Solution Structure of an RNA·2′-*O*-Methylated RNA Hybrid Duplex Containing an RNA·DNA Hybrid Segment at the Center[†,‡]

Tomoko Nishizaki,[§] Shigenori Iwai,[§,||] Eiko Ohtsuka,[§] and Haruki Nakamura*[,||]

Faculty of Pharmaceutical Sciences, Hokkaido University, Kita-ku, Sapporo 060, Japan, and Biomolecular Engineering Research Institute, 6-2-3 Furuedai, Suita, Osaka 565, Japan

Received September 12, 1996; Revised Manuscript Received December 27, 1996[⊗]

ABSTRACT: The solution structure of an RNA·2′-*O*-methylated RNA hybrid duplex containing an RNA·DNA hybrid segment at its center, (ggagaugac)·($G_mU_mC_mATCTC_mC_m$), where lowercase letters, capital letters, and capital letters with the subscript m are RNA, DNA, and 2′-*O*-methylated RNA residues, respectively, was determined by observing the NMR spectra and performing the full relaxation matrix refinement. The 2′-*O*-methylation gives several characteristic features to oligoribonucleotides. In addition, this hybrid duplex is cleaved at a specific position by *Escherichia coli* ribonuclease HI, and so the role of the tertiary structure during the substrate recognition by the enzyme is of interest. The NOE connectivities among the proton resonances revealed that the duplex was a right handed helix. The 2′-*O*-methylated RNA segments had a typical C3′-endo conformation, and the 2′-*O*-methyl groups were directed to the minor groove of this duplex, taking the torsion angles ϕ (C1′−C2′−O2′−CH_3) that were all gauche(+). The DNA residues in the central RNA·DNA hybrid duplex formed the C3′-endo conformation, except for the middle thymine residue. No remarkable structural discontinuities were observed around the junction sites at either the 5′- or 3′-end of the DNA. The overall structure was close to the typical A-form duplex.

It is well-known that 2′-*O*-methylation of RNA residues generally increases the thermal stability of RNA tertiary structures. For example, 2′-*O*-methylation at position 54 of the yeast phenylalanine tRNA can stabilize the intrinsic tertiary structure (Agris et al., 1973), and a natural modification was found at this position in extremely thermophilic archaebacteria (Edmonds et al., 1987). Since an RNA·2′-*O*-methylribonucleotide duplex has greater thermal stability than the corresponding RNA·DNA hybrid duplex (Inoue et al., 1987a; Lesnik et al., 1993), the 2′-*O*-methylated oligoribonucleotides are also of interest as antisense probes (Lamond & Sproat, 1993). The substituted methyls inhibit the hydrolysis of the oligomers, both *in vitro* and *in vivo*. For example, the cleavage reaction of RNase H[1] is inhibited by a 2′-*O*-methylated RNA residue (Inoue et al., 1987b). As of yet, the mechanism of the hydrolysis inhibition by the substituted methyls has not been determined.

Escherichia coli ribonuclease HI (*E. coli* RNase HI) hydrolyzes the RNA strand in an RNA·DNA hybrid duplex. Its specificity is independent of the base sequence, but it cleaves neither DNA nor RNA duplexes (Crouch, 1990; Kanaya & Ikehara, 1995). An RNA·DNA hybrid duplex containing partial 2′-*O*-methylated oligoribonucleotides was subjected to a series of studies of the enzymatic mechanism of the protein, since this oligomer limits the digestion site to a specific position (Inoue et al., 1987b; Nakamura et al., 1991; Uchiyama et al., 1994; Iwai et al., 1995, 1996). It has been suggested that the structure of the 2′-*O*-methylated RNA and the steric hindrance at the interface between the enzyme and the hybrid duplex could inhibit random digestion by the enzyme, and thereby yield specific cleavage.

Thus far, several spectroscopic studies of 2′-*O*-methylated RNA oligomers have been performed (Bobst et al., 1969; Lesnik et al., 1993; Cummins et al., 1995), but there have been only a few precise structural studies of these oligomers. A fiber structural analysis of single-stranded 2′-*O*-methylated oligomers was reported and determined that the conformations of the 2′-*O*-methylated RNA residues were C3′-endo, and the 2′-*O*- methyl groups were located on the outside of the helix (Leslie & Arnott, 1978). Blommers et al. (1994) analyzed the 2′-*O*-methylated RNA·DNA hybrid duplex by NMR, where the phosphate backbone of the DNA strand was modified by amido-3 at its center. In their structure, the sugar conformations of the 2′-*O*-methylated RNA strand were predominantly N-type, and the 2′-*O*-methyl groups were positioned at the surface of the minor groove. Popenda et al. (1995) observed the NMR spectra of the 2′-*O*-methylated RNA duplex and determined the sugar puckers. No other solution structure of a duplex containing 2′-*O*-methylated segments has been determined yet, except within our previous report (Nakamura et al., 1991).

Here, we would like to address two questions. (i) What is the tertiary structure of the 2′-*O*-methylated oligoribonucleotide segment in a duplex? (ii) How is the structure

[†] This research was supported by a Grant-in-Aid from the Ministry of Education, Science, and Culture, Japan.

[‡] Coordinates of all final structures are deposited at the Brookhaven Protein Data Bank, Brookhaven National Laboratory, Upton, NY 11973. The ID code is 1NAO.

* To whom correspondence should be addressed.

[§] Hokkaido University.

[||] Biomolecular Engineering Research Institute.

[⊗] Abstract published in *Advance ACS Abstracts,* February 1, 1997.

[1] Abbreviations: RNase HI, ribonuclease HI; NMR, nuclear magnetic resonance; TBAF, tetrabutylammonium fluoride; THF, tetrahydrofuran; TEAA, triethylammonium acetate; NOESY, nuclear Overhauser effect spectroscopy; TPPI, time-proportional phase incrementation; E-COSY, exclusive correlated spectroscopy; DQF-COSY, double-quantum filtered correlated spectroscopy; TOCSY, total correlated spectroscopy; rmsd, root mean square deviation; rMD, restrained molecular dynamics; RR, refined relaxation matrix.

S0006-2960(96)02297-0 CCC: $14.00

of the central RNA·DNA hybrid duplex influenced by 2′-*O*-methylated RNA residues at both the 5′- and 3′-ends? In order to answer these points, we prepared a hybrid duplex (ggagaugac)·($G_mU_mC_mATCTC_mC_m$), where lowercase letters, capital letters, and capital letters with the subscript m are RNA, DNA, and 2′-*O*-methylated RNA residues, respectively. The solution structure was determined by NMR measurements and full relaxation matrix refinement. In addition, *E. coli* RNase HI recognizes this duplex as its substrate and cleaves at a specific position (Inoue et al., 1987b; Nakamura et al., 1991), because of the central RNA·DNA hybrid region. The mechanism of the substrate recognition by the enzyme is discussed, on the basis of the resultant conformation of the duplex.

MATERIALS AND METHODS

Oligonucleotide Synthesis

(ggagaugac). Solid phase oligonucleotide synthesis was carried out on an Applied Biosystems model 394 DNA/RNA synthesizer using standard β-cyanoethyl chemistry, according to the manufacturer's protocol. The RNA phosphoramidites and the 3′-terminal nucleoside on the controlled pore glass support (CPG) used in the synthesis were purchased from MilliGen/Biosearch. The scale of the synthesis was 10 μmol. The final detritylation was not performed on the synthesizer, and the oligonucleotide was cleaved from the support and was deprotected by treatment with 10 mL of ammonium hydroxide/ethanol (3:1, v:v) for 16 h at 55 °C. The mixture was evaporated to dryness. The residue was treated with 5 mL of 1 M tetrabutylammonium fluoride (TBAF) in tetrahydrofuran (THF) for 24 h at room temperature. The solution was concentrated to a yellow syrup, and the reaction was quenched by adding 3 mL of 2 M triethylammonium acetate (TEAA, pH 7.0). After deprotection and concentration, this mixture was applied to a C18 silica gel column, and the product was eluted with a gradient of 5 to 40% CH_3-CN in 0.1 M TEAA. The elution of the oligomer was monitored at 254 nm by a UV detector, and the appropriate fractions were collected and concentrated to dryness. The 4,4′-dimethoxytrityl group was removed with 0.01 N HCl, and the solution was neutralized with 0.1 M NH_4OH and evaporated. The residue was dissolved in 1 mL of H_2O and was passed through a Millipore filter (0.45 μm pore size). Purification was performed by reverse phase HPLC with a linear gradient of 9 to 13% CH_3CN in 0.1 M TEAA. The final yield of the oligomer was 3.6 μmol (36%).

($G_mU_mC_mATCTC_mC_m$). Protected 2′-*O*-methylribonucleosides were prepared as described (Inoue et al., 1987b) and were phosphitylated with 2-cyanoethyl *N*,*N*-diisopropyl chlorophosphoramidite (Shinha et al., 1984; Shibahara et al., 1987). The oligonucleotide synthesis was performed in the same manner as for the former one. The oligomer was cleaved from the support and was deprotected by treatment with 10 mL of ammonium hydroxide for 12 h at 50 °C. The reagent was removed by evaporation to dryness. After deprotection and concentration, this mixture was applied to a C18 silica gel column, and the product was eluted with a gradient of 5 to 40% CH_3CN in 0.1 M TEAA. The elution of the oligomer was monitored at 254 nm by a UV detector, and the appropriate fractions were collected and concentrated to dryness. Acetic acid (80%, 10 mL) was added to this residue. After 20 min, the solution was evaporated *in vacuo* and was coevaporated with H_2O three times. The residue was dissolved in water and was washed with ethyl acetate. The aqueous layer was evaporated to dryness, and the residue was dissolved in 1 mL of H_2O and was passed through a Millipore filter (0.45 μm pore size). The purification was performed in the same way. The final yield was 3.2 μmol (32%).

NMR Sample Preparation

After purification, 1.6 μmol of both strands were annealed in 400 μL of buffer containing 50 mM sodium phosphate, 100 mM NaCl, and 3 mM ethylenediaminetetraacetic acid (pH 6.95). The solution was applied to a Sephadex G75 column at 5 °C to remove the single-stranded species. The appropriate fractions were concentrated and finally were converted to the sodium form on an AG50WX8 cation exchange resin. The concentrated oligomers were dissolved at a concentration of 3 mM in a buffer containing 50 mM sodium phosphate, 100 mM NaCl, 3 mM ethylenediaminetetraacetic acid, and 10% D_2O (pH 6.95). The solution was heated at 80 °C for 5 min and was gradually cooled to room temperature. After the experiments in H_2O, the sample was lyophilized with D_2O three times and was dissolved in 400 μL of D_2O.

NMR Experiments

All NMR experiments were carried out on a Bruker AMX-600 spectrometer (600.13 MHz for protons; other parameters are described in the text). The sample temperature was set at 313 K for all experiments in D_2O. Three nuclear Overhauser spectroscopy (NOESY) spectra of the sample in D_2O, with mixing times of 45, 90, and 150 ms, were measured within 4 days, with the sample tube kept in the spectrometer. The NOESY data were acquired in the phase sensitive mode with time-proportional phase incrementation (TPPI) at 1024 complex points in t_2 and 512 points in t_1, utilizing a relaxation delay of 5 s to ensure adequate relaxation of the RNA protons (Salazar et al., 1993b). The spectrum width was 6024.1 Hz. The total correlated spectroscopy (TOCSY) and the double-quantum filtered (DQF)-COSY (Piantini et al., 1982) were acquired at 1024 complex points in t_2 and 512 points in t_1, utilizing a relaxation delay of 5 s at a spectral width of 6024.1 Hz. The exclusive correlated spectroscopy (E-COSY) spectrum (Griesinger et al., 1986, 1987) was acquired with 240 scans at 1024 complex points in t_2 and 800 points in t_1, utilizing a relaxation delay of 2.5 s at a spectral width of 5434.8 Hz, where ^{31}P decoupling was employed using a ^{31}P 180° pulse at the center of t_1 and the WALTZ decoupling sequence during the acquisition period (Shaka et al., 1983), so as to be free from the passive J_{HP} terms. One-dimensional spectra and the NOESY spectra in H_2O were acquired with a 1-1 spin echo pulse sequence for the solvent suppression (Sklenár et al., 1987) from 278 to 313 K. These experiments were performed in the phase sensitive mode at the TPPI at 1024 complex points in *t*2 and 512 points in *t*1, utilizing a relaxation delay of 1 s. The spectral width was 15 151.51 Hz. The NMR data were processed using the programs Felix 2.3 (Hare Research, Inc.) and UXNMR (Bruker). For processing of the NOESY data sets acquired in H_2O and D_2O, a sine bell function shifted by 45° was used as the window function in both dimensions. Zero filling, followed by Fourier transformation and baseline correction with a fifth-

order polynomial, gave the spectra with 2048 real points in both dimensions, without any baseline distortion. The cross-peak intensities were obtained by using the measure volume tool of the Felix software. The spectra of E-COSY, TOCSY, and DQF-COSY were processed with a sine bell function in both dimensions. The final spectral point resolution of E-COSY was 2.65 Hz/point for both dimensions.

Determination of Distances and Other Constraints

The initial buildup rates (R_{ij}) in D_2O were calculated by the cross-peak volumes at the mixing time of 45 ms. These rates were converted to distances (r_{ij}) using the cytosine H5−H6 distance of 2.5 Å (for all protons except those of the methyl groups) and the thymine H6−CH_3 distance of 2.9 Å (for the methyl protons) as the reference distances r_{ref}) with the relationship $r_{ij} = r_{ref}(R_{ref}/R_{ij})^{1/6}$. These distances were given with the upper and lower bounds of 0.3 and 0.4 Å, respectively (Salazar et al., 1994; Schweitzer et al., 1994). Additionally, the following restraints were used in this study. *(i) The first was backbone torsion angles.* In order to preserve the right-handed characteristics of the duplex, the α, β, ϵ, γ, and ζ backbone torsion angles were restrained to a range covering both the right-handed A- and B-form duplexes (Gronenborn & Clore, 1989). *(ii) The second was hydrogen bond distances in base pairs.* To maintain conservative Watson−Crick base pairing, distance restraints between the bases were used. The values were adopted from the X-ray analysis. *(iii) The third was base pair planarity*. All carbon and nitrogen atoms in the A·T, A·U, and G·C base pairs were restrained to be coplanar, using the energy term planarity in X-PLOR.

Structure Determination

The initial A- and B-form duplex structures were generated using the SYBYL molecular modeling system (Tripos Inc.) and were designated as Ini-A and Ini-B. In these initial structures, four different torsion angles of C1′−C2′−O2′−CH_3 (ϕ) were set at 0, 90, 180, and 270°, and only the initial structures with the torsion angle at 90° were able to produce converged structures. All subsequent energy minimizations were performed on a workstation (Silicon Graphics Indigo R400XS24) using X-PLOR 3.1 (Brünger, 1993). All of the initial and the refined structures were displayed by the SYBYL molecular modeling system. The nucleic acid parameters and the force constants used in the calculations were taken from parallnewhdg.dna in X-PLOR 3.1, and the added bonds and angles of the 2′-*O*-methyl were from parnahg.trn. The phosphate charge was reduced to −0.32 e (Tidor et al., 1982). The 600-step conjugate gradient minimization was performed on the system before initiating the molecular dynamics (MD) calculation without the electrostatics and hydrogen bonding energy terms. The hard sphere potential was used for the van der Waals repulsion. A switching function between 7.5 and 8.5 Å was used to allow all of the pairwise nonbonded interactions to progress smoothly to zero. A 600-step conjugate gradient minimization was performed on the system before initiating the MD calculation. The force constant of the distance restraints was set to 50 kcal mol^{-1} $Å^{-2}$. The force constant of the dihedral angle restraints was set to 200 kcal mol^{-1} rad^{-2}. These constants were maintained throughout all of the energy minimization steps. The system was maintained at 100 K during the first 1 ps MD procedure. The subsequent MD calculations were performed with the Lennard-Jones potential for the van der Waals potential. The electrostatics and the hydrogen bonding energies were taken into account. A switching function between 10.5 and 11.5 Å was used to make all pairwise nonbonded interactions. The system was heated to 900 K, gradually cooled during the 20 ps MD procedure, and maintained at 200 K during the final 10 ps MD. The coordinates were averaged and were subjected to 600 cycles of restraint energy minimization to obtain the final structures, rMD-A and rMD-B, corresponding to the initial structures, Ini-A and Ini-B, respectively.

Full Relaxation Matrix Refinement

At the final stage of the structure determination, direct NOE refinements (Yip & Case, 1989) of the rMD-A and rMD-B structures were carried out using the RELAX option of X-PLOR (Nilges et al., 1991; Schweitzer et al., 1994). This option directly fits the coordinates of the input structures to the experimental volumes. From each of the 45, 90, and 150 ms NOESY data sets, 256 experimental volumes were taken. All volumes were given with a relative error of 10% estimated from the symmetric peaks (Schweitzer et al., 1994). The theoretical volumes were calculated from a total relaxation matrix of the coordinate set. For these calculations, a uniform isotropic correlation time (4.24 ns) was used (Mujeeb et al., 1994). With restraints of K_{NOE} (experiment in D_2O) = 0, K_{NOE} (H bond) = 25, K_{CIDH} = 0, and K_{relax} = 1000, rMD-A and rMD-B were refined by the dynamical simulated annealing protocol. In this procedure, the initial velocities were chosen from a Maxwellian distribution at 1000 K. The temperature of the system was then reduced by 25 K for 27 further steps of dynamics. This was repeated until a final temperature of 75 K was attained. During this procedure, K_{NOE} and K_{CIDH} were kept constant. The final step consists of 90 steps of energy minimization, and the full nonbonded interaction terms were included in the calculations. The refined structures starting from rMD-A and rMD-B were designated as RR-A and RR-B, respectively. The *R* factors were calculated as the weighted average of the absolute value of the difference between $^1/_6$ of the power of the observed and calculated intensities (James, 1991; Thomas et al., 1991; Brünger, 1993).

Analysis of Helical Parameters

The helical parameters of the duplex were analyzed with the program *RNA*, which was kindly provided by Dr. M. Babcock and Dr. W. S. Olson, Rutgers University (Babcock et al., 1994).

RESULTS

The base sequence of the duplex was chosen from that used in our previous study of the recognition and catalytic mechanisms of *E. coli* RNase HI (Uchiyama et al., 1994). The numbering of the duplexes is as follows.

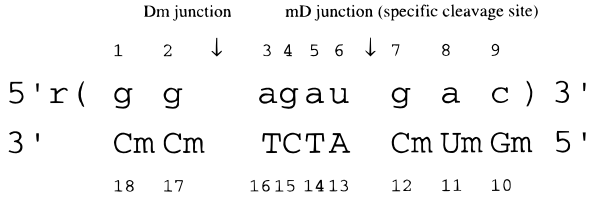

Table 1: Chemical Shifts (Parts Per Million) of the Protons of (ggagaugac)·($G_mU_mC_mATCTC_mC_m$)

residue	H8/H6	H2/CH_3/H5	H1′	H2′	H2″	H3′	H4′	2′-*O*-methyl	imino[a]
1g	7.81		5.59	4.75		4.43	4.38		12.72
2g	7.34		5.80	4.49		4.76	4.48		12.46
3a	7.68	7.57	5.94	4.74		4.60	4.49		
4g	7.07		5.54	4.59		4.47	4.38		12.48
5a	7.74	7.78	6.01	4.56		4.50	4.11		
6u	7.42	5.07	5.43	4.45		4.57	4.07		13.72
7g	7.66		5.78	4.52		4.63	4.18		11.93
8a	7.83	7.83	5.91	4.45		4.60	4.14		
9c	7.28	5.25	5.57	4.15		4.00	4.16		
$10G_m$	7.74		5.74	4.44		4.45	4.05	3.73	12.24
$11U_m$	8.02	5.03	5.89	4.44		4.57	3.73	3.74	14.35
$12C_m$	7.88	5.69	5.84	4.31		4.60	4.04	3.72	
13A	7.99	7.28	6.21	2.55	2.75	4.67	4.35		
14T	7.58	1.13	5.98	2.48	2.57	4.78	4.25		13.89
15C	7.55	5.49	5.87	2.33	2.58	4.57	4.25		
16T	7.71	1.47	5.92	2.55	2.55	4.79	4.22		13.72
$17C_m$	7.78	5.60	5.85	3.91		4.52	4.38	3.68	
$18C_m$	7.70	5.47	5.60	3.70		4.20	3.90	3.57	

[a] Chemical shifts of the imino protons were determined at 283 K. All the other chemical shifts of the nonexchangeable protons were assigned at 313 K.

Table 2: *J* Coupling Constants (Hertz) of Deoxy Residues of the Duplex (ggagaugac)·($G_mU_mC_mATCTC_mC_m$)

residue	$J_{1'2'}$	$J_{1'2''}$	$\Sigma J_{1'}$[a]	$\Sigma J_{3'}$[a]
13A	5.29	10.55	13.28	21.23
14T	7.92	7.85	14.57	18.64
15C	5.29	7.94	13.27	b
16T	b	b	b	b

[a] The values of $\Sigma J_{1'}$ ($=J_{1'2'} + J_{1'2''}$) and $\Sigma J_{3'}$ ($=J_{2'3'} + J_{2''3'} + J_{3'4'}$) were derived from the apparent splitting of the outer peaks of cross-peaks. [b] Could not be determined because of spectral overlap.

Here, at the mD junction between 6u and 7g, the RNA strand is specifically cleaved by *E. coli* RNase HI (Inoue et al., 1987b; Nakamura et al., 1991).

Proton Resonance Assignments. The exchangeable proton resonances were assigned using the NOESY spectrum in a sequential manner. The thymine H3 imino protons gave a NOESY cross-peak with the H2 proton of the base-paired adenine. Including the terminal iminos, all of the imino proton resonances were observed. The assignment of the terminal residues was performed using the one-dimensional spectra at varying temperatures. The imino signals of the terminal residues disappeared when the temperature was raised to 308 K. The nonexchangeable protons were assigned using the DQF-COSY, TOCSY, and NOESY spectra in a sequential manner. The sequential connectivities of the base protons with H1′, H2′, and H2″ could be traced. The proton chemical shifts of these two duplexes at 313 K are shown in Table 1, except for the imino protons whose chemical shifts are the values at 283 K. The chemical shifts of the H2′ and H2″ protons of 16T at the Dm junction completely overlapped, although those of 13A at the mD junction were well separated.

Determination of Scalar Coupling Constants in E-COSY. The *J* coupling constants were derived from the E-COSY experiments with the ^{31}P decoupling in t_1 and t_2 periods. Those for H1′−H2′, and H1′−H2″ of the DNA residues are shown in Table 2. All of the DNA residues, except for 14T, showed small H1′−H2′ cross-peaks in both DQF- and E-COSY (smaller than 6 Hz), as shown in Figure 1 and Table 2. In contrast, the H1′−H2′ coupling constants ($J_{1'2'}$) of 14T were relatively large. The $J_{1'2'}$ and $J_{1'2''}$ values of 16T at the Dm junction could not be determined, since the cross-peaks of H1′−H2′ and H1′−H2″ of 16T overlapped completely (Figure 1). $\Sigma J_{1'}$ and $\Sigma J_{3'}$ are also shown in Table 2. Those coupling constants indicate that the DNA residues do not maintain typical C2′-endo conformations. All of the H1′−H2′ coupling constants ($J_{1'2'}$) of the RNA and the 2′-*O*-methyl RNA residues were smaller than 5 Hz and were undetectable, suggesting that the RNA and 2′-*O*-methyl RNA residues have C3′-endo conformations.

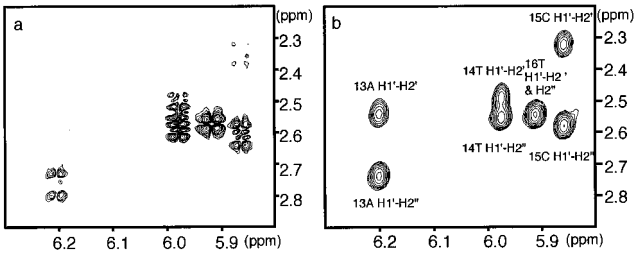

FIGURE 1: DQF-COSY (a) and NOESY (b) spectra in the H1′−H2′ and H2″ region from 13A to 16T. The DQF-COSY spectrum was acquired in the phase sensitive mode with TPPI at 1024 complex points in t_2 and 512 points in t_1. Each FID was collected with 96 scans. The NOESY spectrum was acquired in the same phase sensitive mode and with the same data points as the DQF-COSY spectrum with a mixing time of 45 ms. Each FID was collected with 48 scans. In both measurements, the HDO peaks were irradiated during relaxation delay of 5 s, and the spectrum widths were 6024.1 Hz.

Structure Calculations. The right-handed nature of both duplexes was evident in the NOESY spectra. The NOE connectivities of the imino protons between adjacent bases of the same strands showed normal base stacking throughout the duplexes, except for the termini. To evaluate the structural features of this duplex more quantitatively, structure calculations were carried out. Five calculations were started from the canonical A-form (Ini-A) and another five calculations from the B-form (Ini-B) initial conformations, respectively, with different random seeds for the assignments of the initial velocities of individual atoms. Each of the 10 structures was converted into a lower-energy conformation without any NOE violations greater than 0.5 Å and any dihedral violations above 0.5°. The root mean square deviation (rmsd) values between the final conformations,

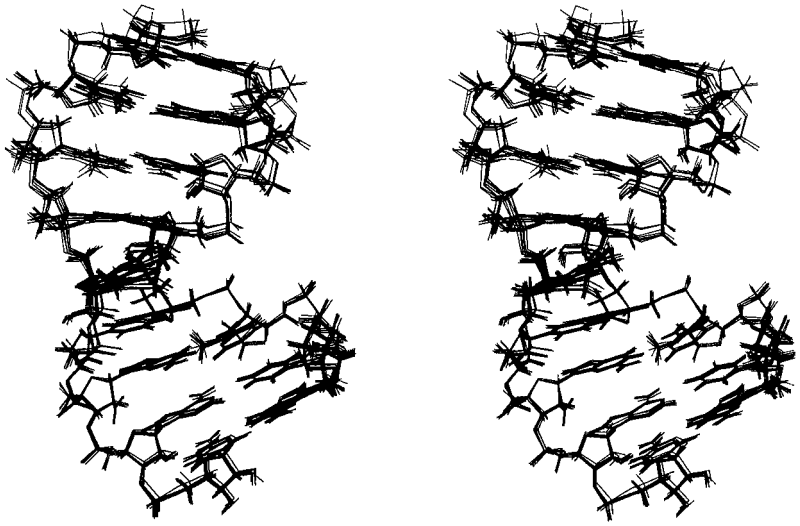

FIGURE 2: Stereoview of the 10 superimposed structures: RR-A1, -A2, -A3, -A4, -A5, -B1, -B2, -B3, -B4, and -B5.

Table 3: Atomic rmsd of All Atoms (Angstroms) and *R* Factors of the rMD-A and -B Structures

	rMD-A					rMD-B					
	A1	A2	A3	A4	A5	B1	B2	B3	B4	B5	*R* factor (%)
rMD-A											
A1											9.59
A2	0.41										9.25
A3	0.15	0.38									9.61
A4	0.50	0.66	0.50								9.49
A5	0.56	0.61	0.53	0.22							9.48
rMD-B											
B1	0.44	0.29	0.41	0.67	0.54						9.48
B2	0.45	0.26	0.46	0.71	0.66	0.29					9.40
B3	0.50	0.34	0.50	0.74	0.70	0.32	0.26				9.42
B4	0.54	0.41	0.55	0.76	0.72	0.43	0.35	0.36			9.31
B5	0.61	0.45	0.60	0.77	0.74	0.47	0.40	0.44	0.54		9.34
Ini-A	1.71	1.75	1.73	1.62	1.67	1.75	1.83	1.81	1.62	1.58	
Ini-B	4.05	4.17	4.05	4.17	4.05	4.19	4.12	4.16	4.77	3.90	

rMD-A from Ini-A and rMD-B from Ini-B, were less than 1.0 Å for all heavy atoms (Table 3).

Full Relaxation Matrix Refinement. In order to refine the obtained structures by taking the spin diffusion into account, the full relaxation matrix refinement with the structure calculation was carried out, without using any dihedral angle restraints. Each of the 256 NOE intensities from the three different mixing times (45, 90, and 150 ms) was used for the refinement, simultaneously. As shown in Figure 2 and Table 4, the structures after the relaxation matrix refinement (RR-A and RR-B) were much closer to Ini-A than to Ini-B. The rmsd values between the structures before and after the full relaxation matrix refinement and the *R* factors of the refined structures are shown in Table 4.

Analysis of the Final Structures. In Figure 3, the averages and the deviations of the backbone torsion angles for all 10 structures of RR-A and RR-B are shown. All of the α, β, γ, δ, ϵ, and χ values are nearly those of a typical A-form duplex. The remarkable feature was the fact that the DNA residues did not show any evidence of the C2′-endo conformation. The pseudorotation phase angles indicate that all of the sugars in the DNA residues have the C3′-endo conformation, except 14T, consistent with the values of $J_{1'2'}$.

The orientation of the 2′-*O*-methyl residues is described by using the torsion angle ϕ (C1′−C2′−O2′−CH_3). As shown in Figure 4, every observed distance between H2′ and the 2′-*O*-methyl (Me) protons in the same residue was very close, about 2.9 Å. In contrast, each distance between H4′ and the Me protons in the same residue was about 4 Å. These are consistent with the final refined structures, where the torsion angles ϕ were around 80° as shown in Figure 5.

The helical parameters are shown in Figure 6. This duplex has two heterogeneous junction sites, which are termed the mD junction and the Dm junction, respectively, as described above. At the junctions between the DNA and the RNA, no distinct discontinuities were observed between 2g and 3a or between 6u and 7g.

DISCUSSION

The solution structure of an oligonucleotide duplex, which contains an RNA·DNA hybrid sandwiched by two RNA·2′-*O*-methylated RNA hybrids, has been determined by NMR measurements and full relaxation matrix refinement. This duplex is cleaved by *E. coli* RNase HI only at a specific position, and so the precise structural features are important for a detailed understanding of the enzymatic reaction of the protein. We first answer the questions mentioned at the beginning of this paper and then discuss the recognition mechanism of *E. coli* RNase HI.

Conformations of RNA and 2′-O-Methylated RNA. The $J_{1'2'}$ values of the riboses and the 2′-*O*-methylated riboses in the current duplex were too small to be detectable, and every ribose of the final refined duplex structure had the C3′-endo

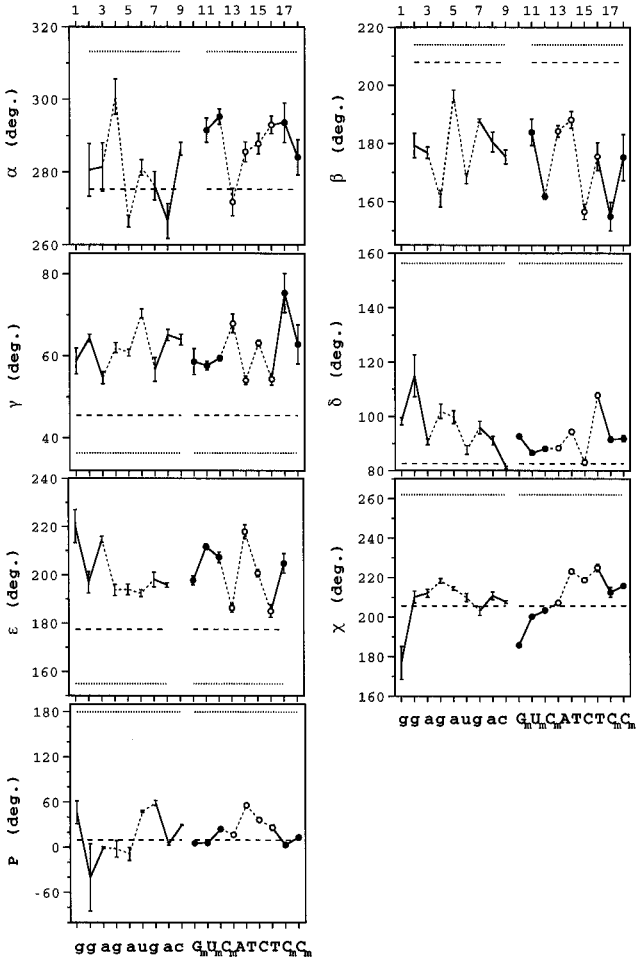

FIGURE 3: Averages and the deviations of the dihedral and the pseudorotation phase angles of the final structures. The RNA·2′-*O*-methylated RNA duplex and the RNA·DNA hybrid duplex region are indicated by lines and dotted lines, respectively. The 2′-*O*-methylated RNA (●) and DNA (○) regions are marked. The values for the standard A-form (- - -) and B-form (···) DNA (Arnott & Hukins, 1972, 1973) are also shown.

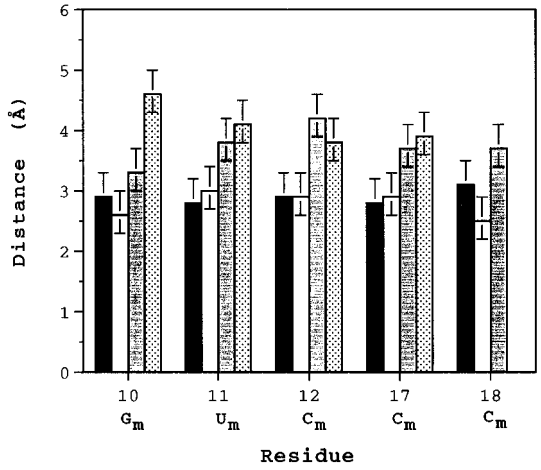

FIGURE 4: Observed distances between the sugar protons and the 2′-*O*-methyl protons (Me), estimated from the initial buildup rates obtained by the cross-peak volumes of the NOESY spectra. Black, white, dark-hatched, and light-hatched bars indicate the distances between H1′ and Me, between H2′ and Me, between H3′ and Me, and between H4′ and Me, respectively. The cross-peak between H4′ and Me in $18C_m$ could not be observed because of spectral overlap. The experimental errors are indicated in the figure.

conformation. They are consistent with the sugar pucker conformation of the 2′-*O*-methyl nucleotides reported by

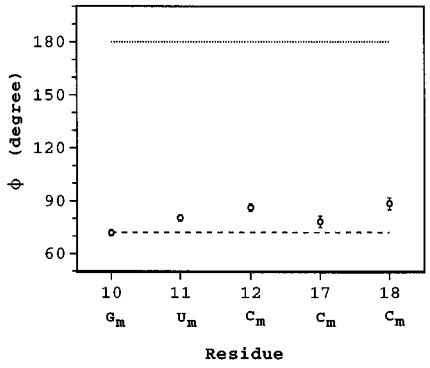

FIGURE 5: Averages and the deviations of the ϕ angles (C1′−C2′−O2′−CH_3) in the final 10 structures. The trans and gauche orientations expected in the 2′-*O*-methyl monomers (Kawai et al., 1992) are indicated by the lines (··· and - - -, respectively).

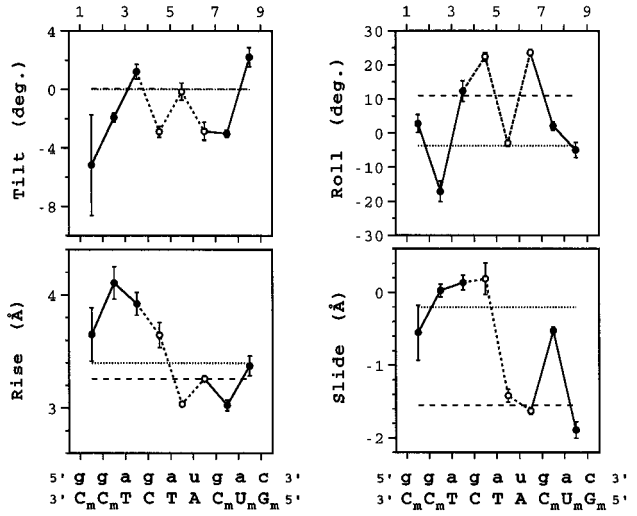

FIGURE 6: Comparison of helical parameters. The averages of the 10 final structures in the RNA·2′-*O*-methylated RNA duplex (●), and the RNA·DNA hybrid duplex regions (○), are indicated along with those of standard A-form (- - -) and B-form (···) DNA.

Popenda et al. (1995). Kawai et al. (1992) examined the conformations of the 2′-*O*-methylated uridines in the monomer and the dimer units by NMR and by energy calculations. They also found that the 3′-phosphorylated 2′-*O*-methyluridine preferred to take the C3′-endo sugar conformation, because of steric repulsion between the 2-carbonyl group of the base, the 2′-*O*-methyl group, and the 3′-phosphate group.

The orientation of the 2′-*O*-methyls in the mononucleotides was estimated by an energy minimization study (Kawai et al., 1992). In the case of the monomeric nucleotides, the 2′-*O*-methyl groups were found to take two orientations, with torsion angles ϕ around 180 and 72°. In the single-stranded poly(C_m) in the fiber form, Leslie and Arnott (1978) reported that the 2′-*O*-methyl groups had trans torsion angles near 180°. Recently, Blommers et al. (1994) estimated that the 2′-*O*-methyl groups relative to the A-form duplex had the orientation with the gauche(+) torsion angles near 60°, by examining the H1′−Me cross-peak intensities in the NOESY spectrum.

As shown in Figure 5, the 2′-*O*-methyl groups in our refined duplex structures take only the orientations with the gauche(+) torsion angles. The distances between the intra- and inter-residue protons and the average position of the three methyl protons of the 2′-*O*-methyls on the typical A-form duplex were calculated as a function of the torsion angle ϕ.

The intraresidue distances between H4′ and Me and between H2′ and Me are largely dependent upon ϕ. The strengths of the current NOE cross peaks indicated in Figure 4 agree well with the observed ϕ values in the final calculated conformation (Figure 5). Therefore, we conclude that the 2′-*O*-methyl groups form the gauche(+) conformation in our duplex.

The conformation of the 2′-*O*-methyl should originate from the steric constraint between the methyl group and the adjacent 3′-phosphate group, as mentioned by Kawai et al. (1992). They showed that only the combination of the 3′-endo sugar conformation and the gauche(+) orientation of the 2′-*O*-methyl group could make a stable RNA duplex structure. The gauche(+) orientation with the 3′-endo sugar conformation orients the bulky methyl groups directly toward the minor groove, which is opposite of the intraresidue 3′-phosphate group, and so no steric constraint occurs. With the 2′-endo sugar conformation, any structure of the 2′-*O*-methyl groups would interfere with the base or the 3′-phosphate group.

In our previous study, the ribonucleotides linked to the 3′-end of the DNA segments did not adopt a pure C3′-endo conformation, where no 2′-*O*-methylated RNAs were inserted (Nishizaki et al., 1996). At the Dm junction in the current duplex, a 2′-*O*-methylated ribonucleotide linked to the 3′-end of the DNA segments ($17C_m$) maintains the C3′-endo conformation. The substituted 2′-*O*-methyl group in $17C_m$ may turn the ribose conformation left, as C3′-endo. Another possible reason is that the total structure of the hybrid duplex prefers the C3′-endo conformation at this junction, because the duplex consists of many more riboses than our previous RNA·DNA hybrids (Nishizaki et al., 1996).

Conformations of DNA. Several solution structures of RNA·DNA hybrid duplex segments were reported previously (Salazar et al., 1993a,b, 1994; Gonzáletz et al., 1994, 1995; Zhu et al., 1995; Nishizaki et al., 1996). Their remarkable features were that the sugar conformations of the ribonucleotides were C3′-endo and that those of the deoxyribonucleotides were intermediate between the C2′- and C3′-endo conformations.

However, the current duplex has a very different feature from those results. In fact, the DNA residues, except 14T, have the typical C3′-endo conformation, as do the other RNA and 2′-*O*-methyl RNA residues. Especially, the $J_{1'2'}$ values in 13A and 15C were significantly smaller than that of typical B-form DNA (Rinkel & Altona, 1987), and they were also smaller than the corresponding $J_{1'2''}$ values, as shown in Table 2. The *J* coupling constants in 16T could not be obtained because of overlapping of the cross-peaks. However, from the strong intraresidue NOE signals between H6 and H3′, and the final refined structures shown in Figures 2 and 3, the deoxyribose of 16T is considered to take the C3′-endo conformation, rather than the C2′-endo conformation.

Only the central residue, 14T, had a $J_{1'2'}$ value larger than those of 13A and 15C, and $J_{1'2'}$ is nearly equal to $J_{1'2''}$, indicating different characteristics from the pure C3′-endo conformation. In fact, the S conformation fraction (%S) of 14T calculated from $\Sigma J_{1'}$ (Rinkel & Altona, 1987) is 81%, if multiple conformations are assumed. In addition, the torsion angles shifted toward those of the B-form duplex at 14T. This means that not only the sugar conformation but also the backbone structure changed at this central thymine.

Thus, the DNA residues near the RNA segments showed more C3′-endo characteristics than the central DNA residue. The conformations of the furanose rings of the DNA residues are strongly influenced by the A-form 2′-*O*-methylated RNA.

Salazar et al. (1993b) and Fedoroff et al. (1993) studied the solution structure of an RNA·DNA hybrid duplex in detail. They concluded that the sugars of the DNA residues have an O4′-endo conformation that is intermediate between the C3′-endo and C2′-endo conformations, on the basis of strong H1′−H4′ NOEs and other features of the coupling constants among the sugar protons. In the current duplex, the distances between H1′ and H4′, as analyzed from the NOE spectra, were all longer than 3 Å, and the coupling constants show the characteristic features of the C3′-endo conformation, except 14T, as discussed above. From the *J* coupling values, only 14T may have an O4′-endo sugar conformation. However, since the distance between H1′ and H4′ of 14T is 3.05 ± 0.35 Å, which is slightly longer than the typical distance of the O4′-endo conformation, the N/S interconversions are not completely denied for this 14T, although a long relaxation delay (5 s) was used to ensure adequate relaxation of the RNA protons (Salazar et al., 1993b).

Structure of the Hybrid Duplex. The overall structure of the current duplex is similar to that of a typical A-form double helix (Figure 2), according to the χ, ϵ, δ, and ζ angles (Figure 3) and the rmsd values from the initial A- and B-forms (Table 4).

As indicated in Figure 6, the hybrid duplex has larger roll angles than those of a typical A-form between 4g and 5a and between 6u and 7g. In contrast, the adjacent roll angle at the middle between 5a and 6u is small enough to keep the entire duplex conformation from bending.

No remarkable structural change was observed at either Dm or mD junction, and a continuous helical structure was maintained. In other RNA·DNA hybrids with junctions between RNA and DNA, it is common to observe a local bent structure at the junction where an RNA residue is linked to the 5′-end of a DNA residue (Salazar et al., 1994; Zhu et al., 1995; Nishizaki et al., 1996). In the current duplex, almost all of the nucleotides have the pure C3′-endo sugar conformation, due to the 2′-*O*-methyl substitution, and the entire duplex is close to the typical A-form. Therefore, it may not be necessary to produce any structural deformation at the mD junction.

Substrate Recognition by E. coli RNase HI. Our duplex is cleaved at one site by *E. coli* RNase HI with a rate similar to that of pure RNA·DNA hybrids (E. Ohtsuka, personal communication; Iwai et al., 1996). The 2′-*O*-methyl substitution restricts the cleavage site. Recent reports suggest that the enzyme recognizes not only the phosphate groups in the RNA strand but also those in the DNA strand. The 2′-*O*-methylated RNA in the DNA strand effectively disrupts the cleavage reaction (Inoue et al., 1987b; Larrouy et al., 1995). According to the current duplex structure, the 2′-*O*-methyls are positioned in the minor groove, and they are near the intraresidue phosphate. The bulky 2′-*O*-methyls should shield the phosphate from the enzyme, to some extent, and thereby disturb the recognition.

So far, *E. coli* RNase HI is considered to bind the minor groove of a hybrid duplex (Nakamura et al., 1991; Fedoroff et al., 1993; Uchiyama et al., 1994). From the solution structure determined by Salazar et al. (1993b), Fedoroff et

Table 4: Atomic rmsd of All Atoms (Angstroms) and *R* Factors of the RR-A and RR-B Structures

	RR-A					RR-B						
	A1	A2	A3	A4	A5	B1	B2	B3	B4	B5	rMD-A/B[a]	*R* factor (%)
RR-A												
A1											0.61	6.90
A2	0.36										0.61	6.79
A3	0.15	0.32									0.62	6.91
A4	0.48	0.63	0.47								0.61	6.94
A5	0.54	0.59	0.54	0.22							0.62	6.79
RR-B												
B1	0.45	0.31	0.42	0.69	0.66						0.61	6.89
B2	0.44	0.30	0.45	0.71	0.69	0.26					0.59	6.88
B3	0.45	0.34	0.46	0.73	0.63	0.27	0.13				0.60	6.96
B4	0.46	0.35	0.47	0.73	0.69	0.30	0.17	0.18			0.60	6.84
B5	0.55	0.43	0.54	0.76	0.74	0.40	0.24	0.31	0.32		0.58	6.97
Ini-A	1.76	1.80	1.77	1.65	1.71	1.79	1.77	1.85	1.80	1.75		
Ini-B	4.10	4.20	4.09	4.15	4.19	4.20	4.26	4.19	4.24	4.25		

[a] RR-A and RR-B structures were compared with the corresponding rMD-A and rMD-B structures, e.g. RR-A1 with rMD-A1, and RR-A2 with rMD-A2, etc.

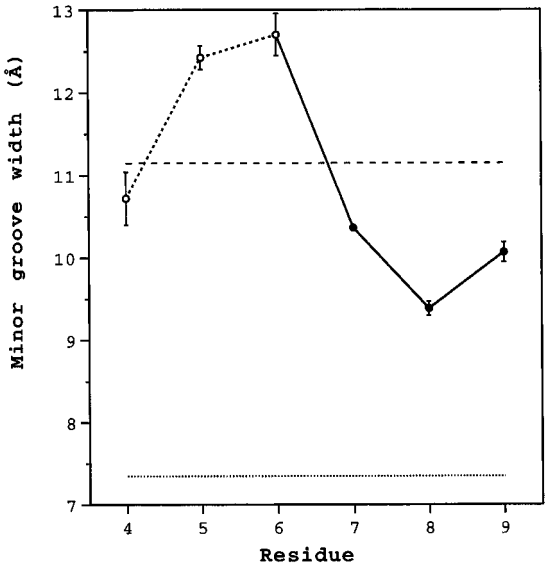

FIGURE 7: Interstrand phosphorus−phosphorus separation (minor groove width) compared to that of standard A-form (- - -) and B-form (···) DNA. The averages in the RNA·2′-*O*-methylated RNA duplex (●), and the RNA·DNA hybrid duplex regions (○), are indicated, along with the deviations.

al. (1993) assumed that its narrow minor groove width (9−10 Å) was crucial for the substrate recognition, as a consequence of the characteristic O4′-endo sugar conformations of the DNA. Our duplex has the RNA·DNA hybrid region, which the enzyme recognizes as its substrate and cleaves at a specific position (Inoue et al., 1987b; Nakamura et al., 1991). As shown in Figure 7, this duplex has a wider minor groove width (11.5−12.5 Å) than that of the RNA·-DNA hybrid duplex reported by Salazar et al. (1993b). This wide minor groove width is very similar to that of a typical A-form RNA duplex. The enzyme can also bind the RNA duplex but cannot cleave it (Oda et al., 1993). Therefore, we do not consider the intrinsic minor groove width to be itself is substantial in the process of substrate recognition.

A specific structural change in either the hybrid duplex or the enzyme could occur during the formation of the enzyme−hybrid complex. In fact, Oda et al. (1993) observed a gradual change in the CD spectrum, corresponding to the deformation of the hybrid, when *E. coli* RNase HI bound the hybrid duplex. In addition, a recent study by Iwai et al. (1996) strongly suggests that the basic protrusion in the enzyme−substrate complex should be closer to the active site than in the free enzyme, as analyzed with several substrates, including 2′-*O*-methylated RNA, that were longer than the current duplex.

In conclusion, from the solution structure of an RNA·2′-*O*-methylated RNA hybrid duplex containing an RNA·DNA hybrid segment at its center, the conformation of the 2′-*O*-methyl groups was clearly determined to be the gauche(+) orientation. The overall structure of the duplex was similar to an A-form RNA duplex, even at the central RNA·DNA hybrid segment, although the hybrid duplex can work as a substrate of *E. coli* RNase HI, whereas double-stranded RNA cannot. No remarkable discontinuities of the helical structure were observed at the junctions between the DNA and the 2′-*O*-methylated RNA.

ACKNOWLEDGMENT

The authors thank Profs. Akira Matsuda and Satoshi Shuto of Hokkaido University for their discussion and Drs. Yasuo Komatsu, Yukiko K. Doi, and Hirokazu Hosoyama of Hokkaido University for their help in the NMR experiments.

REFERENCES

Agris, P. F., Koh, H., & Söll, D. (1973) *Arch. Biochem. Biophys. 154*, 277−283.
Arnott, S., & Hukins, D. W. L. (1972) *Biochem. Biophys. Res. Commun. 47*, 1504−1510.
Arnott, S., & Hukins, D. W. L. (1973) *J. Mol. Biol. 81*, 93−105.
Babcock, M., & Olson, W. K. (1994) *J. Mol. Biol. 237*, 98−124.
Blommers, M. J. J., Pieles, U., & Mesmaeker, A. D. (1994) *Nucleic Acids Res. 22*, 4187−4194.
Bobst, A. M., Cerutti, P. A., & Rottman, F. (1969) *J. Am. Chem. Soc. 91*, 1246−1248.
Brünger, A. T. (1993) *X-PLOR 3.1*: *A system for Crystallography and NMR*, Yale University Press, New Haven, CT.
Crouch, R. J. (1990) *New Biol. 2*, 771−777.
Cummins, L. L., Owens, S. R., Risen, L. M., Lesnik, E. A., Freier, S. M., McGee, D., Guinosso, C. J., & Dan Cook, P. (1995) *Nucleic Acids Res. 23*, 2019−2024.
Edmonds, C. G., Crain, P. F., Hashizume, T., Gupta, R., Stetter, K. O., & McCloskey, J. A. (1987) *J. Chem. Soc., Chem. Commun.*, 909−910.
Fedoroff, O. Y., Salazar, M., & Reid, B. R. (1993) *J. Mol. Biol. 233*, 509−523.
González, C., Stec, W., Kobylanska, A., Hogrefe, R. I., Reynolds, M., & James, T. L. (1994) *Biochemistry 33*, 11062−11072.

González, C., Stec, W., Reynolds, M. A., & James, T. L. (1995) *Biochemistry 34*, 4969−4982.
Griesinger, C., Sørensen, O. W., & Ernst, R. R. (1986) *J. Chem. Phys. 85*, 6837−6852.
Griesinger, C., Sørensen, O. W., & Ernst, R. R. (1987) *J. Magn. Reson. 75*, 474−492.
Gronenborn, A. M., & Clore, G. M. (1989) *Biochemistry 28*, 5978−5984.
Inoue, H., Hayase, Y., Imura, A., Iwai, S., Miura, K., & Ohtsuka, E. (1987a) *Nucleic Acids Res. 15*, 6131−6148.
Inoue, H., Hayase, Y., Iwai, S., & Ohtsuka, E. (1987b) *FEBS Lett. 215*, 327−330.
Iwai, S., Kataoka, S., Wakasa, M., Ohtsuka, E., & Nakamura, H. (1995) *FEBS Lett. 368*, 315−320.
Iwai, S., Wakasa, M., Ohtsuka, E., Kanaya, S., Kidera, A., & Nakamura, H. (1996) *J. Mol. Biol. 263*, 699−706.
James, T. L. (1991) *Curr. Opin. Struct. Biol. 1*, 1042−1053.
Kanaya, S., & Ikehara, M. (1995) in *Subcellular Biochemistry Vol. 24. Proteins: Structure, Function, and Engineering* (Biswas, B. B., & Roy, S., Eds.) pp 377−422, Plenum Press, New York.
Kawai, G., Yamamoto, Y., Kamimura, T., Masegi, T., Sekine, M., Hata, T., Iimori, T., Watanabe, T., Miyazawa, T., & Yokoyama, S. (1992) *Biochemistry 31*, 1040−1046.
Lamond, A. I., & Sproat, B. S. (1993) *FEBS Lett. 325*, 123−127.
Larrouy, B., Boiziau, C., Sproat, B., & Toulmé, J.-J. (1995) *Nucleic Acids Res. 23*, 3434−3440.
Leslie, A. G. W., & Arnott, S. (1978) *J. Mol. Biol. 119*, 399−414.
Lesnik, E. A., Guinosso, C. J., Kawasaki, A. M., Sasmor, H., Zounes, M., Cummins, L. L., Ecker, D. J., Dan Cook, P., & Freier, S. M. (1993) *Biochemistry 32*, 7832−7838.
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.
Nilges, M., Hazabettl, J., Brünger, A. T., & Holak, T. A. (1991) *J. Mol. Biol. 219*, 499−510.
Nishizaki, T., Iwai, S., Ohkubo, T., Kojima, C., Nakamura, H., Kyougoku, Y., & Ohtsuka, E. (1996) *Biochemistry 35*, 4016−4025.
Oda, Y., Iwai, S., Ohtsuka, E., Ishikawa, M., Ikehara, M., & Nakamura, H. (1993) *Nucleic Acids Res. 21*, 4690−4695.
Piantini, U., Sørensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800−6801.
Popenda, M., Milecki, J., Biala, E., & Adamiak, R. W. (1995) *Nucleosides Nucleotides 14*, 983−984.
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., Miller, J. M., Ribeiro, N. S., & Reid, B. R. (1993b) *Biochemistry 32*, 4207−4215.
Salazar, M., Fedoroff, O. Y., Zhu, L., & Reid, B. R. (1994) *J. Mol. Biol. 241*, 440−455.
Schweitzer, B. I., Mikita, T., Kellog, G. W., Gardner, K. H., & Beardsley, G. P. (1994) *Biochemistry 33*, 11460−11475.
Shaka, A. J., Keeler, J., & Freeman, R. (1983) *J. Magn. Reson. 53*, 313−340.
Shibahara, S., Mukai, S., Nishihara, T., Inoue, H., Ohtsuka, E., & Morisawa, H. (1987) *Nucleic Acids Res. 15*, 4403.
Shinha, N. D., Biernat, J., McMaus, J., & Koster, H. (1984) *Nucleic Acids Res. 12*, 4539−4557.
Skelnár V., Brooks, B. R., Zon, G., & Bax,. A. (1987) *FEBS Lett. 216*, 249−252.
Thomas, P. D., Bauses, V. J., & James, T. L. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 1237−1241.
Tidor, B., Irikura, K., Brooks, B. R., & Karpuls, M. (1983) *J. Biomol. Struct. Dyn. 1*, 231−306.
Uchiyama, Y., Miura, Y., Inoue, H., Ohtsuka, E., Ueno, Y., Ikehara, M., & Iwai, S. (1994) *J. Mol. Biol. 243*, 782−791.
Yip, P., & Case, D. A. (1989) *J. Magn. Reson. 83*, 643−648.
Zhu, L., Salazar, M., & Reid, B. R. (1995) *Biochemistry 34*, 2372−2380.

BI962297C