain the mechanism of action of this enzyme (Fedoroff et al., 1993).

On the other hand, modified oligonucleotides have attracted considerable attention because of their potential use as chemotherapeutic agents. Among the many types of oligonucleotide analogs, phosphorothioates are especially interesting due to their enhanced nuclease resistance and ability to permeate cell membranes. Phosphorothioate-modified DNA forms hybrid duplexes with RNA that are very susceptible to ribonuclease H digestion. The inhibitory activity of antisense oligonucleotides is related to their ability to form hybrid duplexes with target RNA that can be attacked by ribonuclease H (Zon, 1988; Uhlmann & Peyman, 1990). Knowledge of the solution structure of DNA·RNA hybrids as well as their modified analogs is of primary importance to understand their biological activity and to develop new modified oligonucleotides with enhanced activity.

Different studies have been carried out on hybrid duplexes in solid state. In the X-ray structure of a chimeric DNA·RNA strand paired with its complementary DNA, both strands adopt a pure A-form conformation (Wang et al., 1982; Egli et al., 1992). Studies on poly(dA)·poly(rU) and poly(dI)·poly(rC) carried out by X-ray fiber diffraction concluded that both duplexes have an heteronomous structure, where the DNA strand adopts a B-form conformation, with sugars in C2′-endo, and the RNA strand assumes an A-form conformation, with sugars in C3′-endo (Arnott et al., 1986). Until very recently not very much was known about the structure of hybrid duplexes in solution (Hall, 1993). Several studies have shown that the sugar conformations in the RNA and DNA strands are different. Riboses adopt an N-type (C3′-endo) conformation typical of an A-form, whereas deoxyriboses in the DNA strand are closer to the *S*-type domain (C2′-endo) (Chou et al., 1989; Katahira et al., 1990; Salazar et al., 1993b;

† This work was supported by NIH Grants GM39247 and GM41639.
* Author to whom correspondence should be addressed. Telephone: (415) 476-1569. Fax: (415) 476-0688.
‡ University of California, San Francisco.
§ Department of Bioorganic Chemistry, Centre of Molecular and Macromolecular Studies, Polish Academy of Sciences, Łódź, Poland.
∥ Genta Inc., 3550 General Atomics Court, San Diego, CA 92121.
® Abstract published in *Advance ACS Abstracts,* September 1, 1994.

0006-2960/94/0433-11062$04.50/0

Lane et al., 1993). However, some controversy has arisen about the structure of the sugars in the DNA strands of hybrids and chimeric duplexes. Salazar et al. (1993a,b) conclude that deoxyriboses assume an unusual E-type conformation (approximately O1′-endo) with pseudorotation phase angle around 90°. On the other hand, Lane et al. (1993) found that their NMR data could be better explained by assuming a two-state approximation with a major population of sugars in *S*-type conformation.

In spite of the considerable number of studies concerning the chemical synthesis and biochemical properties of modified oligonucleotides, not much is known about the structural effects of these modifications. The few studies carried out to date by crystallographic methods focus on DNA·DNA duplexes (Cruse et al., 1986; Heinemann et al., 1991). In both cases, the modification appears to fit well in the general B-type structure of the duplex. A similar conclusion is reached by Gao et al. (1992) and by Stolarsky et al. (1992) in modified DNA·DNA duplexes investigated in solution by NMR. The lack of structural information is even more pronounced in the case of modified hybrids. Although mRNA is the most attractive target for the antisense chemotherapeutical strategy, virtually nothing is known about the structure of oligonucleotide analogs in DNA·RNA duplexes.

The aim of this work is to determine the structure of the fragment d(GCTATAA$_s$TGG)·r(CCAUUAUAGC) in solution by NMR, where the phosphate between A7 and T8 is modified by substituting one oxygen atom for sulfur. After this substitution, two different stereoisomers are possible (*R* and *S*). Two samples with different stereochemistry in the modified phosphate have been studied. The central part of this sequence is the Pribnow box, whose double-stranded DNA form has been studied previously in our laboratory (Schmitz et al., 1992). A pure DNA·RNA hybrid duplex with very similar sequence has been studied by Chou et al. (1989) by NMR.

In the present study, we apply the recent proton NMR methodology that has been successfully employed in our lab to study several DNA fragments (Weisz et al., 1992; Mujeeb et al., 1992; Bishop et al., 1994). Special care has been taken in determining accurate interproton distance constraints using a complete relaxation matrix analysis of the 2D NOE[1] intensities. Quantitative simulations of COSY spectra have been employed to calculate *J*-coupling values by comparing theoretical and experimental cross-peaks (Widmer & Wüthrich, 1986, 1987). While this study introduced a minor variation in the method, it has been previously used to determine sugar puckering in several DNA fragments (Celda et al., 1989; Gochin et al., 1990; Schmitz et al., 1990).

MATERIALS AND METHODS

Sample Preparation. Pure stereoisomers of the modified DNA strand were obtained in two steps. First, synthesis and separation of pure diastereoisomers of d[5′-DMT-A^{Bz}$_{P(S)OEt}$T] were carried out as described earlier (Lesser et al., 1992). ^{31}P NMR analysis had shown that diastereomeric purity of the *R*p isomer was >96%, while the purity of the *S*p isomer was >99%. Independently, samples of both isomers were exhaustively deprotected (25% aqueous NH_3 for 48 h, followed after evaporation by 20% acetic acid for 0.5 h) and analyzed by means of reverse-phase HPLC (PRP-1 column, 305 × 7 mm) with a linear gradient of 0–50% buffer B [40% CH_3CN in 0.1 M triethylammonium bicarbonate (TEAB), pH 7.2] eluting in 17 min. Integration of HPLC traces indicated that [*R*p]-d[ApsT] (RT 11.86 min) was >98%, while [*S*p]-d(ApsT) (RT 12.41 min) was 99% diastereomerically pure compound. Each diastereomer of d[5′-DMT-A^{Bz}$_{P(S)OEt}$T] (25 mg in 250 μL of dry CH_3CN) was phosphitylated at the 3′-oxygen with O-(2-cyanoethyl)-*N*,*N*,*N*′,*N*′-tetraisopropylphosphorodiamidite (30 μmol) in the presence of 5′-ethylthiotetrazole (Koziolkiewicz & Wilk, 1993) and used without isolation as substrate in the appropriate coupling step of automated oligonucleotide synthesis (3 μmol scale) on the 380B DNA synthesizer (Applied Biosystems) using *O*-β-cyanoethyl phosphoramidite monomers (Zon & Stec, 1992). The solid support was 5′-DMT-G^{IPA}LCA-CPG (Uznanski et al., 1989). 3′-*O*-Phosphitylated [*R*p]- and [*S*p]-d[A$_{P(S)OEt}$T], respectively, were introduced on the column following 5′-*O*-detritylation of d(GG) dimer anchored to solid support; the coupling time for this particular step was 10 min. After washing, capping, oxidation, and detritylation (DMT-cation assay in both cases was >95%), the synthesis of each diastereomeric decamer was continued with average efficiency of coupling steps, according to DMT assay, of >97%.

After cleavage from the solid support (25% aqueous NH_3, 2 h, 25 °C) and deprotection of bases (25% aqueous NH_3, 48 h, 50 °C), semipreparative HPLC purification of 5′-*O*-DMT-(*R*p) and -(*S*p) decamers was performed on PRP-1-column 305 × 7 mm with a linear gradient of 0–60% buffer B for 5 min, an increase of B to 75% in the following 2 min, and to 100% in next 5 min, and returning to 100% of buffer A in the 19th min; flow rate and buffers were as above (RT for both diastereomers, 9.86 min). After evaporation of collected fractions containing the desired oligomers, the 5′-O-DMT groups were removed with 20% aqueous acetic acid.

The second HPLC purification was performed on the same column with a linear gradient of 0–60% of buffer B (3%/min), then returned to 100% of buffer A within the following 3 min. Peaks of the desired oligonucleotides (RT 16.77 min) were collected and solvents removed under reduced pressure. Solid residues were dissolved in water with subsequent cation exchange on a Dowex ion exchange column (Na^+ form). The yield of *R*p-labeled diastereoisomer was 150.5 A_{260}OD units, while that of the *S*p-diastereoisomer was 152.3 A_{260}OD units.

The presence and diastereomeric purity of the phosphorothioate moiety between A and T in each oligonucleotide were proved by digestion with snake venom phosphodiesterase (EC 3.1.4.1) or with nuclease P1 (EC 3.1.30.1), which are known to be stereoselective (Bryant & Benkovic, 1979; Potte et al., 1983) toward P chiral phosphorothioate dinucleotides. For nuclease P1 digestion, each oligonucleotide was dissolved in 200 μL of 100 mM Tris-HCl, pH 7.2, and 1 mM $ZnCl_2$ and received 1 μg of nuclease P1. For snake venom phosphodiesterase digestion, each oligomer was dissolved in 200 μL of 100 mM Tris-HCl, pH 8.5, and 15 mM $MgCl_2$ and received 10 μg of phosphodiesterase. The digestion mixtures were incubated for 12 h at 37 °C. After additional treatment of each digest with alkaline phosphatase (1 μg of protein, 1 h at 37 °C) HPLC analysis was performed. Oligonucleotide containing the *R*p-phosphorothioate substitution was completely digested to mononucleotide by venom phosphodiesterase and yielded *R*p-d[A$_{PS}$T] upon nuclease P1 digestion. Oligonucleotide containing *S*p-phosphorothioate substitution

[1] Abbreviations: 1D NMR, one-dimensional NMR; 2D NMR, two-dimensional NMR; 2D NOE, two-dimensional nuclear Overhauser effect; ATP, adenosine triphosphate; COSY, correlation spectroscopy; DQF-COSY, double-quantum filtered COSY; EDTA, ethylenediaminetetraacetate; FID, free induction decay; RT, retention time; *R*-duplex, hybrid duplex d(GCTATAA$_{ps}$TGG)·r(CCAUUAUAGC) where the modified phosphate has an *R*-type chirality; *S*-duplex, hybrid duplex where the modified phosphate has an *S*-type chirality; TBAF, tetrabutylammonium fluoride; TEAA, triethylammonium; THF, tetrahydrofuran.

was completely digested to mononucleotides by nuclease P1 and yielded Sp-d[$A_{PS}T$] upon digestion with venom phosphodiesterase. The nucleotide d($A_{PS}T$) undigested by enzyme was isolated and compared by HPLC (coinjection) with an authentic sample of this compound with known absolute configuration at phosphorus.

The oligoribonucleotide was synthesized on a Milligen 8750 DNA synthesizer using 2′-silyl protected RNA phosphoroamidites, also from Milligen (Milford, MA), with a modified method that allows for the longer RNA amidite coupling times (12 min). All the synthesis procedure was carried out under sterile conditions. The scale of the synthesis was 15 μmol. The average coupling efficiency was 98% by trityl absorbance. The oligonucleotide was detritylated at the end of the synthesis. It was cleaved from the support and partially deprotected by treatment with 4 mL of ammonium hydroxide/ethanol (3:1) overnight at 55 °C (Stawinski et al., 1988). The reagent was decanted and the beads rinsed six times with 1 mL each of sterile water. The rinses were combined with the decanted ammonium hydroxide/ethanol reagent in a sterile glass test tube (13 × 100 mm) and evaporated to dryness. (Note: Do not use plastic tubes. We observed an unidentified modification of G nucleosides during TBAF treatment in plastic tubes.) The residue was treated with 3 mL of 1 M TBAF in THF for 24 h (Hogrefe et al., 1993). The reaction was quenched by adding 3 mL of 2 M TEAA. The sample was desalted by passage through a Bio-Rad, Inc., Richmond, CA, P-6 DG polyacrylamide desalting column (20 mL bed in a 1 cm diameter column) using sterile water as the eluent. Yield of crude material was 800 OD units.

The oligonucleotide was eluted in the void volume and purified on eight 6 mm thick 20% polyacrylamide/7 M urea gels. Each gel was loaded with 100 OD units of crude oligoribonucleotide and run overnight at 330 V. The product bands were excised, and the oligoribonucleotide eluted with 0.3 M ammonium acetate overnight. The samples were desalted using a Waters, Milford, MA, C-18 Sep-pak environmental cartridge according to the instructions supplied by the manufacturer. The yield of purified oligoribonucleotide was 450 OD units, 69% of theoretically available oligoribonucleotide based on trityl absorbance. The oligoribonucleotide was labeled with [α-^{32}P]ATP and analyzed on a 20% polyacrylamide/7 M urea gel using standard kinasing and running conditions (Sambrook et al., 1989). Bands were visualized by autoradiography and found to be pure. Excess polyacrylamide was found in the sample upon analysis by NMR which was removed by precipitation in ethanol. A portion of the oligoribonucleotide (320 OD units) was resuspended in 300 μL of 0.3 M NaOAc, pH 5, and precipitated by addition of 2.5 volumes of cold ethanol. After sitting overnight at −20 °C, the pellet was spun down, the solvent decanted, and the precipitation repeated. The resulting pellet was passed through a P-6 DG column as described above, yielding 290 OD units; 91% recovery.

The hybrid duplex for NMR experiments was prepared by combining equimolar amounts of the phosphorothioate-modified DNA strand and the RNA strand in phosphate buffer, pH 7.0. The proper stoichiometry for 1:1 duplex formation was determined by performing a titration of one strand with the other. Duplex formation was monitored by UV spectroscopy. The resulting buffer solution (10 mM phosphate, 100 mM NaCl, 0.5 mM EDTA, pH 7.0) was about 2 mM in duplex.

NMR Spectroscopy. NMR spectra were acquired at 500 MHz on either a GN-500 or a GE Omega-500 spectrometer and processed using locally written software (STRIKER and SPARKY) on Sun workstations. All experiments were carried out at 30 °C, except acquisition of 2D NOE spectra in H_2O, which were recorded at 10 °C. Pure absorption-phase 2D NOE spectra in D_2O were acquired in hypercomplex mode (States at al., 1982) with spectral widths of 4000 Hz in both dimensions, and data size of 2048 points in ω_2 and 400 in ω_1. A total of 16 scans were collected for each t_1 value with a repetition delay of 10 s between accumulations. Three 2D NOE spectra at different mixing times (100, 200, and 300 ms) were run for each sample. For both samples (R- and S-duplex), the experiments were carried out using exactly the same experimental conditions without removing the sample from the magnet.

Acquisition of 2D-NOE spectra in H_2O were carried out using NODE-1, a notched excitation sequence containing a train of DANTE pulses (Liu et al., 1993b). This pulse sequence provides a uniform excitation profile across the spectrum, including the exchangeable protons with no excitation of the water resonances, allowing for a more reliable cross-peak integration than other related sequences such as 1331. Spectra in H_2O were recorded with 200 ms mixing time and with spectral width of 10000 Hz. A total of 32 scans were collected for each of the 400 t_1 increments. Other acquisition and processing parameters were the same than in the D_2O experiments.

Pure absorption DQF-COSY were recorded with the TPPI scheme (Marion & Wüthrich, 1983) using 4K complex data points in t_2 and 512 in t_1. The spectral width was 4000 Hz in both dimensions, the number of scans 48, and the repetition delay between accumulations 2.5 s. After zero-filling, a frequency domain data set of 4K × 4K was obtained with a digital resolution of 1.0 Hz/point in ω_2 and 3.9 Hz/point in ω_1. In both dimensions, a sine-squared window function shifted by 30° was used for resolution enhancement.

Longitudinal relaxation experiments were carried out using the inversion–recovery method with a 180° composite pulse (Freeman et al., 1980). A total of 14 experiments with different variable delays (t) were performed. For each time point, 32 scans were collected with 15 s of repetition delay between scans. The intensities of the isolated peaks were fitted to the expression $S(t) = A + Be^{-t/T_1}$. Spin–spin relaxation times were estimated using the spin-echo method in which 32 scans for each of the six variable delays were recorded, allowing 15 s between scans for complete relaxation. The data were fitted to a single-exponential function: $S(t) = Ae^{-t/T_2}$. Imino proton exchange rates with bulk water at 10 °C were determined by analyzing the exponential decay in the intensity of imino protons while a spin-lock field is applied selectively to the water signal (Adams & Lerner, 1992).

DQF-COSY Simulation. DQF-COSY cross-peaks for the deoxyribose sugar protons in the DNA strand were simulated with the programs SPHINX and LINSHA (Widmer & Wüthrich, 1986, 1987). The stick spectra for the spin system of each individual sugar, composed of H1′, H2′, H2″, H3′, and ^{31}P coupled to H3′, were calculated with the program SPHINX. Weak coupling was always assumed except between H2′ and H2″. Line shapes were added to the stick spectra by using the program LINSHA. Digital resolution, apodization functions, and acquisition times were the same as in the experimental spectra. The natural line width for each proton was determined to best reproduce the observed cross-peaks.

Measurement of NOE Cross-Peak Intensities and Generation of Distance Constraints. Integration of 2D NOE cross-peaks was carried out with the locally written program SPARKY. Overlapping peaks were integrated by performing

Table 1: Chemical Shifts (ppm) of Different Protons in d(GCTATAA$_s$TGG)·r(CCAUUAUAGC) at 30 °C for the *S*-Duplex[a]

	NH_2[b,c]	NH[b]	6-8	5/M-2	1′	2′	2″	3′	4′	5′/5″[d]	
1 G	–	–	8.05		6.04	2.79	2.83	4.83	4.28	3.84	3.87
2 C	8.64/6.89		7.65, *7.64*	5.38	6.13	2.44	2.70	4.77	4.37	4.24	4.28
3 T		13.59	7.62	1.70	6.06	2.49, *2.50*	2.73	4.95	4.29	4.22	4.24
4 A	7.71/6.23		8.12	7.09	6.24	2.58	2.89	4.94	4.45	4.27	4.27
5 T		13.17, *13.15*	7.26, *7.25*	1.36	5.88	2.23	2.60	4.90	4.25	4.27	4.19
6 A	–		8.02	6.58, *6.56*	6.02, *6.03*	2.62, *2.65*	2.80, *2.79*	5.02, *5.01*	4.41	4.14	4.24
7 A	7.67/6.19		7.89	7.61	6.14, *6.15*	2.53, *2.50*	2.82	5.00, *5.08*	4.42, *4.45*	4.30, *4.31*	4.27, *4.25*
8 T		13.30	7.13	1.26, *1.31*	5.91, *5.88*	2.20, *2.18*	2.55	4.83, *4.79*	4.25	4.27, *4.25*	4.24, *4.19*
9 G	–	12.27	7.55		5.93	2.49	2.69	4.92	4.38	4.12, *4.11*	4.24, *4.19*
10 G	–	–	7.57, *7.56*		6.17	2.38	2.38	4.61	4.24	4.16	–
11 C	8.11/6.96		8.02, *8.01*	5.92, *5.91*	5.47, *5.46*	4.51		4.44, *4.43*	4.31	3.91	4.05
12 C	8.37/7.06		7.96	5.67, *5.66*	5.53	4.71		4.63	4.45	4.17	–
13 A	7.73/6.58		8.10	7.38	5.91	4.63, *4.62*		4.57	4.45	4.22	–
14 U		13.96, *13.94*	7.62	5.09	5.39	4.34		4.39	4.43	4.10	–
15 U		13.16, *13.13*	8.06	5.60	5.64	4.67, *4.65*		4.61	4.46	4.01	4.17
16 A	–		8.16	6.93, *6.94*	5.92, *5.93*	4.47		4.55	4.60	4.21	–
17 U		13.26, *13.24*	7.61	5.09, *5.10*	5.35	4.47		4.50	4.41	4.10	–
18 A	7.65/6.21		8.06	6.83	5.92, *5.93*	4.54		4.71	4.49	4.24	4.15
19 G	–	13.32, *13.31*	7.43		5.65, *5.64*	4.21		4.56	4.40	4.08	4.06
20 C	8.25/6.96		7.53	5.31	5.80	3.99		4.17	4.17	–	–

[a] When assignments for the *R* stereoisomer are different, they are shown in italics. Blank spaces denote assignments are not applicable, and "–" denotes an unassigned resonance. [b] T = 10 °C. [c] Hydrogen-bonded NH_2 protons resonate at lower field. [d] Nonstereospecific assignment.

a multiple line fitting of the cross-peaks to a Gaussian function and subsequent calculation of the area below the theoretical curve. The line-fitting method has the supplementary advantage of being independent of the specific region selected for integration. In addition to this line-fitting procedure, nonoverlapping peaks were also integrated by a sum-of-points (boxing) method. When volumes obtained by both methods were different, in general weak peaks defined by only a few data points, the boxing method was considered more reliable. The final intensity used was the average of the corresponding peaks on both sides of the diagonal. Cross-peaks involving exchangeable protons were integrated in the 2D NOE spectra in H_2O. Due to the flat excitation profile of the reading pulse used in these experiments, no adjustment of the intensities was necessary.

Interproton distances were calculated from the experimental intensities by using the program MARDIGRAS (Borgias & James, 1990), which makes use of complete relaxation matrix analysis. The program CORMA was used to calculate theoretical NOE intensities from model structures according to the same linear algebra method (Keepers & James, 1984). In all MARDIGRAS and CORMA calculations, a single correlation time for the whole molecule (isotropic motion) was assumed. No internal motions were considered except for methyl groups where a three-state jump model was used (Liu et al., 1992). The scaling factor necessary to normalize experimental and theoretical intensities was determined from those intensities involving fixed proton distances, i.e., cytosine and uracyl H5–H6 and thymidine methyl-H6.

Model Structures. Three different model structures were used for MARDIGRAS calculations. Two of these structures correspond to classical A and B-form and the third one to a heteronomous duplex, where the RNA strand is in standard A-form and the DNA strand adopts an intermediate conformation with the sugar pseudorotation angle around 90°, in the E-type range. This kind of conformation has been recently proposed as the general solution structure in DNA·RNA hybrid duplexes (Salazar et al., 1993a,b). All the initial models were built with the program DNAminiCarlo (Ulyanov et al., 1993) which allows generation of structures by specifying generalized helical parameters. The heteronomous structure was created with glycosidic angles of −130° in the DNA strand and −150° in the RNA strand. Pseudorotation phase angles were 90° and 25° respectively.

To avoid any unfavorable atom contact, the structures were energy minimized with the program package AMBER Version 4.0 (Pearlman et al., 1991) using the all-atom representation. Hexahydrated Na^+ counterions were added at a distance of 5 Å to the phosphorus atoms to neutralize the negative charge. The electrostatic term was calculated using a distance-dependent dielectric constant to mimic bulk solvent effects. After 1000 cycles of energy minimization with the method of steepest descent, the potential energy was decreased to reasonable values without significant distortion of the initial structures.

RESULTS AND DISCUSSION

The numbering scheme of the duplex used in the following discussion is

```
        1  2  3  4  5  6  7  8  9  10
5′-d(G  C  T  A  T  A  Aps T  G  G)-3′
3′-r(C  G  A  U  A  U  U  A  C  C)-5′
       20 19 18 17 16 15 14 13 12 11
```

where ps indicates the phosphorothioate moiety. Molecules containing each of the two possible stereoisomers (*S* or *R*) are designated as *S*-duplex or *R*-duplex.

Proton Resonance Assignments. (a) Nonexchangeable Protons. Nonexchangeable protons were assigned using established techniques for right-handed, double-stranded nucleic acids (Feigon et al., 1983; Scheek et al., 1984) using DQF-COSY and 2D NOE spectra. The chemical shifts of the protons in both duplexes (*S* and *R* isomers) are given in Table 1.

Spin systems in the sugar moiety of the DNA strand could be easily identified in the DQF-COSY spectrum. In the RNA strand, lack of a H1′–H2′ cross-peak (except for terminal residues) and extreme overlapping in the H2′–H3′–H4′ region prevent the use of the DQF-COSY spectrum for spin system identification. However, resonances in the ribose sugars could be identified by their strong NOE peaks with H1′ and confirmed by the cross-peak with the corresponding base proton. In this way, all the resonances, including all H4′ and most of the H5′ and H5″, were assigned. The different intensities of the cross-peaks of H1′ with H2′ and H2″ at short mixing time was used to stereospecifically assign H2′

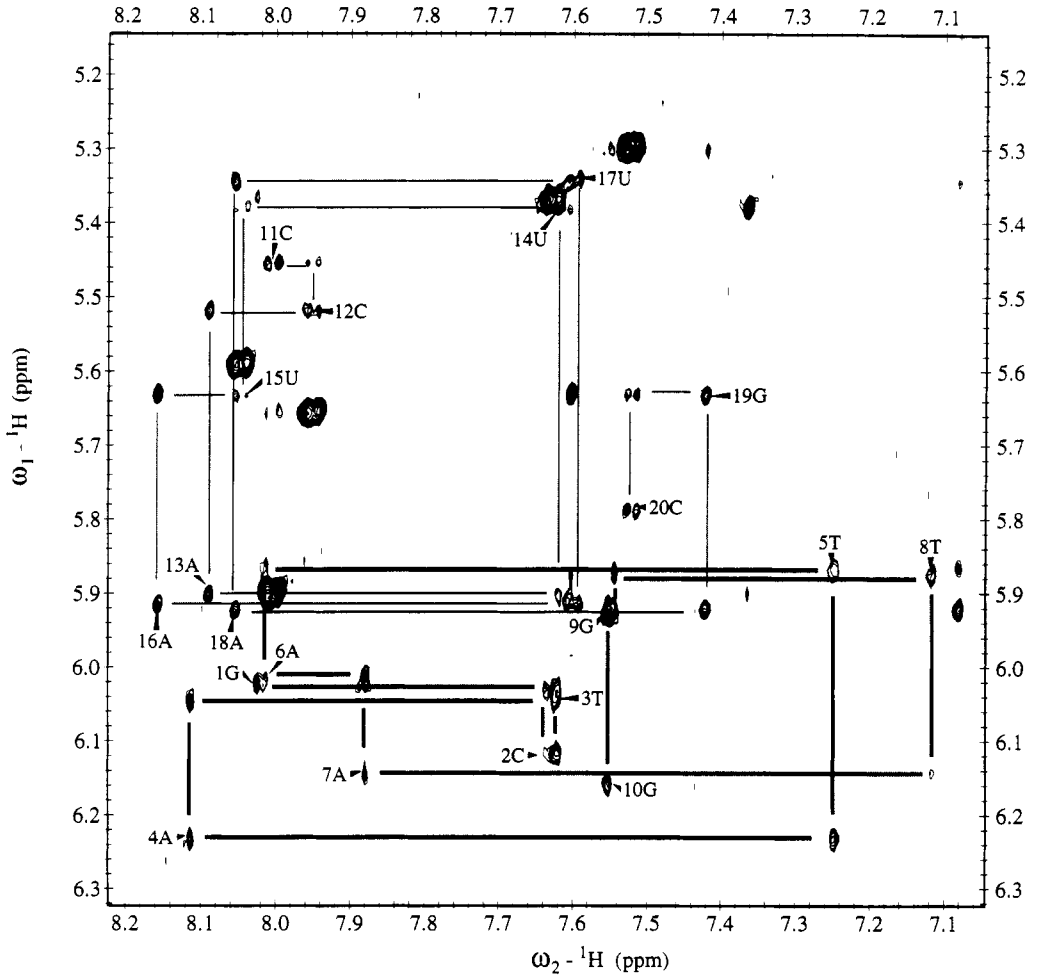

FIGURE 1: Aromatic–H1′ region of the 2D NOE spectrum at 200 ms mixing time of the *R*-duplex in D_2O, T = 30 °C. Sequential assignment pathways are drawn for the strands (thin lines for RNA and thick lines for DNA strand). Labels indicate intranucleotide base–H1′ cross-peak. 14U H1′–H6 overlaps with 2C H5–H6.

and H2″ resonances in deoxyriboses. Stereospecific assignments of H5′ and H5″ could not be performed without further structural assumptions.

Assignment pathways between cross-peaks connecting each base with its own sugar and the 3′ neighbor could be followed in the base–H1′ region, for both strands (Figure 1), and in the base–H2′/H2″, for the DNA strand. Connectivities in the RNA strand could be also determined by using the strong H2′–base sequential and base–H3′ intranucleotide cross-peaks.

Adenine H2 resonances were identified by their strong cross-peaks with H1′ protons. These include sequential connectivities with the 3′-linked nucleotide and interstrand contacts with the base-paired nucleotide to the 5′-neighboring residue. Sequential cross-peaks H1′–H2 are unusually strong as has been previously reported by Chou et al. (1989) in another hybrid duplex. Assignments of H2 protons were also confirmed by their long spin–lattice relaxation times (Table 2).

Both stereoisomers present very similar chemical shifts. Differences are only significant in protons located very close to the modified phosphate: A7H2′, A7H3′, T8H1′, T8M, T8H3′, and T8H5′ (see Figure 2). Therefore, the global structure of the two duplexes must be identical. Any possible difference is only expected in the A7–T8 region.

(b) Exchangeable Protons. Exchangeable protons could be easily assigned by following well-established strategies (Boelens et al., 1985). Imino H3 protons in thymines and uracyls were identified by their strong interstrand cross-peaks with H2 of the base-paired adenines (see Supplementary Material). Additional cross-peaks to the H2 of the 3′ neighbor were also observed. Amino protons of adenines were assigned by their cross-peaks with H3. Unfortunately, the chemical shift degeneration between H3 in T5 and U15 prevents the assignment of amino protons in A6 and A16.

Labile protons in GC base pairs were assigned by following their connection to H5 of cytosines (H5C → HN4C → H1G). Except for the two terminal residues, all the imino protons were identified. Although an extra signal in this region is present at 13.65 ppm, no cross-peak is observed, making its assignment to any of the two remaining imino protons impossible. No amino proton resonances from guanines were observed.

Table 2: T_1 and T_2 Relaxation Times and Calculated Correlation Times[a]

proton	T_1 (s)	T_2 (ms)	τ_c (ns)	proton	T_1 (s)	T_2 (ms)	τ_c (ns)
C2H1′	1.1	20.0	2.6	C11H1′	2.2	64.0	2.0
A4H1′	1.1	15.0	3.0	C12H1′	2.9	41.0	2.9
A4H8	1.4	25.0	2.6	C12H6	2.2	27.0	3.2
A4H2	4.2	85.0	2.4	A13H8	3.3	42.0	3.1
T5H1′	1.3	32.0	2.2	A13H2	5.5	81.0	2.9
T5H6	1.4	55.0	1.7	A16H8	3.0	65.0	2.3
A6H2	4.7	163.0	1.9	A16H2	4.9	88.0	2.6
A7H1′	1.7	32.0	2.5	U17H1′	2.6	33.0	3.1
A7H8	1.1	34.0	2.0	A18H2	4.6	130.0	2.1
T8H6	1.3	50.0	1.8	G19H8	2.3	60.0	2.1
G10H1′	1.2	130.0	1.0	C20H6	1.6	72.0	1.6

[a] Presented T_1 and T_2 estimations are the average values obtained in both duplexes (*R* and *S*).

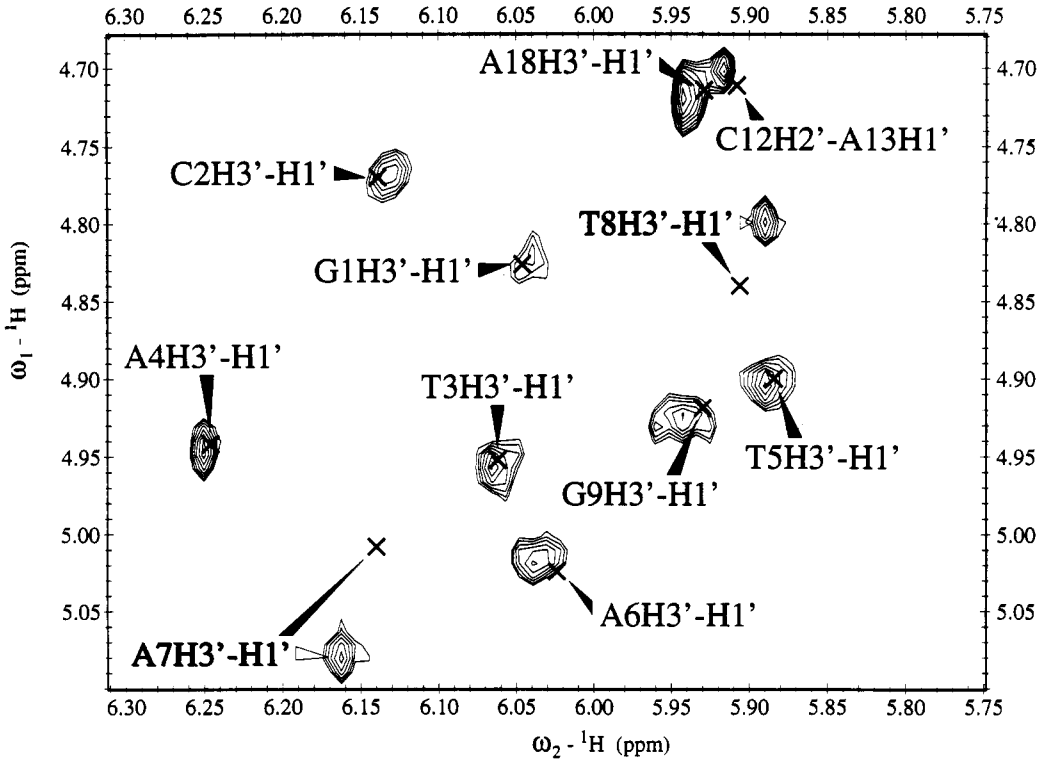

FIGURE 2: Region of the 2D NOE spectrum of the *R*-duplex corresponding to H1′–H3′ cross-peaks (T = 30 °C, mixing time 300 ms). Contour plots belong to the *R*-duplex spectrum, with crosses and labels indicating the peak position in the *S* stereoisomer.

Table 3: Scalar Constants (Including Experimental Sums), Pseudorotation Phase Angles (P_s), Pucker Amplitudes (ψ), and Percentage of *S* Conformers[a]

	J(1′,2′)	J(1′,2″)	J(2′,3M)	J(2″,3′)	ΣH1′-ex	ΣH2′-ex	ΣH2″-ex	P_s	ψ	%*S*
1 G[b]										
2 C	6.1[c]	6.1[c]	6.3[c]	4.0[d]	12.2	26.7	25.7	150–190	35	55–67
3 T	7.1	6.1	5.9	3.5[c]	13.1	27.1	23.0	160–190	35	57–69
4 A	7.1	6.1	5.9	3.5[c]	13.1	27.3	23.0	160–190	35	57–69
5 T	8.5	6.0	6.5[c]	<3.0	14.5	~28.5	~22.0	140–170	35	78–86
6 A	8.3	6.5[d]	5.5[c]	<3.0	14.4	28.5		165–205	35	79–100
7 A R	6.3	6.3	6.3	4.0[d]	12.6	25.9	24.9	160–190	35	56–67
S	7.3	6.3	6.6	3.0[d]	13.7	28.0	22.5	145–165	35	67–71
8 T					~14.3	~26.7	~22.0			
9 G	8.5	6.0	6.0[c]	<3.0	14.5	24.5	20.4	150–205	35	79–100
10 G[b]										

[a] Only in the case of A7, the two isomers (*S* and *R*) present different values. Estimated error ±0.3 Hz. [b] H2′ and H2″ have the same chemical shifts. [c] Error ±0.5 Hz. [d] Error ±1.0 Hz.

Analysis of DQF-COSY. DQF-COSY is an essential experiment for sugar puckering determination. Sugar conformation can be reliably established from the information contained in the vicinal proton–proton *J*-coupling constants, which can be related to different dihedral angles by means of the Karplus equation (Rinkel & Altona, 1987).

(a) Determination of the Coupling Constants. *J*-coupling constants were obtained from DQF-COSY spectra. The first remarkable feature of these spectra is the lack of H1′–H2′ cross-peaks in the RNA strand, indicating a predominant N-type conformation for the ribose sugars. Only one peak, corresponding to the 3′-terminal nucleotide (see Supplementary Material), is observable, probably due to extra mobility in this residue. Signal overlapping in the H2′–H3′–H4′ region in the RNA strand prevents the determination of these *J*-coupling constants. For deoxyriboses, instead, $J_{1'2'}$, $J_{1'2''}$, $J_{2'3'}$, and $J_{2''3'}$ could be obtained by simulating H2′/H2″-(ω_1)–H1′(ω_2), H1′(ω_1)–H2′/H2″(ω_2), and H3′(ω_1)–H2′/H2″(ω_2) cross-peaks. H2′(ω_1)–H3′(ω_2) cross-peaks were, in general, too weak to be detected and appear close to the residual water ridge. Signal overlapping is not severe in these regions, and cross-peaks are, except for the terminal residues, well-resolved. For T8, cross-peak positions impeded a reliable *J*-coupling determination for this residue. The strategy for the cross-peak simulation in the DNA strand is similar to the one used by Weisz et al. (1992) for a double-stranded DNA duplex. In general, line widths are found to be slightly larger than in double-stranded DNA (7–8 Hz for H1′, ~8 Hz for H2″, and 9 Hz for H2′).

Coupling constant values are shown in Table 3 together with an experimental estimation of the sum of the *J*-couplings for each proton. This estimation is useful to confine the range of *J* values to explore during simulations. Figure 3 illustrates the quality of the fit for nucleotide A4. Similar figures for other nonterminal nucleotides are included as supplementary material. The two terminal nucleotides exhibit similar chemical shifts for H2′ and H2″ (complete degeneration in the case of the 3′-terminal), preventing any *J*-constant estimation. Adenine 7 is the only nucleotide which presents a different cross-peak pattern in the two stereoisomers.

Coupling constants between H3′ and H4′ protons could not be determined accurately. H3′–H4′ cross-peaks are well resolved in both dimensions, but the lack of fine structure in these peaks together with the number of passive couplings

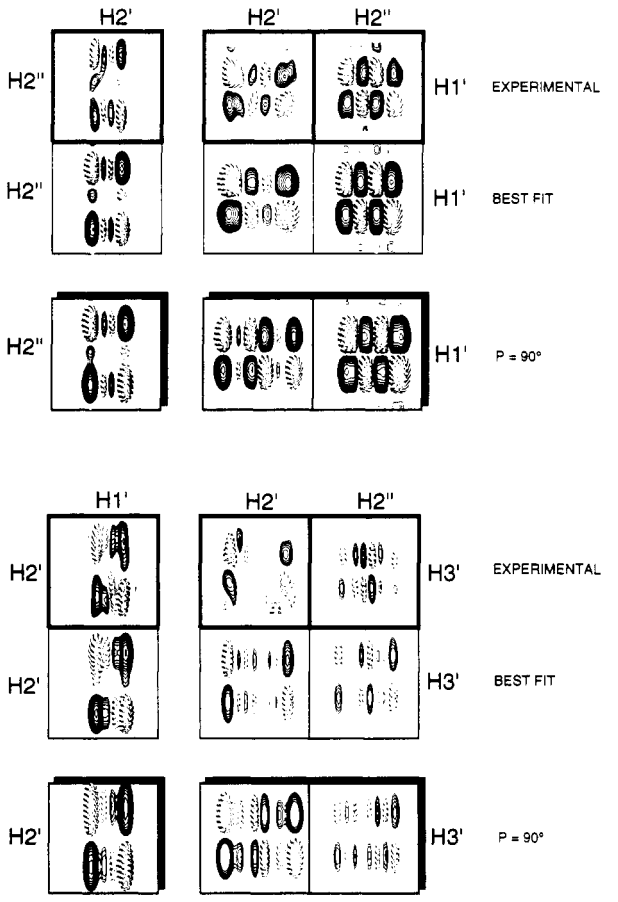

FIGURE 3: Plots of experimental (broad line boxes) and simulated (thin line boxes) double-quantum-filtered COSY cross-peaks for nucleotide A4. Negative peaks are dashed. The best fit simulated plots correspond to J values given in Table 3. Plots of simulated cross-peaks for a single conformer with a sugar pseudorotation phase angle of 90° and amplitude of 35° are obtained with the following J values derived from the Karplus equation: $J_{1'2'} = 9.4$, $J_{1'2''} = 7.5$, $J_{2'3'} = 9.5$, and $J_{2''3'} = 4.0$ Hz.

involved made the exact determination of the active coupling constant impossible. H3′(ω_2)–H2′/H2″(ω_1) cross-peaks also contain information about $J_{3'4'}$ but these peaks were too weak to be used. $J_{3'4'}$, values in the range of 3–6 Hz with $J_{3'^{31}P}$ between 5 and 8 Hz equally reproduce the experimental H3′–H4′ cross-peaks. Variation of these two coupling constants did not affect the cross-peak pattern of H3′(ω_1)–H2′/H2″-(ω_2), since H3′ couplings appear here in the poorly resolved dimension.

(b) Sugar Pucker Determination. Ranges of pseudorotational parameters consistent with the experimental coupling constants were determined by comparing the previously obtained experimental values of each $J_{\alpha\beta}$ with the average calculated according to

$$J_{\alpha\beta} = (1 - f_s)J_{\alpha\beta}(\mathrm{N}) + f_s J_{\alpha\beta}(\mathrm{S})$$

where f_s is the fraction of S-type conformer and J(S) and J(N) are the respective coupling constants between proton α and β for the S and N conformers. The J constants for each conformer were obtained from tables published by Rinkel and Altona (1987). To reduce the number of variables, equal pucker amplitudes for S and N conformers were assumed, and the phase angle for the N conformer was set to 9°. The three remaining parameters were varied within certain ranges: fractional population of S conformer (f_s) from 50 to 100%, pseudorotation phase angle (P_S) from 0° to 360°, and pucker amplitude (Ψ_S) of the major conformer from 20° to

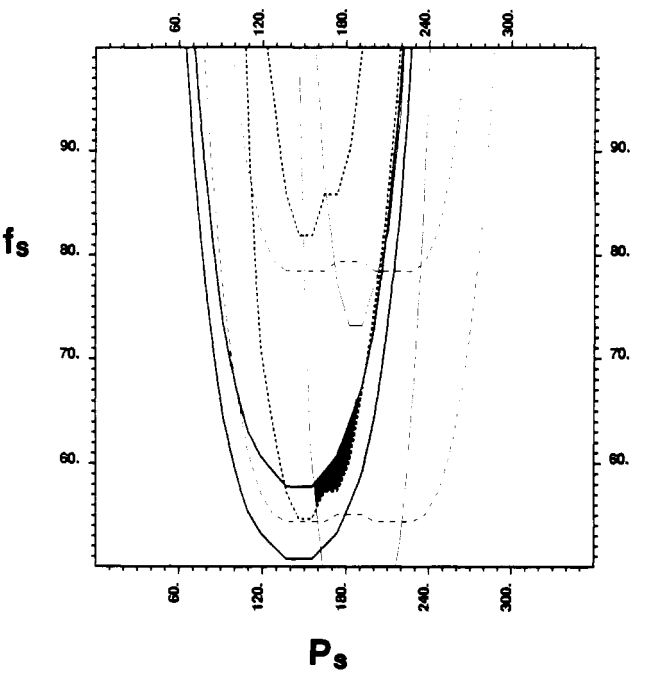

FIGURE 4: Contour plots of the estimated J-coupling values for A4 versus pseudorotation phase angle (P_s) and population of the major S conformer (f_s) for a pucker amplitude of 35°. Solid bold lines represent $J_{1'2'} = 7.1 \pm 0.3$ Hz; dashed bold lines, $J_{1'2''} = 6.1 \pm 0.3$ Hz; solid thin lines, $J_{2'3'} = 5.9 \pm 0.3$ Hz; and dashed thin lines, $J_{2''3'} = 3.5 \pm 0.5$ Hz. The region consistent with all the coupling constants measured for A4 is marked in black.

40° (20°, 30°, 35°, and 40°). For each pucker amplitude, contour lines corresponding to the upper and lower bounds of the experimental $J_{\alpha\beta}$ values were drawn in a (f_s,P_s) map. The intersection of the allowed regions for all the $J_{\alpha\beta}$ is the range of sugar pucker parameters consistent with the experimental data (see Figure 4 for an example). Intersection of the permitted regions differed from zero only for pucker amplitudes of 35°. Ranges of pseudorotational parameters are presented in Table 3; these figures clearly represent extreme values. The likely ranges are probably smaller because the region is not square (Figure 4). Plots of the (f_s,P_s) map consistent with the experimental data for each nucleotide in the DNA strand are included as supplementary material. Obtaining pseudorotation parameters by this graphical method, instead of calculating the single values that best fit the experimental constants, has the advantage of properly taking into account the error bounds in the estimation of the J values. Also, this method does not presuppose a two-state equilibrium since, in those cases where one single conformer can accommodate all the J-coupling values in a sugar, the points where $f_s = 100\%$ must be contained in the intersection of allowed regions of the (f_s,P_s) map. In our hybrid duplex, this occurs only for A6 and G9.

In recent studies of hybrid duplexes in solution (Salazar et al., 1993a,b), an unusual E-type conformation ($P_S \sim 90°$) for the deoxyribose sugars has been proposed. From the contour plot in Figure 4, it can be seen that this conformation (f_s = 100%, $P_s = 90°$) is not in agreement with our J-coupling data. To further check this result, a cross-peak simulation was carried out with the J-coupling constants calculated from the Karplus equation (Rinkel & Altona, 1987) for a phase angle of 90° and a pucker amplitude of 35°. Simulated cross-peaks are shown in Figure 3. Although H1′(ω_2)–H2′/H2″(ω_1) and H1′-(ω_1)–H2″(ω_2) are similar to the experimental ones, H3′(ω_1)–H2′(ω_2) and, especially, H1′(ω_1)–H2′(ω_2) present a rather different cross-peak pattern. The *in phase* splitting in the H1′(ω_1)–H2′(ω_2) cross-peak produced by the large $J_{2'3'}$

coupling constant (~9 Hz), distinctive of this conformation, does not appear in our experimental peaks. Strong H3′–H4′ DQF-COSY cross-peaks have also been proposed as an indication of pseudorotation phase angle around 90°, since $J_{3'4'}$ is very small for pure S-type rings. Although our experimental H3′–H4′ cross-peaks are fairly strong, they can be reproduced considering a population of N-type conformer within the range determined by the other coupling constants. Simulations carried out for $J_{3'4'}$ in the range of 3–5 Hz gave rise to strong H3′–H4′ peaks depending on the value of H3′–^{31}P coupling.

Relaxation Time Measurements. Spin–lattice relaxation times (T_1) and spin–spin relaxation times (T_2) were measured utilizing 1D NMR techniques. T_1 was determined for isolated base and H1′ protons with the inversion–recovery method. T_2 was estimated by spin-echo experiments. The T_1 and T_2 relaxation times are listed in Table 2. The difference in spin–lattice relaxation times between protons in the DNA and the RNA strand has been previously reported by other authors (Wang et al., 1992) and seems to be a general feature of DNA·RNA hybrid duplexes. According to these authors, the large T_1 difference for H1′ between the DNA and the RNA strand is due to the lack of H2″, which is one of the most efficient relaxation pathways for H1′ in the deoxyribose sugars. Therefore, there is no reason to assume that isotropic tumbling is not a correct approximation for hybrid duplexes as it is for double-stranded DNA duplexes. Consequently, the single correlation time can be estimated from T_1 and T_2 according to Suzuki et al. (1986). The average τ_c obtained by this method is 2.3 ± 0.7 ns. In the subsequent relaxation matrix calculations, global correlation times of 1.5, 2.0, and 2.5 ns were used.

The large difference in longitudinal relaxation times between DNA and RNA has an important implication in the extraction of interproton distance constraints. Serious errors in distances involving RNA protons may be present if a sufficiently long time is not allowed for full relaxation. Thus, long repetition delays between scans must be used in the NMR experiments. With the spin–lattice relaxation times reported here, repetition delays of more than 13 s would be necessary to assure complete equilibration. In practice, a compromise between complete relaxation and nonprohibitive experimental times must be reached.

Analysis of 2D NOE Intensities. To obtain accurate and reliable structures of nucleic acids in solution, it is necessary to have a large number of accurate distance constraints. Interproton distances can be accurately obtained from 2D NOE intensities by using the complete relaxation matrix approach implemented in the MARDIGRAS algorithm (Borgias & James, 1988, 1990). This algorithm properly takes into account spin diffusion effects and has been demonstrated to be extremely useful in structure determination of oligonucleotides (Gochin & James, 1990; Kerwood et al., 1991). In the present study, distance constraints were obtained from NOE data sets acquired for each duplex at three different mixing times. Although it has been shown that final distances do not depend too much either on the initial model or the correlation time, several MARDIGRAS calculations were run with different starting structures and correlation times to avoid any possible bias coming from initial conditions. Three different initial models were used, i.e., standard A- and B-form duplexes and the heteronomous duplex (see Materials and Methods). Correlation times of 1.5, 2.0, and 2.5 ns, in an interval around the average τ_c estimated from the calculated T_1 and T_2 values, were employed. Final distances were obtained by averaging a total of 27 MARDIGRAS outputs (3 mixing times × 3 starting structures × 3 correlation times). MARDIGRAS calculates upper and lower bounds for each distance by performing an error propagation analysis from the estimated uncertainty in the corresponding intensity. Final distance errors were determined by averaging upper and lower bounds in all 27 individual runs.

Table 4: Imino Proton Exchange Rates in Water

	k (s^{-1})		k (s^{-1})
T3	1.4	8T/G19	2.9
T5/U15	2.8	U14	2.6
G9	2.0	U17	2.9

Exchangeable protons were also included in MARDIGRAS calculations. Exchange effects can be considered in addition to dipole–dipole relaxation: $R_T = R_{DD} + K$, where K represents the exchange rates of the labile protons (Liu et al., 1993a). The distance bounds estimated from MARDIGRAS should account for exchange. Upper distance bounds can be obtained by neglecting the exchange of labile protons, and lower bounds can be calculated by assuming large exchange rates. This method has been used previously in our group (Weisz et al., 1994), but the range between upper and lower bounds can be substantially reduced by determining the exchange rate of the labile protons (Bishop et al., 1994). Imino proton exchange rates in water were determined with a 1D NMR method (Adams & Lerner, 1992) which utilizes a spin-lock field applied selectively to the water signal. The values obtained are listed in Table 4. Based on these results, a conservative value of 4 s^{-1} was assumed for T and U imino protons in MARDIGRAS calculations, and 10 s^{-1} was used for imino C and G protons as well as amino protons, which exchange rates could not be determined due to signal overlap. We have used these large exchange rate upper limits to set lower distance bounds conservatively.

MARDIGRAS calculations for labile protons were carried out from 2D NOE intensities obtained at a mixing time of 200 ms using a pulse sequence with a virtually flat excitation profile in the imino and amino proton region (Liu et al., 1993b). Several calculations were repeated with different model structures and correlation times, first, neglecting exchange rates and, secondly, using the exchange rate limits obtained above. Upper distance limits were obtained by averaging output distances in the first set of calculations, and lower bounds were determined from the second set. Around 30 distances involving exchangeable protons were obtained for each duplex.

The total number of distance constraints was 322, including both exchangeable and nonexchangeable protons. This set was comprised of 182 intraresidue, 115 interresidue, and 25 interstrand constraints. The average number of constraints per base pair is 28. This large number of distance restraints, together with torsion angle constraints obtained from correlated experiments, will be used to calculate the structure of these hybrid duplexes by restrained molecular dynamics.

Theoretical NOE intensities were calculated for each of the model structures by using complete relaxation matrix analysis with the program CORMA. The calculated intensities were compared with the experimental ones using two of the numerical indices (R factors) commonly employed in the literature (Thomas et al., 1991; González et al., 1991). Both the crystallographic-like R factor and sixth-root R factor indicate similar results. The first yields values of 0.44, 0.72, and 0.47 for A, B, and heteronomous structures (summing all the NOE intensities). For the sixth-root R factor, these values are 0.10, 0.15, and 0.10. According to these results, experimental intensities appear to fit better with a global

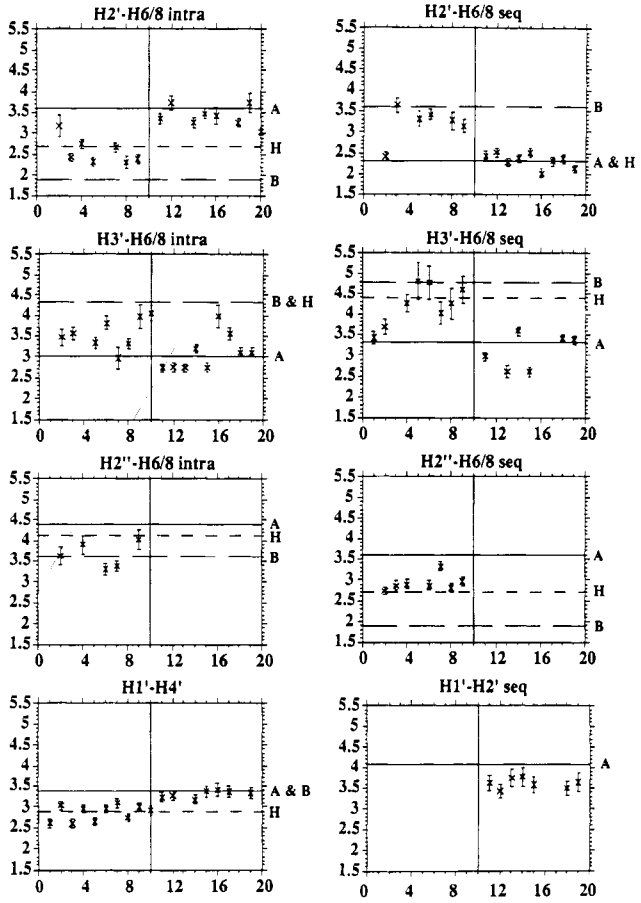

FIGURE 5: Interproton distances (Å) resulting from MARDIGRAS calculations plotted versus the nucleotide sequence for some significant contacts. Residues from 1 to 10 correspond to the DNA and, separated by the vertical line, from 11 to 20 to the RNA strand. Horizontal lines labeled A, B, and H indicate the average distance in each of the initial models (see text). The value for H (heteronomous structure) applies only to the DNA strand. Distances in the RNA strand of this structure are equivalent to the A-form.

A-form than with the H structure, but the difference might not be significant. In contrast, the B-form duplex clearly does not represent the experimental data.

Differences between the DNA and the RNA strand can be assessed by inspecting the resulting distances. Some of them are plotted in Figure 5 together with the average model interproton distances in the three starting structures. Specially significant are H2′–base contacts, both sequential and intranucleotide. In the RNA strand the average sequential distance is around 2.3 Å and the intraresidue around 3.5 Å. Both values agree quite well with an A-form structure. In the DNA strand the situation is not so clear. Sequential contacts are around 3.3 Å and intranucleotide around 2.5 Å in both cases an intermediate value between the A- and B-form. The heteronomous structure seems to match the intranucleotide distances better than the other two models. The same applies for the H2″–base sequential distances, the average value of 2.7 Å agrees with an intermediate structure better than with standard A- or B-forms. Another clear distinction between both strands comes from the H1′–H2′ sequential contacts which are only observable in the RNA strand. These NOEs are only expected in A-form structures. The sequential distances between H1′ and H2 of adenines match with an A or heteronomous form better than with a B-type structure in both strands. The intensity of these sequential contacts in the RNA strand is unusually high, giving rise to distances around 3.1 Å while the model distance in A or heteronomous forms is around 4.1 Å. Adenine H2 protons also exhibit strong interstrand NOE peaks with H1′ of the base-paired nucleotide to the 5′ neighbor with distances around 3.5 Å, typical of A-form structures.

Among the different interproton contacts, intranucleotide H1′–H4′ and H2″–H4′ are the most directly related to the sugar conformation. Other contacts either do not change significantly with sugar pucker parameters or depend on additional variables such as the glycosidic angle. The variation in H1′–H4′ distance with the pseudorotation phase angle is not very high (~0.5 Å), but, as shown in Figure 5, a clear difference between both strands can be observed. In this case, N- and S-type sugars exhibit the same distance (3.4 Å). Experimental distances in ribose sugars match this value, but, in deoxyriboses, the distances are shorter, around 3.0 Å on average. Only an E-type conformation, with P_s ~90°, may account for these strong NOEs. The other meaningful intrasugar distance, H2″–H4′, exhibits an average value of 3.4 Å, intermediate between N- and S-type conformers (data not shown). These contacts can be explained either by a single intermediate puckering or by two conformers (N and S) in rapid exchange. More information can be obtained from base–sugar contacts, but their interpretation is more difficult since they are affected by the glycosidic angle as well. H3′–H6/8 distances are especially sensitive to sugar pseudorotation parameters (Wijmenga et al., 1993; Ulyanov et al., 1994), yielding short values only in N-type sugars. In two of our initial models, B- and H-form, this distance is 4.3 Å; in the A-form structure, the value is 3.0 Å. The experimental distances are in the range between 3.0 and 4.0 Å and are definitely shorter than expected for an E-type conformation. As pointed out by Wijmenga et al. (1993), these short distances could be explained by assuming some population of N-type as well as S-type conformer. Other distance constraints also contain information about the sugar puckering, but a restrained molecular dynamics calculation is necessary to extract proper conclusions from them. At present neither a single sugar conformation in the E-type domain nor two conformers in rapid exchange between N and S domain can fully account for the distance information. Quite possibly, more than two sugar pucker states in equilibrium may be necessary to explain all the experimental data; a study utilizing molecular dynamics with time-average distance restraints indicated that some residues could have a significant population of sugars with P = 60°–120° in equilibrium with a dominant S-conformer (Schmitz et al., 1993).

CONCLUSIONS

Although restrained molecular dynamics calculations are necessary to obtain a detailed three-dimensional structure of this molecule, some features of the general conformation can be assessed from the distances and torsion angle constraints obtained in this work. Both strands are clearly heterogeneous. Whereas all the interproton contacts agree with the RNA strand being in an A-type structure, the DNA strand adopts a different conformation which cannot be directly classified as either A- or standard B-form. A model structure with the RNA strand in A-type conformation and the DNA strand in an intermediate between A- and B-form with pseudorotation phase angles set to 90° and glycosidic angles of −130° seems to account better for some of the experimental distances, but at present it appears that one single conformation cannot represent all the experimental data.

The conformation of riboses and deoxyriboses in our hybrid duplexes is clearly different. Both J-coupling information and distances obtained from 2D NOE experiments indicate that the sugar pucker in the RNA strand is N-type, typical

of the A-form family. On the other hand, cross-peak patterns in the DQF-COSY as well as experimental distances show that the deoxyribose sugars do not adopt a pure C3′-endo conformation. This result is not consistent with the structure of another hybrid chimeric duplex studied in the crystal state (Egli et al., 1992) but agrees with the X-ray fiber diffraction structure obtained by Arnott et al. (1986) and with most of the recent studies in solution (Chou et al., 1989; Katahira et al., 1990; Salazar et al., 1993a,b). Distance information is not conclusive in determining the conformation of the deoxyriboses. Although H1′–H4′ contacts suggest a pseudorotation phase angle around 90°, intranucleotide H3′–base distances match better with a mix of N- and S-type conformers in rapid exchange.

Sugar pucker is, in general, more accurately determined from the *J*-coupling information. The analysis of the DQF-COSY cross-peaks gives us a more precise description of the sugars in the DNA strand. With the exception of A6 and G9, the scalar coupling data from the deoxyribose ring cannot be justified by one single conformer (f_s = 100%). In most of the nucleotides, some population of N-type sugar is required to account for the *J*-coupling data, this population being up to 40% in some cases. This result contradicts recent studies of DNA·RNA hybrids in solution by NMR (Salazar et al., 1993a; Fedoroff et al., 1993) which postulate a single sugar conformation in the E-type domain ($P_S \sim 90°$). Although our experimental values of $J_{1'2'}$, $J_{1'2''}$, and $J_{2''3'}$ are consistent with an E-type conformation, the $J_{2'3'}$ values (5.9–6.5) are always much lower than the expected one for $P_S \sim 90°$ ($J_{2'3'} \sim 9$ Hz, according to the Altona's tables). Unfortunately, none of the previous studies reports $J_{2'3'}$ values, making it difficult to know if multiple conformers should have been invoked in those cases as well. It is interesting to note that Lane et al. (1993), analyzing only sums of coupling constants, also reached a conclusion similar to ours: the deoxyribose sugars in a DNA·RNA hybrid duplex are in rapid exchange between S- and N-type conformations.

There has recently been a report that determination of scalar coupling constants can be compromised by strong dipolar coupling (Harbison, 1993). For example, dipolar coupling between the geminal H2′ and H2″ protons could affect scalar coupling constant values derived from coupling entailing those protons. It would appear, however, that for the short correlation times typical for the size of nucleic acid duplexes such as those studied here, the modifications would be much smaller than the other experimental errors; Schmidt et al. (1994) have shown that correlation time of <7 ns will result in coupling constant corrections of <1 Hz.

The populations of minor N conformer obtained in our hybrid duplex seem to be significantly higher than the fractions of minor conformer found in DNA duplexes studied by our lab. Specially interesting is a comparison with the sugar puckering found in the analogous Pribnow sequence in double-stranded DNA (Schmitz et al., 1992). Populations of the major *S* conformer in the hybrid duplex are systematically lower than the corresponding ones in double-stranded DNA (around 85%). Pseudorotation phase angles are determined with less precision than in Schmitz et al. (1992), but it can be established that their values are closer to 180°.

In summary, the DNA strand of this DNA·RNA hybrid duplex appears to be more flexible than its corresponding double-stranded DNA. This flexibility could be the reason for the different structures of hybrids and chimeric duplexes found in solid state by crystallographic methods. On the other hand, there is no evidence of this augmented flexibility in the RNA strand since the presence of a minor *S*-conformer in the riboses with a population similar to the one found in the deoxyriboses would increase the $J_{1'2'}$ coupling constant to detectable values. Moreover, even a small population of *S*-type conformer in the RNA strand would lower the base–H2′ intranucleotide distances since these contacts are very strong in B-form structures. This effect has not been observed, as shown in Figure 5, where base–H2′ intranucleotide distances match an A-form structure very well.

In this study a comparison between the two possible stereoisomers resulting from the substitution in the phosphate between A7 and T8 has been carried out. The great similarity in chemical shifts between the *S* and *R* stereoisomers is a definitive indication of nearly identical structures for both molecules. Since the only protons that exhibit different chemical shifts are located near the modified phosphate (A7H2′, A7H3′, T8H1′, T8M, T8H3′, and T8H5′/H5″), any structural dissimilarity must be very local. MARDIGRAS output distances involving these protons in the *S* and *R* stereoisomers were carefully checked, and no significant variation was observed. In contrast, *J*-coupling constants in the 5′-linked neighbor (A7) to the thiophosphate are slightly different (Table 3). This difference is explained by a higher minor conformer population in the *R*-duplex. Overlapping in the T8 cross-peaks prevented a proper *J*-coupling constant determination but an estimation of the sums of the coupling constants indicates a similar conformation for both duplexes.

Electrostatic or steric effects due to the presence of the sulfur atom in different positions could account for the variation in chemical shifts without invoking any other structural change. The sulfur atoms appear to exist in either position without need for any structural accommodation. A similar conclusion was previously reached for thiophosphate-modified double-stranded DNA duplexes by LaPlanche et al. (1986). More insight on the effect of the sulfur substitution will be gained when the three-dimensional structures of the two duplexes are obtained.

ACKNOWLEDGMENT

We gratefully acknowledge the help of Dr. Klaus Weisz during the initial phase of this project, the technical expertise of Lionella Borozdina in preparing the RNA, and Drs. Nikolai Ulyanov and Uli Schmitz for many useful discussions. C.G. acknowledges a postdoctoral fellowship from the Spanish Consejo Superior de Investigaciones Científicas (CSIC).

SUPPLEMENTARY MATERIAL AVAILABLE

Six figures showing the imino–base region of the 2D NOE spectrum in H_2O of the *S*-duplex (T = 10 °C, mixing time 200 ms); sections of the DQF-COSY of d(GCTATAA$_{ps}$-TGG)·r(CCAUUAUAGC) for the *R*-duplex (T = 30 °C), showing the H2′(ω_1)–H1′(ω_2) region in the DNA strand and the H1′(ω_1)–H2′(ω_2) in the RNA strand; plots of experimental (broad line boxes) and simulated (thin line boxes) double-quantum-filtered COSY cross-peaks for nonterminal nucleotides; and contour plots of the estimated *J*-coupling values for every nucleotide versus pseudorotation phase angle (P_s) and population of major *S* conformer (f_s) for an amplitude of 35°, showing the allowed regions of the (P_s,f_s) map (6 pages). Ordering information is given on any current masthead page.

REFERENCES

Adams, B., & Lerner, L. (1992) *J. Magn. Reson. 96*, 604–607.
Arnott, S., Chandrasekaran, R., Millane, R. P., & Park, H. S. (1986) *J. Mol. Biol. 188*, 631–640.
Bishop, K. D., Blocker, F. J. H., Egan, W., & James, T. L. (1994) *Biochemistry 33*, 427–438.

Bryant, F. R., & Benkovic, S. J. (1978) *Biochemistry 18*, 2825–2827.
Boelens, R., Scheek, R. M., Dijkstra, K., & Kaptein, R. (1985) *J. Magn. Reson. 62*, 371–386.
Boelens, R., Koning, T. M. G., & Kaptein, R. (1988) *J. Mol. Struct. 173*, 299–311.
Boelens, R., Koning, T. M. G., Marel, G. A. v. d., Boom, J. H. v., & Kaptein, R. (1989) *J. Magn. Reson. 82*, 290–308.
Borgias, B. A., & James, T. L. (1988) *J. Magn. Reson. 79*, 493–512.
Borgias, B. A., & James, T. L. (1990) *J. Magn. Reson. 87*, 475–487.
Celda, B., Widmer, H., Leupin, W., Chazin, W. J., Denny, W. A., & Wüthrich, K. (1989) *Biochemistry 28*, 1462–1471.
Chou, S.-H., Flynn, P., & Reid, B. (1989) *Biochemistry 28*, 2435–2443.
Cruse, W. B. T., Salisbury, S. A., Brown, T., Cosstick, R., Eckstein, F., & Kennard, O. (1986) *J. Mol. Biol. 192*, 891–905.
Egli, M., Usman, N., Zhang, S., & Rich, A. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 534–538.
Fedoroff, O. Y., Salazar, M., & Reid, B. R. (1993) *J. Mol. Biol. 233*, 509–523.
Feigon, J., Leupin, W., Denny, W. A., & Kearns, D. R. (1983) *Biochemistry 22*, 5943–5951.
Freeman, R., Kempsell, S. P., & Levitt, M. H. (1980) *J. Magn. Reson. 38*, 453–479.
Gao, X., Brown, F. K., Jeffs, P., Bischofberger, N., Lin, K., Pipe, A. J., & Noble, S. A. (1992) *Biochemistry 31*, 6228–6235.
Gochin, M., & James, T. L. (1990) *Biochemistry 29*, 11172–11180.
González, C., Rullmann, J. A. C., Bonvin, A. M. J. J., Boelens, R., & Kaptein, R. (1991) *J. Magn. Reson. 91*, 6228–6235.
Hall, K. B. (1993) *Curr. Opin. Struct. Biol. 3*, 336–339.
Harbison, G. S. (1993) *J. Am. Chem. Soc. 115*, 3026–3027.
Heinemann, U., Rudolph, L. N., Claudis, A., Morr, M., Heikens, R. F., & Blocker, H. (1991) *Nucleic Acids Res. 19*, 427–433.
Hogrefe, R. I., McCaffrey, A. P., Borozdina, L. U., McCampbell, E. S., & Vaghefi, M. M. (1993) *Nucleic Acids Res. 21*, 4739–4741.
Katahira, M., Lee, S. J., Kobayashi, Y., Sugeta, H., Kyogoku, Y., Iwai, S., Ohtsuka, E., Benevides, J. M., & Thomas, G. J. (1990) *J. Am. Chem. Soc. 112*, 4508–4512.
Katayanagi, K., Miyagawa, M., Matsushima, M., Ishikawa, M., Kanaya, S., Ikehara, M., Matsuzaki, T., & Morikawa, K. (1990) *Nature 347*, 306–309.
Keepers, J. W., & James, T. L. (1984) *J. Magn. Reson. 57*, 404–426.
Kerwood, D. J., Zon, G., & James, T. L. (1991) *Eur. J. Biochem. 197*, 583–595.
Koziolkiewicz, M., & Wilk, A. (1933) in *Methods in Molecular Biology* (Agrawal, S., Ed.) pp 207–224, Humana Press, Inc., Clifton. NJ.
Lane, A. N., Ebel, S., & Brown, T. (1993) *Eur. J. Biochem. 215*, 297–306.
LaPlanche, L. A., James, T. L., Powell, C., Wilson, W. D., & Uznanski, B. (1986) *Nucleic Acids Res. 14*, 9081–9093.
Lesser, D. R., Grajkowsky, A., Kurpiewski, M. R., Koziolkiewicz, M., Stec, W. J., & Jen-Jacobson, I. (1992) *J. Biol. Chem. 267*, 24810–24818.
Liu, H., Thomas, P. D., & James, T. L. (1992) *J. Magn. Reson. 98*, 163–175.
Liu, H., Kumar, A., Weisz, K., Schmitz, U., Bishop, K., & James, T. L. (1993a) *J. Am. Chem. Soc. 115*, 1590–1591.
Liu, H., Weisz, K., & James, T. L. (1993b) *J. Magn. Reson. 105*, 184–192.
Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95.
Madrid, M., Llinas, E., & Llinas, M. (1991) *J. Magn. Reson. 93*, 329–346.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967.
Mujeeb, A., Kerwin, S. M., Egan, W., Kenyon, G. L., & James, T. L. (1992) *Biochemistry 31*, 9325–9338.
Nakamura, H., Oda, Y., Iwai, S., Inoue, H., Ohtsuka, E., Kanaya, S., Kimura, S., Katsuda, C., Katayanagi, K., Morikawa, K., Miyashiro, H., & Ikehara, M. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 11535–11539.
Pearlman, D. A., Case, D. A., Caldwell, J. C., Seibel, G. L., Singh, U. C., Weiner, P., & Kollman, P. A. (1991) *AMBER4.0*, University of California, San Francisco.
Potter, B. V. J., Connolly, B. A., & Eckstein, F. (1983) *Biochemistry 22*, 1369–1377.
Rinkel, L. J., & Altona, C. (1987) *J. Biomol. Struct. Dyn. 4*, 621–649.
Salazar, M., Champoux, J. J., & Reid, B. R. (1993a) *Biochemistry 32*, 739–744.
Salazar, M., Fedoroff, O. Y., Miler, J. M., Ribeiro, N. S., & Reid, B. R. (1993b) *Biochemistry 32*, 4207–4215.
Sambrook, J, Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor University Press, Cold Spring Harbor, NY.
Scheek, R. M., Boelens, R., Russo, N., van Boom, J. H., & Kaptein, R. (1984) *Biochemistry 23*, 1371.
Schmidt, P., Schawalbe, H., & Griesinger, C. (1994) ENC 35th Experimental Nuclear Magnetic Resonance Conference, poster WP114.
Schmitz, U., Zon, G., & James, T. L. (1990) *Biochemistry 29*, 2357–2368.
Schmitz, U., Sethson, I., Egan, W. M., & James, T. L. (1992) *J. Mol. Biol. 227*, 510–531.
Schmitz, U., Ulyanov, N. B., Kumar, A., & James, T. L. (1993) *J. Mol. Biol. 234*, 373–389.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Stawinski, J., Stromberg, R., Thelin, M., & Westman, E. (1988) *Nucleic Acids Res. 16*, 9285–9298.
Stolarsky, R., Egan, W., & James, T. L. (1992) *Biochemistry 31*, 7027–7042.
Suzuki, E., Pattabiraman, N., Zon, G., & James, T. L. (1986) *Biochemistry 25*, 6854–6865.
Thomas, P. D., Basus, V. J., & James, T. L. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 1237–1241.
Uhlmann, E., & Peyman, A. (1990) *Chem. Rev. 90*, 543–584.
Ulyanov, N. B., Schmitz, U., & James, T. L. (1993) *J. Biomol. NMR 2*, 547–568.
Ulyanov, N. B., Schmitz, U., Kumar, A., & James, T. L. (1994) *Biophys. J.* (submitted).
Uznanski, B., Grajkowski, A., & Wilk, A. (1989) *Nucleic Acids Res. 17*, 4863–4971.
Varmus, H. (1988) *Science 240*, 1427–1435.
Wang, A. C., Kim, S. G., Flynn, P. F., Chou, S.-H., Orban, J., & Reid, B. (1992) *Biochemistry 31*, 3940–3946.
Weisz, K., Shafer, R. H., Egan, W., & James, T. L. (1992) *Biochemistry 31*, 7477–7487.
Weisz, K., Shafer, R. H., Egan, W., & James, T. L. (1994) *Biochemistry 33*, 354–366.
Widmer, H., & Wüthrich, K. (1986) *J. Magn. Reson. 70*, 270–279.
Widmer, H., & Wüthrich, K. (1987) *J. Magn. Reson. 74*, 331–336.
Wijmenga, S. S., Mooren, M. M. W., & Hilbers, C. W. (1993) in *NMR in Macromolecules* (Robert, G. C., Ed.) pp 217–288, IRL Press, Oxford.
Yang, W., Hendrickson, W. A., Crouch, R. J., & Satow, Y. (1990) *Science 249*, 1398–1405.
Zon, G. (1988) *Pharm. Res. 5*, 539–549.
Zon, G., & Stec, W. J. (1991) in *Oligonucleotides and Analogues: A Practical Approach* (Eckstein, F., Ed.) pp 87–108, IRL Press, London.