Solution Structures of DNA Duplexes Containing a DNA•RNA Hybrid Region, $d(GG)r(AGAU)d(GAC) • d(GTCATCTCC) \ and \\ d(GGAGA)r(UGAC) • d(GTCATCTCC)^{\dagger,\ddagger}$

Tomoko Nishizaki,[§] Shigenori Iwai,^{§,||} Tadayasu Ohkubo,^{⊥,#} Chojiro Kojima,[▽] Haruki Nakamura,[⊥] Yoshimasa Kyogoku,[▽] and Eiko Ohtsuka*,[§]

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

Received August 21, 1995; Revised Manuscript Received January 10, 1996[⊗]

ABSTRACT: The solution structures of two DNA duplexes containing a DNA•RNA hybrid region at different sites, d(GG)r(AGAU)d(GAC)•d(GTCATCTCC) (DHD, where D and H represent the DNA and DNA•RNA hybrid segments, respectively) and d(GGAGA)r(UGAC)•d(GTCATCTCC) (DDH), were determined by nuclear magnetic resonance spectroscopy to clarify the structural features of the D—H and H—D junctions. All proton—proton distances were derived from the NOESY spectra, with mixing times of 45 ms, and the restrained molecular dynamics were carried out starting from the typical A- and B-form conformations. Both duplexes were converged from the respective initial structures into structures with RMSD values of less than 1.0 Å. These structures were subjected to full relaxation matrix refinement to produce the final structures. In the case of the D—H junction, where the ribonucleotide was linked to the 3′-end of the DNA, the H2′ and H2″ signals of the deoxynucleotide overlapped completely, and the ribonucleotide had a H1′—H2′ coupling constant larger than that of the normal C3′-endo sugar pucker. The dihedral angles, the pseudorotation phase angles, and the helical parameters changed at the H—D junction, but not at the D—H junction. A detailed comparison of the two duplexes revealed the structural heterogeneity between the DNA segment and the DNA•RNA hybrid region and the transitions at the junctions.

The DNA•RNA hybrid duplex is an important intermediate during replication. Ribonuclease H (RNase H), ¹ e.g., from *Escherichia coli*, recognizes the DNA•RNA hybrid duplex as its substrate and hydrolyzes the phosphodiester linkages in the RNA strand (Kanaya & Ikehara, 1995). The reverse transcriptase of human immunodeficiency virus, which synthesizes the proviral DNA using its RNA as a template, contains an RNase H domain that is structurally similar to *E. coli* RNase H. These enzymes recognize only DNA•RNA hybrid duplexes, and not DNA or RNA duplexes. Although the mechanism of the substrate specificity of RNase H has not yet been determined, the tertiary structure of the hybrid

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

* To whom correspondence should be addressed.

[⊥] Protein Engineering Research Institute.

[▽] Institute for Protein Research, Osaka University.

[®] Abstract published in Advance ACS Abstracts, March 15, 1996.

duplex is assumed to determine the specificity (Fedoroff et al., 1993; Salazar et al., 1994). The three-dimensional structures of DNA•RNA hybrid duplexes have been determined, mainly by nuclear magnetic resonance (NMR) spectroscopy (Hall, 1993; Lane et al., 1993; Salazar et al., 1993a; Fedoroff et al., 1993; González et al., 1994, 1995). In these studies, it was found that the RNA strand adopted an A-form geometry with the C3′-endo sugar conformation. The sugar puckers of the deoxyriboses were in the S domain, but not exactly in the normal C2′-endo conformation (Lane et al., 1993). Reid and his co-workers reported that the sugar puckers adopted the O4′-endo (Salazar et al., 1993a; Fedoroff et al., 1993), but the existence of multiple conformers in dynamic exchange was assumed by González et al. (1995).

Another type of oligonucleotide duplex of structural interest is that containing one or more ribonucleotide residues incorporated into the DNA, including the Okazaki fragment in which the deoxynucleotides at the 5'-end of one strand of a DNA duplex are converted into ribonucleotides. X-ray crystallographic studies have revealed that the incorporation of even one ribonucleotide per DNA strand transformed the whole duplex to the A-form geometry, with predominantly C3'-endo sugar puckers (Egli et al., 1993; Ban et al., 1994), and the Okazaki fragment adopted an A-type conformation (Egli et al., 1992, 1993). On the other hand, recent NMR studies revealed that the overall conformation remained as the canonical B-form containing the C3'-endo pucker only

[‡] Coordinates of all final structures have been deposited at the Protein Data Bank, Brook Haven National Laboratory, Upton, NY 11973. ID codes are 1DHH and 1DRN for DHD and DDH, respectively.

[§] Faculty of Pharmaceutical Sciences, Hokkaido University.

^{||} Present address: Biomolecular Engineering Research Institute, 6-2-3 Furuedai, Suita, Osaka 565, Japan.

^{**} Present address: Center for New Materials, Japan Advanced Institute of Science and Technology, Hokuriku, 15 Asahidai, Tatsunokuchi, Ishikawa 923-12, Japan.

¹ Abbreviations: RNase H, ribonuclease H; 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, relaxation matrix refined.

at the ribose-containing residue, when a ribonucleotide was incorporated into DNA (Jaishree et al., 1993), and that the Okazaki fragment was composed of the hybrid form and the B-form, with a bend at the junction (Salazar et al., 1993b, 1994). Even when a DNA duplex is flanked by DNA•RNA hybrids, the transition from A-form RNA to B-form DNA conformations involves only one nucleotide (Zhu et al., 1995). The disagreement between the crystal structures and the solution structures determined by NMR could be attributed to the effects of dehydration and the crystal packing force.

Here we report the structures of two duplexes each containing a DNA·RNA hybrid region at different sites in the DNA. One is a nine-base-pair (9 bp) oligodeoxynucleotide duplex containing a 4 bp hybrid region in its center, which is designated DHD. In the other duplex, a stretch of ribonucleotides is located at the 3'-end of one strand (DDH). The structural features of these duplexes are reported and discussed. The main purpose of this study is to elucidate the influence of ribonucleoside incorporation on the conformation of the DNA duplex, especially at the junction sites. There have been no reports about the conformation of the D-H junction. The mechanism for substrate recognition by RNase H may also be derived from our study, since it has been reported that a 4 bp DNA·RNA hybrid flanked by DNA duplexes is the minimum substrate recognized efficiently by E. coli RNase H (Hogrefe et al., 1990).

MATERIALS AND METHODS

Oligonucleotide synthesis. 5'-d(GG)r(AGAU)d(GAC)-3'. 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 DNA and RNA phosphoramidites used in the synthesis were purchased from Perkin Elmer and MilliGen, respectively. The scale of the synthesis was 10 μ mol. The oligomer was detritylated at the end of the synthesis and was cleaved from the support and 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). The residue was desalted on a Sephadex G-25 column. The residue was dissolved in 1 mL of H₂O and was passed through a Millipore filter (0.45 μ m pore size). Purification was performed by reverse-phase HPLC with a linear gradient of 9-13% CH₃CN in 0.1 M TEAA. The sample was desalted on a Sephadex G-25 column. The oligomer was finally converted to the sodium form on an AG50WX8 cation exchange resin. The final yield of the oligomer was 2.49 μ mol (25%).

5'-d(GGAGA)r(UGAC)-3'. The oligonucleotide synthesis was performed in the same manner as for the former, except that the final detritylation was not performed on the synthesizer. The oligomer 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 TBAF in 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 TEAA. After concentration, this mixture was applied to a column of C18 silica gel, and the product was eluted with a gradient of 5-40% CH₃CN in 0.1 M TEAA. The elution of the oligomer was monitored at 254 nm using 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 NH₄OH and evaporated. The residue was dissolved in 1 mL of H₂O and was passed through a Millipore filter (0.45 μm pore size). Purification was performed by reverse-phase HPLC with a linear gradient of 9-13% CH₃CN in 0.1 M TEAA. The sample was desalted on a Sephadex G-25 column. The oligomer was finally converted to the sodium form on an AG50WX8 cation exchange resin. The final yield of the oligomer was $2.65 \mu \text{mol}$ (27%).

5'-d(GTCATCTCC)-3'. The oligonucleotide synthesis was performed in the same manner as for d(GGAGA)r(UGAC). 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. The residue was dissolved in water and applied to a column of C18 silica gel, and the product was eluted with a gradient of 5-40% CH₃CN in 0.1 M TEAA. The appropriate fractions were concentrated to dryness. Acetic acid (80%, 10 mL) was added to this residue. After 20 min, the solution was evaporated in vacuo and coevaporated with H₂O three times. The residue was dissolved in water and washed with ethyl acetate. The aqueous layer was evaporated to dryness, and the residue was dissolved in 1 mL of H₂O and was passed through a Millipore filter (0.45 μ m pore size). The purification and desalting were performed in the same manner.

NMR Sample Preparation and Spectroscopy. Both strands were dissolved at a concentration of 3 mM in a buffer containing 50 mM sodium phosphate, 100 mM NaCl, and 3 mM ethylenediaminetetraacetic acid (pH 6.95). The solution was heated at 80 °C for 5 min and was gradually cooled down to room temperature. All NMR experiments were carried out on a Bruker AMX-600 spectrometer (600.13 MHz for protons) at 298 K in D₂O or 288 K in H₂O. Three nuclear Overhauser spectroscopy (NOESY) spectra with mixing times of 45, 90, and 150 ms were measured within 4 days, without removing the sample from the spectrometer, for each DHD and DDH duplex in D₂O. 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 (during which the residual HDO peak was irradiated), and a spectrum width of 6024.1 Hz. The exclusive correlated spectroscopy (E-COSY) spectrum, the total correlated spectroscopy (TOC-SY), and the double quantum filtered correlated spectroscopy (DQF-COSY) were acquired at 1024 complex points in t_2 , and 512 points in t_1 , utilizing a relaxation delay of 5 s and a spectral width of 6024.1 Hz. All spectra in H₂O were acquired using a 1-1 spin echo pulse sequence for the solvent suppression (Sklenár et al., 1987). NOESY in H₂O was obtained with 1024 complex points in t_2 and 512 points in t_1 , utilizing a relaxation delay of 1 s and a spectral width of 15151.51 Hz. Mixing time was set to 100 ms. The NMR data were processed using the program Felix 2.3 (Hare Research, Inc.) and UXNMR (Bruker). For processing the

NOESY data sets acquired in either H_2O and D_2O , a sine-bell function shifted by 45° was used as a window function in each dimensions. Zero-filling followed by a Fourier transform and a baseline correction with a fifth order polynomial gave the spectra with 2048 real points in both dimensions without any baseline distortion. Their cross peak intensities were obtained using a measure-volume tool of Felix software. E-COSY, TOCSY-, and DQF-COSY were processed with a sine-bell function in both dimensions. Their spectral point resolution was 1.47 Hz/point for both dimensions.

Determination of Distances and Other Restraints. In both DHD and DDH, the initial buildup rates (R_{ii}) in D_2O were calculated by the cross-peak volumes of NOESY at the mixing time of 45 ms. These rates were converted to distances (r_{ii}) using the cytosine H5-H6 distance of 2.5 Å (for all protons except those of the methyl group of thymine) and the thymine H6-CH₃ distance of 2.9 Å (for the methyl protons of thymine) as the reference distances (r_{ref}) with the relationship $r_{ij} = r_{ref}(R_{ref}/R_{ij})^{1/6}$. These distances were given with upper and lower bounds of 0.3 and 0.4 Å, respectively (Salazar et al., 1994; Schweitzer et al., 1994). In DHD, the initial buildup rates in H₂O were also calculated from the cross-peak volumes at 100 ms. For the sets of volumes in H_2O , the offset dependence of the 1–1 echo excitation (sin³ $\omega \tau$) was taken into account, and the volumes were calibrated. The rates (R_{ii}) were converted to distances (r_{ii}) as described above. These distances were given with upper bounds of r_{ii} + 0.3 Å. Lower distance bounds were not obtained because of their exchange. A list of NOE-derived distance restraints in both D₂O and H₂O for the DHD and DDH duplexes is provided as Supporting Information. Additionally, following restraints were used in this study. (i) Backbone torsion angles. In order to preserve the right-handed characteristics of these duplexes, the α , β , ϵ , γ , and ζ backbone torsion angles were restrained to a range covering both the righthanded A- and B-DNA forms. (ii) Hydrogen-bond distances in base pairs. To maintain the conservative Watson-Crick base pairing, distance restraints between the bases were used. The values were adopted from the X-ray analysis. (iii) 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. (i) DHD. The initial A- and B-form structures were generated using the INSIGHT molecular modeling program (BIOSYM Technologies) and were designated Ini-DHD-A and Ini-DHD-B. All subsequent energy minimization were performed on a Silicon Graphics INDIGO R4000XS24 workstation using the X-PLOR 3.1 program (Brünger, 1993). All initial and refined structures were displayed by using the SYBYL molecular modeling system. Nucleic acid parameters and force constants used in the calculation were taken from the parallhdg.dna (X-PLOR 3.1). The phosphate charge was reduced to -0.32e(Tidor et al., 1983). The switching function between 10.5 and 11.5 Å was used to make all pairwise nonbonded interactions (electrostatic and van der Waals) progress smoothly to zero. A 600-step conjugate gradient minimization was performed on the system before initiating the dynamics. The force constant of 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 kept throughout all energy minimization steps.

Table 1: J Coupling Constants (Hz) of DHD residue $J_{1'-2'}$ $J_{1'-2''}$ $\sum J_{1'}$ $\sum J_{3'}$ residue $J_{1'-2'}$ $J_{1'-2''}$ $\sum J_{1'}$ $\sum J_{3'}$ 1G 8.5 6.8 15.2 а 18C a 2G 17C 16.1 13.3 а a а a а а 3A 5.5 nd 16T 15.9 13.2 4G 15C 7.5 6.3 13.8 15.9 b nd 5A b nd 14T 8.6 6.4 14.9 10.6 6U b nd 13A 7.7 6.3 14.0 a 14.2 a 7G 7.6 6.7 12C 8.0 6.6 14.4 13.3

^a Could not be determined because of spectral overlap. ^b The coupling constant was too small to be observed. The values of $\sum J_{1'}$ (= $J_{1'2'}$ + $J_{1'2''}$) and $\sum J_{3'}$ (= $J_{2'3'}$ + $J_{2''3'}$ + $J_{3'4'}$) were derived from the apparent splitting of the outer peaks of cross-peak. nd, not determined.

11T

10G

9.1

5.7

14.9 13.3

8A

9C

8.0

6.6

6.3

14.6 15.9

15.4 15.6

Table 2: J Coupling Constants (Hz) of DDH											
residue	$J_{1'-2'}$	$J_{1'-2''}$	$\sum J_{1'}$	$\sum J_{3'}$	residue	$J_{1'-2'}$	$J_{1'-2''}$	$\sum J_{1'}$	$\sum J_{3'}$		
1G	7.4	6.6	13.6	а	18C	а	а	а	a		
2G	9.2	6.6	14.6	9.2	17C	6.6	6.7	13.3	14.6		
3A	7.6	8.3	15.6	9.2	16T	7.9	6.6	14.5	8.2		
4G	9.3	6.7	13.9	9.3	15C	7.9	6.6	13.3	a		
5A	a	a	a	a	14T	8.6	6.0	14.6	9.4		
6U	6.5				13A	9.3	8.0	15.9	9.3		
7G	b				12C	8.0	6.6	13.2	9.3		
8A	b				11T	8.0	8.0	14.6	12.0		
9C	5.7				10G	a	a	a	a		

^a Could not be determined because of spectral overlap. ^b The coupling constant was too small to be observed. The values of $\sum J_{1'}$ (= $J_{1'2'}$ + $J_{1'2''}$) and $\sum J_{3'}$ (= $J_{2'3'}$ + $J_{2''3'}$ + $J_{3'4'}$) were derived from the apparent splitting of the outer peaks of cross peak. nd, not determined.

The system was initially heated to 1200 K and was gradually cooled down to 200 K over 20 ps and was maintained at 200 K during the subsequent 10 ps. The coordinates at the last 8 ps were averaged and subjected to 600 cycles of restrained energy minimization to obtain the final structures, rMD-DHD-A and rMD-DHD-B. (ii) DDH. The initial structures were generated using the SYBYL molecular modeling program (TRIPOS Inc.) and were designated Ini-DDH-A and Ini-DDH-B. All subsequent energy minimization was performed in the same manner as the former one, except that the minimization was started at 900 K. The coordinates at the last 8 ps were averaged and subjected to 600 cycles of restrained energy minimization to obtain the final structures, rMD-DDH-A and rMD-DDH-B.

Full Relaxation Matrix Refinement. (i) DHD. At the final stage of the structure determination, direct NOE refinements (Yip & Case, 1989) of the rMD-DHD-A and rMD-DHD-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. The experimental volumes were taken from the 150 ms NOESY data set. For the rMD-DHD-A and rMD-DHD-B calculations, 256 volumes were used. All volumes were given with a relative error of 10%. The theoretical volumes were calculated from a relaxation matrix of the coordinate set. For these calculations, a uniform isotropic correlation time of 4.25 ns was used. With the restraints of K_{NOE} (exp. in D_2O) = 0, K_{NOE} (hydrogen bond, exp. in H_2O) = 25, K_{CIDH} = 0, and K_{relax} = 1000, rMD-DHD-A and rMD-DHD-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 more steps of dynamics. This was repeated until a final temperature of 75 K was reached. During these procedures, the K_{NOE} and the K_{CIDH} were kept constant. The final step consisted of 90 steps of energy minimization, and the full nonbonded interaction terms were included in the calculations. The refined structures were designated as RR-DHD-A and RR-DHD-B. The R factors were calculated as the weighted average of the absolute value of the difference between 1/6th power of the observed and calculated intensities. (James, 1991; Thomas et al., 1991; Brünger, 1993). (ii) DDH. The procedure of the relaxation matrix refinement was performed in the same manner as the DHD. For the rMD-DDH-A and rMD-DDH-B calculations, 260 volumes were used. With the restraints of K_{NOE} (exp. in D_2O) = 0, K_{NOE} (hydrogen bond) = 25, $K_{\text{CIDH}} = 0$, and $K_{\text{relax}} = 1000$, rMD-DDH-A and rMD-DDH-B were refined by the dynamical simulated annealing protocol to yield the final structures, RR-DDH-A and RR-DDH-B.

The analyses of the helical parameters were performed 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 two duplexes 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.

Proton Resonance Assignments. The exchangeable proton resonances were assigned using the NOESY spectrum. All imino proton resonances were observed and assigned, including the terminal ones. In DHD, the amino proton signals of the cytosines and adenines showed connectivities to the imino ones of guanines and thymines, respectively. In DDH, any amino resonances were not assigned because of their overlappings. Adenine H2 signals were identified by NOESY cross peaks from base-paired thymine H3 imino protons. Other nonexchangeable proton resonances were assigned using the DQF-COSY, TOCSY, and NOESY spectra. In both duplexes, the sequential connectivities of the base protons with H1', H2', and H2" were traced. The proton chemical shifts of these two duplexes are shown in the Supporting Information.

The chemical shifts of the H2' and H2'' protons of the deoxynucleoside 5'-adjacent to the ribo residue at the D-H junction (2G of DHD and 5A of DDH) were completely overlapped, although those of the deoxynucleoside at the H-D junction (7G of DHD) were separated. Overlapping

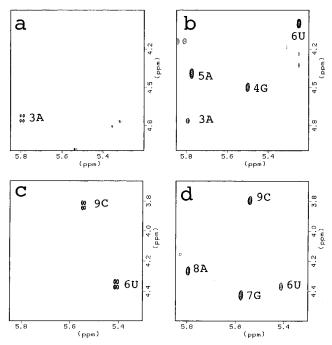


FIGURE 1: DQF-COSY (a and c) and NOESY (b and d) spectra in the H1'-H2'region for 3A, 4G, 5A, and 6U of DHD (a and b) and DDH (c and d).

Table 3: Atomic RMSD (Å) and R Factors of the rMD-DHD-A and -B Structures^a

		rN	1D-DI	łD	Ini-	Ini-	R	
	A1	A2	A3	B1	B2		DHD-B	factor
rMD-DHD-A1						2.19	3.82	0.086
rMD-DHD-A2	0.81					2.30	4.37	0.087
rMD-DHD-A3	0.68	0.66				2.43	4.05	0.087
rMD-DHD-B1	0.97	0.94	0.78			2.26	3.73	0.086
rMD-DHD-B2	0.91	0.96	0.60	0.62		2.46	3.87	0.086
rMD-DHD-B3	0.64	0.94	0.79	0.05	0.63	2.27	3.72	0.086

^a rMD-DHD-A1-3 and rMD-DHD-B1-3 were obtained by restrained molecular dynamics from Ini-DHD-A and Ini-DHD-B, respectively.

Table 4: Atomic RMSD (Å) and R Factors of the rMD-DDH-A and -B Structures^a

		rN	1D-DE	Н	Ini-	Ini-	 R	
	A1	A2	A3	B1	B2	DDH-A	DDH-B	factor
rMD-DDH-A1						3.14	2.40	0.082
rMD-DDH-A2	0.16					3.24	2.35	0.080
rMD-DDH-A3	0.83	0.73				3.51	2.02	0.080
rMD-DDH-B1	0.13	0.12	0.74			3.29	2.37	0.082
rMD-DDH-B2	0.22	0.09	0.69	0.16		3.30	2.43	0.082
rMD-DDH-B3	0.30	0.37	0.68	0.37	0.32	3.23	2.39	0.082

^a rMD-DDH-A1-3 and rMD-DDH-B1-3 were obtained by restrained molecular dynamics from Ini-DDH-A and Ini-DDH-B, respectively.

of the H2' and H2" signals was also observed at the 3' termini of the deoxy strands (10G and 18C of both duplexes).

Determination of the Scalar Coupling Constants in E-COSY. The J-coupling constants were derived from the E-COSY spectra using the fitting tool of FELIX with a 1 Hz precision. Those for H1'-H2' and H1'-H2" are shown in Tables 1 and 2, with the $\sum J_{1'} (= J_{1'2'} + J_{1'2''})$ and $\sum J_{3'} (= J_{2'3'} + J_{2''3'} + J_{3'4'})$ values which were determined experimentally. The H1'-H2' coupling constants $(J_{1'2'})$ of the riboses at the D-H junction site (3A of DHD and 6U of DDH) and at the 3' terminus (9C of DDH) were relatively large (>5 Hz), although those of the other ribo residues were

DHD DDH

FIGURE 2: Stereoview of the best fit superposition of the RR-DHD-A1, -A2, -A3, -B1, -B2, and -B3 (upper) and RR-DDH-A1, -A2, -A3, -B1, -B2, and -B3 (lower).

not detectable (Figure 1). In all deoxyriboses, the H1′—H2″ coupling constants ($J_{1'2''}$) were below 7 Hz, but the $J_{1'2'}$ values were greater than 7 Hz. The $J_{1'2'}$ and $J_{1'2''}$ values of the deoxynucleosides at the D—H junctions (2G of DHD and 5A of DDH) and at the termini of the deoxy strands (10G and 18C) were not identified because of signal overlapping. The $J_{2'3'}$ and $J_{2''3'}$ values were also not determined due to signal overlapping and broadening, although ³¹P decoupling was employed during the acquisition period.

Structure Calculation. The right-handed nature of both duplexes was evident in the NOESY spectra, and the NOE connectivities of the imino protons showed the normal base stacking throughout the duplexes. To evaluate the structural

features of these two duplexes, a structure calculation was carried out. In the case of DHD, three calculations were started from the canonical A-form (Ini-DHD-A) and B-form (Ini-DHD-B) initial conformations, with different random seeds for the assignments of the initial velocities of atoms. Each of the three structures was converted into a lower energy conformation. The root mean square deviation (RMSD) values between the final conformations, rMD-DHD-A from Ini-DHD-A and rMD-DHD-B from Ini-DHD-B, were less than 1.0 Å for all heavy atoms. As shown in Table 3, the DHD structures were closer to Ini-DHD-A than to Ini-DHD-B. For DDH, the structure calculation was carried out in the same manner as for DHD. Using three calculations from the canonical A-form (Ini-DDH-A) and

Table 5: Atomic RMSD (Å) and R Factors of the RR-DHD-A and -B Structures

			RR-DHD						
	A1	A2	A3	B1	B2	Ini-DHD-A	Ini-DHD-B	$rMD-DHD^a$	R factor
RR-DHD-A1						2.19	3.92	0.55	0.069
RR-DHD-A2	0.85					2.36	4.50	0.57	0.068
RR-DHD-A3	0.66	0.65				2.45	4.15	0.55	0.069
RR-DHD-B1	0.67	0.99	0.76			2.31	3.83	0.59	0.068
RR-DHD-B2	0.93	0.99	0.62	0.59		2.53	3.98	0.56	0.069
RR-DHD-B3	0.66	0.99	0.77	0.08	0.60	2.31	3.82	0.58	0.067

"Compared with the rMD structure from which the respective RR structure was derived, e.g., RR-DHD-A1 with rMD-DHD-A1 and RR-DHD-A2 with rMD-DHD-A2.

Table 6: Atomic RMSD (Å) and R Factors of the RR-DDH-A and -B Structures

			RR-DDH						
	A1	A2	A3	B1	B2	Ini-DDH-A	Ini-DDH-B	$\mathrm{rMD}\text{-}\mathrm{DDH}^a$	R factor
RR-DDH-A1						3.21	2.40	0.64	0.062
RR-DDH-A2	0.17					3.32	2.39	0.64	0.061
RR-DDH-A3	0.18	0.09				3.59	2.13	0.65	0.063
RR-DDH-B1	0.23	0.21	0.67			3.21	2.37	0.62	0.062
RR-DDH-B2	0.31	0.21	0.62	0.16		3.34	2.42	0.62	0.061
RR-DDH-B3	0.35	0.32	0.60	0.25	0.32	3.17	2.41	0.60	0.061

^a Compared with the rMD structure from which the respective RR structure was derived, e.g., RR-DDH-A1 with rMD-DDH-A1 and RR-DDH-A2 with rMD-DDH-A2.

B-form (Ini-DDH-B) conformations, final conformations (rMD-DDH-A and rMD-DDH-B, respectively) with RMSD values less than 1.0 Å for all heavy atoms were obtained. As shown in Table 4, the DDH structures were closer to Ini-DDH-B than to Ini-DDH-A.

Full Relaxation Matrix Refinement. In order to refine the obtained structures, taking the spin diffusion into account, a full relaxation matrix refinement with structure calculation was carried out without using any dihedral angle restraints. For this refinement, 256 NOE intensities were used. The refined structures were designated as RR-DHD-A and RR-DHD-B, respectively. As shown in Figure 2 and Table 5, the structures of the RR-DHDs were closer to Ini-DHD-A than to Ini-DHD-B. The RMSD values for all heavy atoms between the structures before and after the full relaxation matrix refinement, and the R factors of the refined structures. are also shown in Table 5. The refinement of DDH was performed in the same way, and the refined structures were designated as RR-DDH-A and RR-DDH-B. The six structures (rMD-DDHs) were converted to refined structures (RR-DDHs, Figure 2) with lower R factors, as shown in Table 6.

Dihedral Torsion Angles. In both DHD and DDH, a structural discontinuity was observed, as shown in Figure 3. In DHD, most of the δ values were smaller than 120°, which is usually observed for the A-form duplex, but the δ values of the nucleotides at the H-D junction site (7G and 13A) were changed significantly, to the range of the B-form. On the other hand, no change was observed at the D-H junction site. In DDH, the δ values of the ribo residues were smaller than those of the deoxynucleotides, and a drastic change was not observed at the junction sites. The ϵ and χ values in both DHD and DDH suggested that an A-like structure was formed in the DNA·RNA hybrid region, although a significantly large χ value (larger than 270°) was obtained at the H–D junction site (13A) of DHD. The δ and γ angles in the DNA duplex regions were intermediate between those observed for the A- and B-forms.

Helical Parameters. Structural discontinuities were also found in the helical parameters (Figure 4). The values of

the rise, the roll, and the slide were increased at the H–D junction site (between the 6U·13A and 7G·12C base pairs of DHD), whereas no obvious changes were observed at the D–H junction sites (between 2G·17C and 3A·16T of DHD and between 5A·14T and 6U·13A of DDH). In the case of the tilt, the values in the DNA·RNA hybrid region deviated from those observed in the normal A- and B-form duplexes, but there was no distinctive feature at the junction sites.

Pseudorotation Phase Angles (Sugar Conformations). The J-coupling constants, within a 1 Hz precision, are shown in Table 1 and 2. The dihedral angle constraints for the sugar conformations were not used throughout the rMD and relaxation matrix refinement calculations of the structures of DHD and DDH, since the dihedral angles of H1'-C1'-C2'-H2' and H1'-C1'-C2'-H2" could not be determined simultaneously using the $J_{1'2'}$ and $J_{1'2''}$ values, respectively. After the structure calculations using only the distance constraints derived from the NOE data, the pseudorotation phase angles for deoxyriboses fluctuated between 180° (typical C2'-endo) and 60° (Figure 3), and coupling constants of some residues could not be explained by the obtained sugar conformations. Our interpretation of the inconsistency observed for the deoxyriboses is described under Discussion. In the case of the ribonucleotide residues, the $J_{1'2'}$ value for 3A of DHD was larger than 5 Hz, suggesting that this ribose did not adopt the pure C3'-endo conformation, but its pucker obtained by the molecular dynamics was C3'-endo, which is typical for riboses. In contrast, the sugar pucker obtained for 4G was O4'-endo, although J coupling was not detected. This kind of inconsistency may be attributed to the relatively high R factors in this region, which are described in the next section and discussion, and it is appropriate for the discussion of the structural features of the D-H junction to use the structure of the DDH duplex mainly.

NOE Constraints and Local R Factors. The number of input NOE restraints was plotted against the residue number in Figure 5 panels a and b for DHD and DDH, respectively. The number of intraresidual constraints from 1G to 5G in

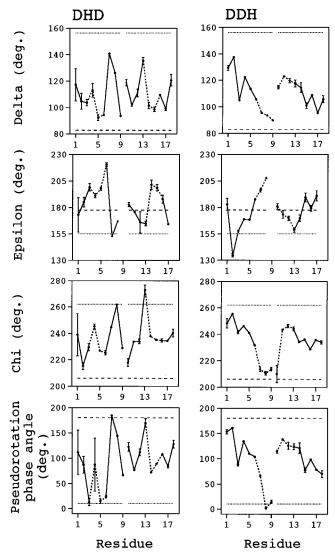


FIGURE 3: Comparison of the dihedral and pseudorotation phase angles of DHD and DDH. The averages of the six RR-DHDs and the six RR-DDHs are presented with deviations. The deoxy (-·-) and hybrid (-·•-) duplex regions are shown. The values for the standard A-form (-·-) and B-form (···) DNA (Arnott & Hukins, 1972, 1973) are also shown.

DHD was less than that in the other parts of DHD and DDH, because of the spectral overlapping. The *R* factors calculated for each residue were shown in Figure 6 panels a and b for DHD and DDH, respectively, to show how these conformations satisfied the experimental NOE restraints locally. In the DDH duplex, *R* factors for 2G, 15C, and 16T were larger than the others. On the other hand, there was no unusual value in DDH.

Hybrid Structure and Minor Groove Width. For the comparison of the DNA•RNA hybrid structure of each DHD and DDH, the minor groove width was defined as the distance from each phosphorus to the second phosphorus on the 3'-side of the other strand (Fedoroff et al., 1993; Salazar et al., 1994). Those of DHD and DDH were shown in Figure 7, as a function of the residue numbers of the chimera strand. In DHD, the DNA•RNA hybrid region had a minor groove wider than that in the DNA segment. The width was larger at 6U (from the phosphorus between 5A and 6U to that between 15C and 16T across the minor groove within the hybrid region) and became smaller as it covered the DNA segment. In the case of DDH, however, no distinctive feature

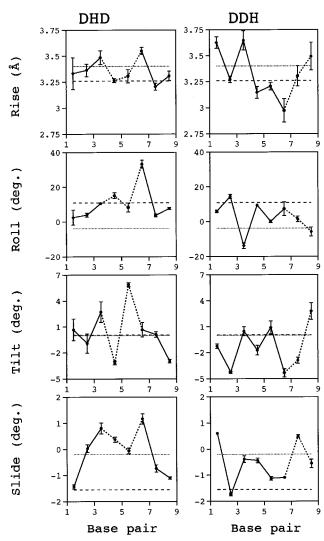


FIGURE 4: Comparison of helical parameters of DHD and DDH. The averages of the six RR-DHDs and the six RR-DDHs in the deoxy (- • -) and hybrid duplex regions (- • • -) are presented and compared to those of the standard A-form (- - -) and B-form (· · ·) DNA.

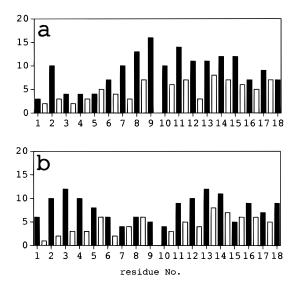


FIGURE 5: Numbers of intra (dark) and inter (blank) residual NOE constraints of DHD (a) and DDH (b) from NOESY spectra in D₂O.

was found. Between DHD and DDH, there was no common feature in the DNA•RNA hybrid structure, in terns of the minor groove widths.

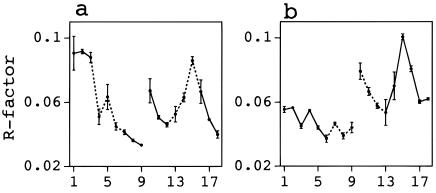


FIGURE 6: R factors calculated for each residue are shown for DHD (a) and DDH (b). Deoxy ($- \cdot -$) and hybrid duplex regions ($- \cdot - \cdot -$) are presented.

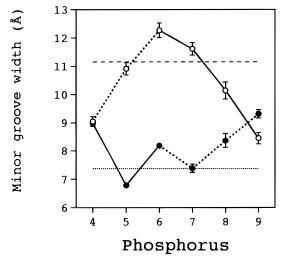


FIGURE 7: Interstrand phosphorus—phosphorus separation (minor groove width) of DHD (○) and DDH (●) compared to the standard A-form (---) and B-form (·--) DNA. The averages in the deoxy (—) and hybrid duplex regions (---) are presented with deviations.

DISCUSSION

The solution structures of two oligonucleotide duplexes containing a DNA·RNA hybrid region at different sites, designated DHD and DDH, have been determined by twodimensional NMR, structure calculation, and full relaxation matrix refinement. The spectral and structural characters of these duplexes are as follows: (1) At the D-H junction site, some unique spectroscopic features were observed. Those are the completely overlapped H2' and H2" signals of the deoxyribose in both duplexes and the $J_{1'2'}$ coupling constant of the ribose at the D-H junction larger than that expected for the C3'-endo conformation. These characters were also found at the 3' termini. (2) Changes in the dihedral angles and the helical parameters were observed at the H-D junction, but not at the D-H junction. (3) In DHD, the minor groove in the DNA·RNA hybrid region was more varied than that of DDH. The global structures of DHD and DDH were closer to the A- and B-forms, respectively. The structure of the hybrid segment of DHD did not resemble to that of DDH.

Unusual spectral features, the overlapped H2' and H2'' signals of the deoxyribose and the large $J_{1'2'}$ coupling constant for the ribose, were observed at the D–H junction sites in our duplexes (2G and 3A of DHD and 5A and 6U of DDH). They also have been reported by Chou et al. (1991) for the solution structures of chimeric DNA–RNA–DNA duplexes.

In their NMR studies of three duplexes, {d(CGCG)r(AAUU)d(CGCG)}₂, {d(CGCG)r(UAUA)d(CGCG)}₂, and d(CGTT)r(AUAA)d(TGCG)·d(CGCA)r(UUAU)d(AACG), the H2' signal of the deoxyribonucleotide 5'-adjacent to the RNA segment appeared slightly downfield of the H2", and the ribonucleotide linked to the 3'-end of DNA exhibited detectable H1'-H2' *J* coupling although A- and B-conformations were proposed for the RNA and DNA segments, respectively. Since these spectral features were not observed at the H-D junction in the DHD or in the Okzaki fragment containing a H-D junction (Salazar et al., 1993b, 1994), they may be unique to the deoxynucleotide—ribonucleotide junction in the 5'-3' direction, regardless of whether the ribo residue is a part of an RNA duplex or a DNA·RNA hybrid.

By means of the structure calculation, followed by the full relaxation matrix refinement, the solution structures of DHD and DDH were determined. The relaxation matrix, which was refined without using any direct dihedral angle restraints on the backbone, enabled us to determine the helical parameters and the backbone torsion angles more accurately (Kaluarachchi et al., 1991). For DHD and DDH, a set of cross peak volumes, derived from the NOESY experiment with a mixing time of 150 ms, was used directly in the calculations. This procedure has been suggested to be superior to that utilizing the distances derived from the NOESY experiment with a short mixing time, since the spin diffusion is explicitly taken into account (Post et al., 1990). Although the dihedral angle restraints were not used in the full relaxation matrix refinement, final structures with small RMSD values were obtained by using more NOE data. From the peak volumes in the NOESY spectra, local R factors were calculated (Figure 6). Since the H-D junction of DHD and the D-H junction of DDH had a low R factor, structural characteristics of these junctions can be discussed using the coordinates obtained in this study. Larger values at the D-H junction of DHD may be attributed to flexibility of the sugar conformations at this site, although the equilibrium between the sugar puckers could not be determined because of signal overlapping. The R factors of the DNA strand seem to be dependent on the sequence.

As shown in Figure 2 and Tables 5 and 6, the global structures of DHD and DDH were closer to the A- and B-forms, respectively. Three of the seven torsion angles, which exhibit the distinctive features of these duplexes, are shown in Figure 3. The pseudorotation phase angles are also shown. These values suggest that the RNA regions adopt an A-type conformation and that the geometries of the deoxynucleosides are intermediate between the A- and

B-form. Drastic changes in the δ , ϵ , and χ values and the pseudorotation phase angle were found at the H-D junction of DHD (between 6U and 7G in the chimeric strand and at 13A in the DNA strand). Such changes in H-D junction were also observed in the helical parameters, tilt, roll, rise, and slide as showed in Figure 4. On the other hand, the smooth transition from the hybrid-form to the B-form was observed in D-H junction. This structural discontinuity at the H-D junction of DHD resembled that at the H-D junction of the Okazaki fragment and of the HDH-type duplex in solution (Salazar et al., 1994; Zhu et al., 1995).

In this study, the dihedral angle constraints for the sugar conformations were not used in the structure calculations because contradiction was found in using the J coupling constants for the determination of the sugar puckers. Besides the discrepant dihedral angles described in Results, another contradiction could be pointed out. Using the $J_{1'2'}$ coupling constant, the pseudorotation phase angle estimated for 12C of DDH (this residue was chosen as an example because of its evident difference) was around 100° (O4'-endo conformation), but that estimated from the $\sum J_{3'}$ value was around 150° (C2'-endo). We assumed that this contradiction was caused by an equilibrium of multiple sugar conformations, which has been shown in the DNA strand of a DNA·RNA hybrid duplex (González et al., 1994). Assuming multiple conformations, the fractional percentage of S conformer (%S) calculated for this residue using its $\sum J_{1'}$ value was 58% (Rinkel & Altona, 1987). This %S value corresponded qualitatively to the sugar conformation with a pseudorotation phase angle of 120° that was obtained for this residue in the final structure of DDH. Moreover, the distance between H1' and H4' obtained for this residue from the NOE data was longer than 3 Å, which was much longer than that estimated from the $J_{1'2'}$ coupling constant in the single conformer model. As described above, our calculation method in which the dihedral angles were not used as constraints was justifiable, and the observed J coupling constants could be related, so some extent, to the obtains structures by using the N/S equilibrium model.

The ³¹P chemical shift value is affected by its bond angle and stereoelectronic effects (Gorenstein, 1994; Schweitzer et al., 1995). It has been reported that the chemical shift dispersion in the DNA·RNA hybrid duplex was relatively wide (Lane et al., 1993). Figure 8 panels a and b show the ³¹P NMR spectra of DHD and DDH, respectively. Both spectra exhibited relatively wide chemical shift variety. This observation suggests a relatively wide distribution of backbone torsion angles in the hybrid region of DHD and DDH. Furthermore, separated and broad signals in the low field (between -3.0 and -3.2 ppm) were observed in both spectra. Since these low field shifts were not reported in the DNA·RNA hybrid duplexes or in the Okazaki fragment that contains the H-D junction (Salazar et al., 1994; Zhu et al., 1995), they may be caused by the D-H junction in DHD and DDH.

The width of the minor groove of the DHD and DDH duplexes was determined, as shown in Figure 7, since it has been suggested that its width intermediate between the A-and B-form structures, which was found in a DNA•RNA hybrid duplex, is very important for the substrate recognition by RNase H (Fedroff et al., 1993). The chimeric strand of DHD was cleaved at a single site between 5A and 6U, and DDH was hydrolyzed at several sites in the hybrid region

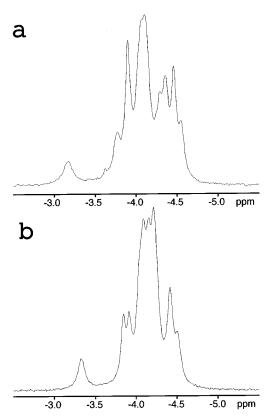


FIGURE 8: ³¹P NMR spectrum of the DHD (a) and DDH (b). The spectrum was recorded (acquisition time = 4 s) at AMX600 with WALTZ decoupling of the ¹H. The temperature was 298 K. The spectral reference was external trimethyl phosphate as 0 ppm. The free induction decay was apodized with the exponential window function (LB = 1 Hz). (Lane et al., 1993)

(M. Wakasa, unpublished result). From these results, there is apparently no relationship between the minor groove width and the RNase H cleavage sites, even through the 2'-OH requirement is taken into account (Iwai et al., 1995). Recently, we determined the solution structure of another DNA•RNA hybrid duplex containing 2'-O-methylribonucleotides, which turned out to have an A-form geometry (manuscript in preparation). Its comparison with the DHD and DDH structures will give a clue to the elucidation of the substrate recognition mechanism of RNase H.

In conclusion, the remarkable finding from the studies of the structures of DHD and DDH is that some drastic changes were observed for the nucleotides at the H–D junction, whereas the D–H junction, which was analyzed for the first time, exhibited no obvious structural change.

ACKNOWLEDGMENT

We thank Mr. Makoto Wakasa with the digestion assay of the RNase H, Dr. Yasushi Oda for help with the NMR measurement at the beginning of this study, and Dr. Soichi Morikawa for helpful discussions.

SUPPORTING INFORMATION AVAILABLE

Proton chemical shift table and a list of NOE-derived distance restraints in both D₂O and H₂O of each DHD and DDH duplex (28 pages). Ordering information is given on any current masthead page.

REFERENCES

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.
 Ban, C., Ramakrishnan, B., & Sundaralingam, M. (1994) *J. Mol. Biol.* 236, 275–285.
- Brünger, A. T. (1993) X-PLOR 3.1: A System for Crystallography and NMR, Yale University, New Haven, CT.
- Chou, S., Flymn, P., Wang, A., & Reid, B. (1991) Biochemistry 30, 5248-5237.
- Egli, M., Usman, M., Zhang, S., & Rich, A. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89, 534–538.
- Egli, M., Usman, M., Zhang, S., & Rich, A. (1993) *Biochemistry* 32, 3221–3237.
- 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.
- Gorenstein, D. G. (1994) Chem. Rev. 94, 1315-1338.
- Gorenstein, D. G. (1992) Methods Enzymol. 211, 254-286.
- Griesinger, C., Sorensen, O. W., & Ernst, R. R. (1986) J. Chem. Phys. 85, 6837–6852.
- Griesinger, C., Sorensen, O. W., & Ernst, R. R. (1987) J. Magn. Reson. 75, 474–492.
- Haasnoot, C. A. G., Westerink, H. P., van der Marel, G. A., & van Boom, J. H. (1983) *J. Biomol. Struct. Dyn. 1*, 131–149.
- Hall, K. B. (1993) Curr. Opin. Struct. Biol. 3, 336-339.
- Hogrefe, H. H., Hogrefe, R. I., Walder, R. Y., & Walder, J. A. (1990) *J. Biol. Chem.* 265, 5561–5566.
- Inoue, H., Hayase, Y., Iwai, S., & Ohtsuka, E. (1987) *FEBS Lett.* 215, 327–330.
- Iwai, S., Kataoka, S., Wakasa, M., Ohtsuka, E., & Nakamura, H. (1995) FEBS Lett. 368, 315-320.
- Jaishree, T. N., van der Marel, G. A., van Boom, J. H., & Wang, H.-J. (1993) *Biochemistry* 32, 4903–4911.
- James, T. L. (1991) Curr. Opin. Struct. Biol. 1, 1042-1053.
- Kaluarachchi, K., Meadows, R. P., & Gorenstein, D. G. (1991) Biochemistry 30, 8785–8797.
- Kanaya, S., & İkehara, 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.

- Lane, A. N., Ebel, S., & Brown, T. (1993) Eur. J. Biochem. 215, 297–306.
- 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.
- Piantini, U., Sorensen, O. W., & Ernst, R. R. (1982) J. Am. Chem. Soc. 104, 6800-6801.
- Post, C. B., Meadows, R. P., & Gorenstein, D. G. (1990) *J. Biomol. Struct. Dyn.* 8, 253–294.
- Rinkel, L. J., & Altona, C. (1987) J. Biomol. Struct. Dyn. 4, 621-
- Salazar, M., Fedoroff, O. Y., Miller, J. M., Ribeiro, N. S., & Reid, B. R. (1993a) *Biochemistry 32*, 4207–4215.
- Salazar, M., Champoux, J. J., & Reid, B. R. (1993b) *Biochemistry* 32, 739-744.
- 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.
- Schweitzer, B. I., Gardner, K. H., & Kellogg, G. T. (1995) *J. Biomol. NMR* 6, 180–188.
- 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.
- Yang, W., Hendrickson, W. A., Crouch, R. J., & Satow, Y. (1990) Science 249, 1398–1405.
- Yip, P., & Case, D. A. (1989) J. Magn. Reson. 83, 643-648.
- Zhu, L., Salazar, M., & Reid, B. R. (1995) *Biochemistry 34*, 2372–2380.

BI9519821