The DNA Strand in DNA·RNA Hybrid Duplexes is Neither B-form nor A-form in Solution†

Miguel Salazar,‡ Oleg Y. Fedoroff,§ Julie M. Miller,‡ N. Susan Ribeiro,‡ and Brian R. Reid*,‡,∥

Department of Chemistry, BG-10, and Department of Biochemistry, SJ-70, University of Washington, Seattle, Washington 98195, and Institute of Mathematical Problems of Biology, Russian Academy of Sciences, Pushchino, Moscow Region 142292, Russia

Received November 30, 1992

ABSTRACT: The structure of the DNA·RNA hybrid (GTCACATG)·(caugugac), where lowercase letters designate RNA residues, has been determined on the basis of *J*-coupling analysis and 2D-NOE studies. The central hexamer in this sequence has been previously studied [Reid, D. G., Salisbury, S. A., Brown, T., Williams, D. H., Vasseur, J.-J., Rayner, B., & Imabach, J.-L. (1983) *Eur. J. Biochem. 135,* 307–314] via one-dimensional NOE methods and circular dichroism studies. Contrary to their results, we find that this duplex *does not* assume a B-form conformation in solution. Instead, the RNA residues retain their C3′-endo (A-form) conformation, as indicated by the absence of H1′-H2′ couplings and by strong H6/H8 to (n – 1) H2′ NOEs. The sugars of the DNA residues, on the other hand, do not assume an A-form (or a B-form) conformation but an *intermediate* conformation in the O4′-endo range ($P \approx 72$–$110°$), as indicated by the presence of strong H1′-H4′ NOEs, medium-strength H2″–H3′ COSY cross peaks, strong H3′–H4′ DQF-COSY cross peaks, and H1′–H2′ coupling constants that are of approximately the same magnitude as the H1′–H2″ coupling constants. These results suggest that the RNA strand not only retains its N-type structure but also exerts an influence on the conformation of the DNA strand. Our results provide strong evidence that DNA·RNA hybrid duplexes *do not* assume an all-C2′-endo B-type conformation; neither do they assume an all-C3′-endo A-type conformation in solution. Furthermore, although not the main focus of this study, a comparison of the longitudinal relaxation times of the DNA and RNA residues indicates the need for extended relaxation delays in two-dimensional NMR spectra of hybrid duplexes, as has been previously observed for DNA·RNA chimeric duplexes (Wang, A. C., Kim, S.-G., Chou, S.-H., Orban, J., Flynn, P., & Reid, B. R. (1992) *Biochemistry 31,* 3940–3946).

Knowledge of the solution structure of DNA·RNA hybrids is of primary importance because of their role in genetic information transfer during transcription of DNA into messenger RNA, reverse transcription of RNA into DNA during retroviral replication, and DNA replication via Okazaki fragments, in which short RNA primer hybrid duplexes of the lagging strand are extended into DNA duplexes.

The occurrence of DNA·RNA hybrid helices was first demonstrated by Rich (1960), who studied the hybrid poly(rA)·poly(dT) using optical density measurements. Milman and co-workers (1967) carried out the first X-ray diffraction study on a hybrid f1 DNA·RNA duplex and concluded that the hybrid *did not* adopt a B-form conformation. A calf thymus DNA·RNA hybrid has been studied via optical rotatory dispersion (Tunis & Hearst, 1968). This hybrid was found to exhibit ORD spectra more similar to double-stranded RNA than to B-form DNA. The X-ray diffraction pattern of poly(rI)·poly(dC) was found to be more similar to that of A-form RNA than to B-form DNA (O'Brien & MacEwan, 1970). The first circular dichroism study of hybrids was carried out by Gray and Ratliff (1975) on poly[d(AC)·r(GU)] and poly[r(AC)·d(GT)]; they concluded that the spectrum of poly[r(AC)·d(GT)] more closely resembled the spectrum of the double-stranded RNA duplex poly[r(AC)·r(GU)] than did that of the reverse poly[d(AC)·r(GU)]. The base protons, anomeric H1′ protons, and the imino protons of the hybrid duplex (CTTTTTG)·(caaaaag), where lowercase letters represent RNA residues, have been studied by Pardi et al. (1981). They concluded that the sugar ring pucker of *both* strands of this hybrid duplex is in an A-form RNA-like conformation.

Zimmerman and Pheiffer (1981) first suggested that DNA·RNA hybrids can have different conformations depending on the degree of solvation. They found that the X-ray diffraction pattern of poly(rA)·poly(dT) resembled that of A-RNA at 79% relative humidity, whereas when this hybrid is fully hydrated at higher humidity, the sugars of the DNA chain assume a B-like C3′-exo conformation ($P = 198°$) and the RNA sugars assume an A-form C3′-endo conformation ($P = 18°$). Reid et al. (1983) carried out a one-dimensional NOE and circular dichroism study of the hybrid duplex (TCACAT)·(auguga). They concluded that the RNA and DNA sugars both assumed a B-form conformation in solution. The first ^{31}P solid-state study on a hybrid duplex was carried out by Shindo and Matsumoto (1984) on poly(rA)·poly(dT). Their conclusions were consistent with those of Zimmerman and Pheiffer (1981) in that they observed an A-form conformation for the entire duplex below 87% relative humidity, but two *distinct* backbone conformations above 92% relative humidity.

† This work was supported by National Institutes of Health Grants GM-42896 and GM-32681 to B.R.R. M.S. acknowledges the generous support of the W. M. Keck Foundation in the form of a Predoctoral Fellowship.

* Address correspondence to this author at the Department of Chemistry, BG-10, University of Washington, Seattle, WA 98195.

‡ Department of Chemistry, University of Washington.

§ Russian Academy of Sciences.

∥ Department of Biochemistry, University of Washington.

[1] Abbreviations: RNA, ribonucleic acid; DNA, deoxyribonucleic acid; NMR, nuclear magnetic resonance; hybrid, nucleic acid duplex in which a pure DNA strand is base paired to an RNA strand; T_1, longitudinal relaxation time; NOE, nuclear Overhauser effect.

0006-2960/93/0432-4207$04.00/0 © 1993 American Chemical Society

Subsequent ^{31}P NMR (Fugiwara & Shindo, 1985), circular dichroism (Steely et al., 1986), and Raman spectroscopy studies (Benevides & Thomas, 1988) have confirmed the observation of two distinct backbone conformations at high hydration for the poly(rA)·poly(dT) hybrid. Gupta et al. (1985) carried out one-dimensional solution NMR studies on the poly(rA)·poly(dT) hybrid duplex. On the basis of the absence of AH8 to H3′ NOEs in the ribo strand they suggested that *both* the DNA and the RNA strand assumed C1′-exo/C2′-endo-type conformations (P = 108–180°) consistent with a B-form structure for the whole duplex; they further suggested that it is not obligatory for the RNA strand to assume a C3′-endo conformation in order to accommodate the 2′-OH group. Thus the conclusions of Gupta et al. (1985) were in agreement with those of Reid et al. (1983).

Arnott et al. (1986) carried out X-ray fiber diffraction studies on poly(dA)·poly(rU) and poly(dI)·poly(rC). They concluded that in both duplexes the DNA strand assumed a B-like C2′-endo conformation, whereas the RNA strand assumed an A-like C3′-endo conformation; thus the dA·rU hybrid duplex was found to have a *heteromerous* secondary structure. More recently, a high-resolution two-dimensional NMR study on the (CGTTATAATGCG)·(cgcauuauaacg) hybrid duplex was carried out in this laboratory (Chou et al., 1989b). The RNA strand in this duplex assumes a C3′-endo conformation in solution, as was observed in hydrated fibers by Arnott et al. (1986) for their hybrid duplex. However, the DNA strand was tentatively concluded to assume a solution conformation that was in the general S domain but detectably different from the classical B-type C2′-endo structure in its sugar conformations.

One of the barriers to a more detailed study of biologically important hybrid duplexes has been the difficulty in chemically synthesizing RNA strands. Recent advances in chemical RNA synthesis (Ogilvie et al., 1988) have led to the commercial availability of suitably protected RNA phosphoramidites, thus making large-scale synthesis of hybrid duplexes feasible. Here we wish to report on the structure of the synthetic octamer hybrid duplex (GTCACATG)·(caugugac). The conflicting results regarding the duplex structure of DNA·RNA hybrids, particularly in solution, prompted the present study of this duplex, in which the central hexamer is exactly the same as that previously studied by Reid et al. (1983). We present evidence that, despite end effects and the relatively small size of this duplex, the RNA strand retains its C3′-endo configuration, contrary to earlier claims (Reid et al., 1983). Furthermore, our results indicate that the sugars in the DNA strand are neither N-type nor S-type but are in the general E-type conformational domain, further contradicting the previous claims of an all-S-type B conformation for this hybrid duplex sequence.

Finally, we find that, as in DNA·RNA chimeric duplexes (Wang et al., 1992), RNA T_1 values are 2–3 times longer than DNA T_1 values in hybrid duplexes. It is quite possible that the differences between our results and those of previous studies may, in part, be due to this difference in relaxation; i.e., incomplete relaxation of the slower-relaxing RNA protons due to inadequate relaxation delays in the data collection protocol.

MATERIALS AND METHODS

Sample Preparation. Ten-micromole RNA and DNA syntheses were carried out separately on an automated DNA synthesizer (Applied Biosystems Model 392) as previously described (Ogilvie et al., 1988; Chou et al., 1989a,b, 1991). Solid-phase RNA synthesis was carried out using base-protected 5′-DMT 3′-(cyanoethyl *N*,*N*-diisopropylphosphoramidites) protected with *tert*-butyldimethylsilyl (TBDMS) groups at the 2′-position; the phosphoramidites were purchased from Biogenex/ABN and Applied Biosystems. RNA deprotection and cleavage from the support was carried out with anhydrous methanolic ammonia for 18–24 h. The TBDMS protecting groups were removed by dissolving the sample in a 1 M solution of tetrabutylammonium fluoride (TBAF) in THF and letting the solution stand overnight at room temperature. Major impurities were removed from the DNA and RNA samples by ethanol precipitation. The residual TBAF in the RNA was removed by passing the sample through a sodium-form ion-exchange column (AG 50W X-2, Bio-Rad). The RNA and DNA samples were further purified by size exclusion column chromatography (Sephadex DNA grade, superfine, Pharmacia). Column fractions were subjected to microscale electrophoretic analysis on a 20% polyacrylamide gel and pooled accordingly. Equal amounts (ca. 12 mg) of purified RNA and DNA strands were then annealed by dissolving the strands in 1 mL of 500 mM NaCl, 0.2 mM EDTA, and sodium phosphate buffer, pH 6.8, heating the mixture to 55 °C, and slowly cooling the mixture to room temperature. Single-stranded material was subsequently removed by hydroxyapatite column chromatography with a phosphate gradient as described previously (Chou et al., 1989a,b, 1991). The sample was then desalted via Sephadex G-10 column chromatography. The purified sample (ca. 18 mg) was dissolved in 0.4 mL of a buffer consisting of 100 mM NaCl, 1.4 mM EDTA, 25 mM NaH_2PO_4, and 25 mM Na_2HPO_4 and having a pH adjusted to 7.0. After repeated lyophilization from 99.96% D_2O, the sample was finally redissolved in 0.4 mL of 99.996% D_2O for the NMR experiments.

NMR Spectroscopy. All NMR experiments were carried out at 25 °C on either a Bruker AM-500 or a home-built 500-MHz NMR spectrometer (J. Gladden and G. P. Drobny, unpublished design). The data were processed on an IRIS 4D computer using the FTNMR and FELIX software programs (Hare Research, Woodinville, WA). Absolute magnitude COSY data were collected from 1024 complex points in t_2 and 384 t_1 points. A relaxation delay of 5 s was used to ensure adequate (≥65°) relaxation of the RNA protons. The COSY data were apodized with an unshifted sine-bell function. NOESY spectra with mixing times of 45, 90, 100, and 150, and 220 ms were collected in the phase-sensitive mode (States et al., 1982) with a 5-s relaxation delay using 1024 complex points in t_2 and 400 pairs of real and imaginary t_1 experiments. The NOESY data were apodized with a 90°-shifted sine-squared function. Symmetrization was not applied to any of the spectra. The double-quantum filtered COSY (DQF-COSY) spectrum was collected in the phase-sensitive mode with time-proportional phase incrementation (Drobny et al., 1979; Marion & Wüthrich, 1983) from 2048 complex points in t_2 and 746 t_1 points. The DQF-COSY data were apodized with a skewed phase-shifted sine-bell function. The exclusive COSY (E. COSY) spectrum (Griesinger et al., 1985, 1986, 1987) was also collected in the phase-sensitive TPPI mode. For the E. COSY spectrum, maximum sensitivity and resolution in t_2 was achieved by acquiring the spectra in 2048 complex points with 48 scans/t_1 experiment and sample spinning. A total of 700 experiments were collected. The E. COSY data were processed using 6 Hz of exponential line narrowing and 6 Hz of Gaussian broadening in the t_2 and t_1 dimensions. Resolution enhancement was achieved by zero-

Table I: Coupling Constants (Hz) for (GTCACATG)·(caugugac)[a]

residue	G1	T2	C3	A4	C5	A6	T7	G8
$J_{1'-2'}$	o	6.2	6.0	6.9	6.7	6.8	8.0	o
$J_{1'-2''}$	o	6.1	6.6	6.2	6.0	6.6	6.5	o
$J_{2''-3'}$	s	m	m	m(–)	m	m	m	m/s
sugar conformation	*	O4′-endo	O4′-endo	O4′-endo	O4′-endo	O4′-endo	O4′-endo/C1′-exo	*
residue	c16	a15	g14	u13	g12	u11	a10	c9
$J_{1'-2'}$	≤2	≤2	≤2	≤2	≤2	≤2	≤2	≤2
sugar conformation	C3′-endo	C3′-endo	C3′-endo	C3′-endo	C3′-endo	C3′-endo	C3′-endo	C3′-endo

[a] Values obtained from an E. COSY experiment with a digital resolution of 1.07 Hz/point; o = overlapped cross peaks for which the J values could not be measured. The lowercase letters represent RNA residues. For $J_{2''-3'}$ the exact value is very difficult to measure accurately due to the lower intrinsic resolution and passive couplings to other protons and to phosphorus; thus only the relative cross-peak intensities are given (s = strong, m = medium, m(–) = medium-weak).

Table II: Longitudinal Relaxation Times (s) for (GTCACATG)·(caugugac)[a]

residue	G1	T2	C3	A4	C5	A6	T7	G8
H1′	1.84	1.47	–	1.61	1.60	1.63	1.67	1.71
H5			2.20		3.33			
H6		–	1.76		1.77		1.52	–
H8	2.12			1.65		–	1.85	
residue	c16	a15	g14	u13	g12	u11	a10	c9
H1′	2.72	3.24	3.97	4.86	4.05	5.41	–	5.14
H5	3.19			6.40		4.08		–
H6	–			3.17		3.03		3.09
H8		–	–		–	3.43		

[a] The lowercase letters represent RNA residues. The dashes represent T_1 values which could not be measured due to spectral overlap.

filling to 4096 points in t_2 and 2048 points in t_1. With a spectral width of 4386 Hz, the digital resolution in t_2 was thus 1.07 Hz/point. The nonselective T_1 experiments were collected using an inversion–recovery pulse sequence with a composite 180° pulse, and 25 different delays, ranging from 5ms to 8s, were used between the inversion and the observe pulse. For each time point, a total of 64 scans were collected, with a relaxation delay of 22-s between scans. Typical standard deviations from the two-parameter fit ($S(t) = A + Be^{-t/T_1}$) were 0.02–0.05 s.

RESULTS

We analyzed the sugar conformations of the DNA and RNA strands, using a combination of J-coupling analysis and NOE information as previously described (Kim et al., 1992). Briefly, the sugar torsion angles, ν_i, can be obtained from a knowledge of the H–C–C–H dihedral angles, ϕ_i. The sugar torsion angles are related to the pseudorotational phase angle (P) and the puckering amplitude (T_m) as $\nu_i = T_m \cos[P + 144(j-2)]$, where $j = 0$–4 (Altona & Sundaralingam, 1972). The dihedral angles can be estimated from a knowledge of the three-bond coupling constant, $^3J_{HH}$, using either a simple Karplus equation or a modified Karplus relationship that incorporates electronegatively and orientation effects (Haasnoot et al., 1980; van de Ven & Hilbers, 1988). While these treatments yield somewhat different results, their conservative application can be used to restrict torsion angles and sugar conformations to moderately narrow ranges. The empirically calibrated $^3J_{HH}$ values as a function of the respective P values are reproduced from Kim et al. (1992) in Figure 1 for convenience and to facilitate interpretation and discussion of the present results; a pucker amplitude of 38.5° is assumed. The diagnostically useful coupling constants are those that vary widely over a small range of P values, namely, the H1′–H2′, the H3′–H4′, and the H2″–H3′ coupling constants.

Because of the uncertainty in the Karplus equation parameters and in the amplitude of sugar pucker, we used a relatively conservative approach in interpreting the experimentally observed J-coupling values. In our approach, we

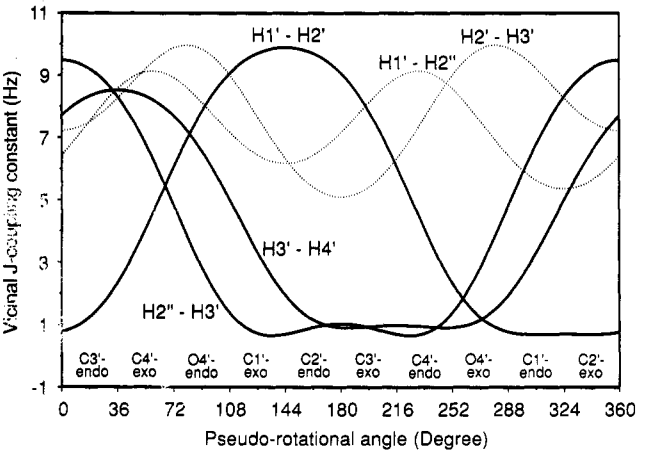

FIGURE 1: Plot of the coupling constant (Hz) as a function of sugar conformation. The P values (pseudorotational phase angle) are in degrees. Reproduced from Kim et al. (1992) and derived from the Karplus constants of Haasnoot et al. (1980).

avoid equating an explicit coupling constant with a specific dihydral angle and we constrain the conformation of the RNA and the DNA sugars to as small a range as possible using a combination of NOE and J-coupling observations. On the basis of the sensitivity of our spectrometer, we used a 1.5-Hz cutoff as the minimum coupling for a coherence cross peak to be experimentally detectable. Thus, an undetectable $J_{1'-2'}$ implies $P < 25$ (ca. C3′-endo sugar conformation), an undectable $J_{2''-3'}$ implies $P > 110°$ (C1′-exo to C2′-endo sugar conformation), and an undectable $J_{3'-4'}$ implies $150° < P \leq 280°$ (C2′-endo to O4′-exo sugar conformation). Furthermore, a comparison of $J_{1'-2'}$ and $J_{2'-2''}$ provides additional constraints. When $J_{1'-2'} > J_{1'-2''}$, then $90° < P$ 190°, and when $J_{1'-2'} \approx J_{1'-2''}$, then $P \approx 90°$ or 200°.

As a first step in the structure analysis of the (GTCACATG)·(caugugac) hybrid duplex, we determined the conformation of the DNA sugars using the above conservative approach. The H1′–H2′ coupling constants can easily be extracted from an E. COSY spectrum provided the sensitivity

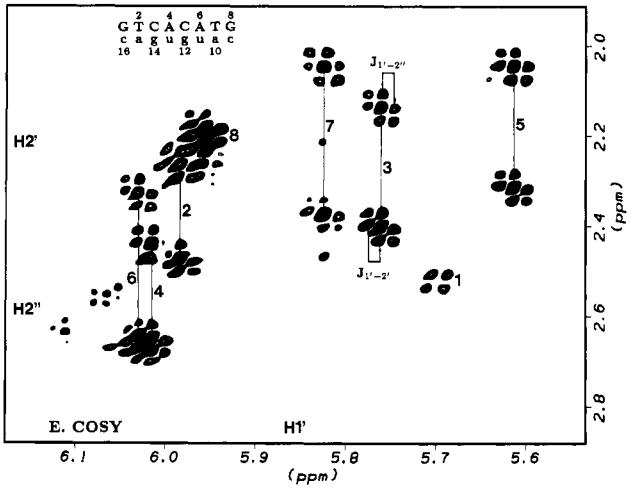

FIGURE 2: H1′–H2′/H2″ above-diagonal region of the E. COSY spectrum of the (GTCACATG)·(caugugac) hybrid duplex. The DNA H1′–H2′ and H1′–H2″ coupling constants were measured from the passive couplings in the t_2 dimension as shown (Griesinger et al., 1987, Bax & Lerner, 1988). The coupling constant values are listed in Table I.

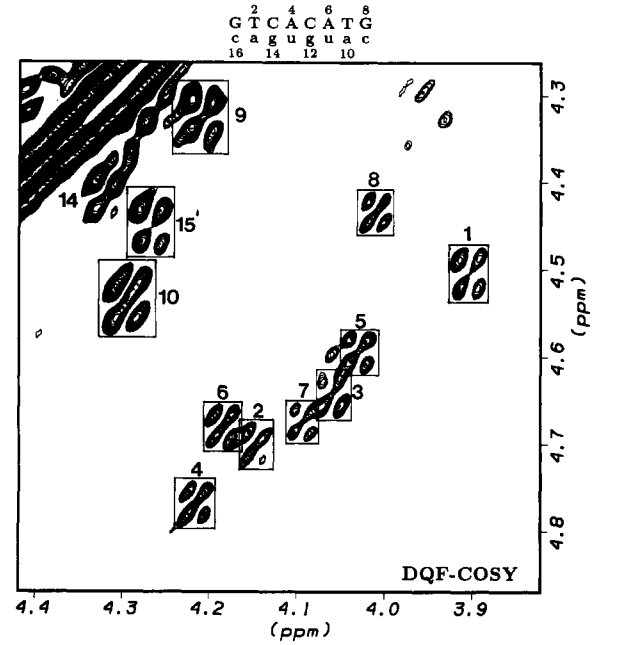

FIGURE 3: DNA H3′–H4′ cross-peak region in the DQF-COSY spectrum of (GTCACATG)·(caugugac). The cross peaks labeled 9, 10, and 15 have been tentatively assigned to the H3′–H4′ cross peaks of the corresponding RNA residues. The close-to-diagonal peak labeled 14 has been tentatively assigned to the H2′–H3′ cross peak of the G14 sugar in the RNA chain. The assignment of the rest of the sugar protons for the RNA residues is complicated by the virtually redundant chemical shifts of their RNA H2′, H3′, H4′, H5′, and H5″ protons. Thus, many of the cross peaks for the RNA residues lie too close to the diagonal to assign with any confidence.

and resolution are adequate. In the case of the (GTCACATG)·(caugugac) hybrid, the resolution in the DNA H1′–H2′/H2″ region was good enough to measure most of the H1′–H2′ and H1′–H2″ coupling constants. Figure 2 shows the H1′–H2′/H2″ E. COSY cross peaks from the DNA strand. To take advantage of the better spectral resolution in t_2, the cross peaks above the diagonal were used to measure the coupling constants as shown. Table I lists the coupling constant values for each residue. It can be seen that the H1′–H2′ and H1′–H2″ coupling constants are approximately equal (6.5 ± 0.4 Hz) for residues T2, C3, A4, C5, and A6 (for T7 they differ by ca. 1.5 Hz). From Figure 1, this indicates that the

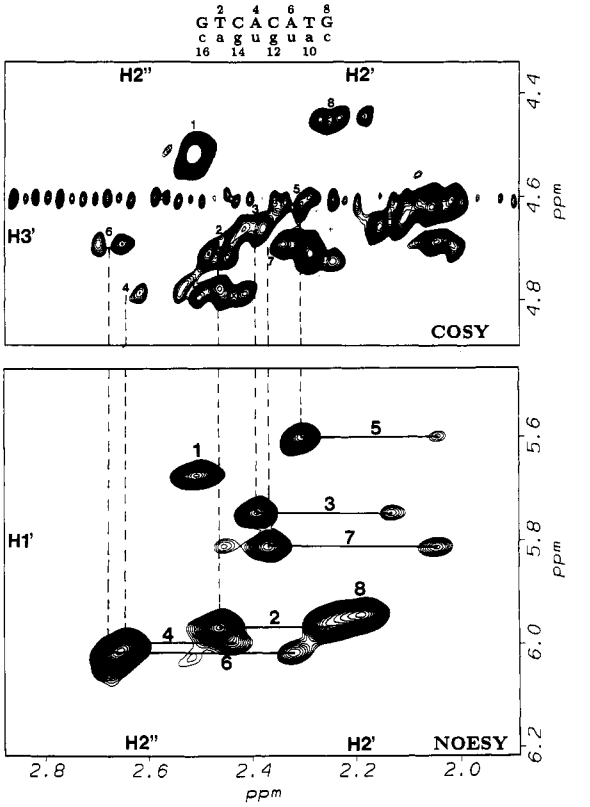

FIGURE 4: H1′–H2′/H2″ (100 ms) NOESY spectrum (bottom) and H3′–H2′/H2″ absolute magnitude COSY spectrum (top) of (GTCACATG)·(caugugac). The dashed vertical lines connect the DNA H1′–H2″ cross peaks in the NOESY spectrum to the H3′–H2″ cross peaks in the COSY spectrum. The H2″–H3′ cross peaks in the COSY spectrum are labeled with small numbers according to the numbering scheme at the top of the figure.

sugar *P* values for these DNA residues are either 90° ± 18°, i.e., in the general E-domain, or 198° ± 18°, i.e., in the general S- to SW-domain.

These two conformational domains can be distinguished by an analysis of the H3′–H4′ coupling constants. Unfortunately, the exact H3′–H4′ coupling constants cannot be determined accurately due to passive couplings to other protons and to phosphorus. However, since the active coupling is the major contributor to the cross-peak *intensity*, the relative intensities of the H3′–H4′ COSY cross peaks, in combination with the H2″–H3′ cross peaks and NOE information, can be used to approximate the range of possible *P* values (Kim et al., 1992). Thus, in the C2′-endo to C3′-exo S-type to SW-type sugar conformational rings (P = 160–216°) the H3′–H4′ coupling is expected to be undetectable ($J \leq 1$ Hz; See Figure 1). Thus, the absence of H3′–H4′ COSY cross peaks would be indicative of *P* values greater than 150° and would favor the S to SW conclusion. Conversely, observation of moderately strong intensity H3′–H4′ COSY cross peaks would imply a *P* value of <150° and favor the E conclusion. Since these cross peaks occur close to the diagonal, the phase-sensitive DQF-COSY is the experiment of choice for making this distinction because it has better resolution than the COSY experiment and better effective sensitivity than the E. COSY experiment.

Figure 3 shows the DNA H3′–H4′ region of the DQF-COSY spectrum of the (GTCACATG)·(caugugac) hybrid duplex. All the DNA H3′–H4′ cross peaks are well resolved and quite strong, and their assignment was straightforward with the aid of the NOESY experiment. From Figure 3, it can be seen that, to a good approximation, all the DNA H3′–H4′ cross peaks have quite similar and quite strong intensities.

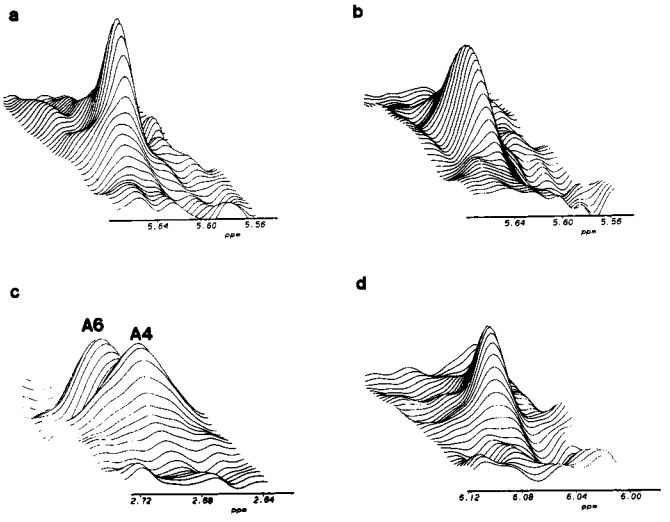

FIGURE 5: Contour plots of NOESY cross peaks. (a) The H1′–H4′ and (b) the H1′–H2′ NOESY cross peaks for residue C5 at 90-ms mixing time. Both cross peaks are of comparable intensity, although the H1′–H2′ cross-relaxation pathway may be contaminated by spin-diffusion. (c) The H2″–H4′ NOESY cross peaks for residues A4 and A6 compared with (d) the H8–H1′ NOESY cross peak for residue A6. These cross peaks are plotted with a scale factor 1.5 times higher than that of (a) and (b). The H2″–H4′ NOESY cross peak intensities are greater than the H8–H1′ NOESY cross peak. Since the AH8–H1′ distance is approximately 3.7–3.9 Å in the anti conformation, the H2″–H4′ distance must be ≤3.7 Å. The observation of strong-intensity H1′–H4′ and H2″–H4′ NOESY cross peaks in combination with the *J*-coupling information rules out the C2′-endo conformation for the DNA residues in the (GTCACATC)·(caugugac) hybrid.

This observation immediately rules out the S/SW-type conformation (P = 180–216°) for the DNA residues in the (GTCACATG)·(caugugac) hybrid. The strong-intensity H3′–H4′ COSY cross peaks, in combination with the H1′–H2′ and H1′–H2″ coupling constants, indicate that the conformation of the DNA sugars is somewhere in the E conformational range. This conclusion can be confirmed by examining the H2″–H3″ region of the COSY spectrum since, from Figure 1, a sugar in an E conformation should have readily detectable H2″–H3′ COSY crosspeaks.

Figure 4 shows the H1′–H2′/H2″ region of the NOESY spectrum (bottom) and the H3′–H2′/H2″ region of the absolute magnitude COSY spectrum (top) for the DNA strand. The most notable feature of the COSY spectrum is the appearance of easily detectable and surprisingly strong H2″–H3′ cross peaks for *all* the DNA residues. In typical B-form DNA duplexes such cross peaks are not detectable. Although the value of the H2″–H3′ coupling constant (and thus a more accurate value for P) could in principle be obtained from an E. COSY spectrum, the presence of passive couplings to phosphorus and to other protons, combined with the lower signal-to-noise ratio in the E. COSY experiment, makes this very difficult. Nevertheless, this region of the absolute magnitude spectrum corroborates the sugar conformations deduced from the other coupling constant data (vide supra). From Figure 1, the presence of these H2″–H3′ cross peaks implies a sugar conformation for the DNA residues with a pseudorotational P value of less than 110°. Indeed, from this observation and the observation of $J_{1'-2'}$ and $J_{1'-2''}$ values of similar magnitude, it is safe to conclude that the DNA residues have rather similar sugar conformations in the O4′-endo range (P = 80–110°). This is the only solution to the conditions $J_{1'-2'} = J_{1'-2''}$ = ca. 7 Hz, $J_{3'-4'} > 1.5$ Hz, and $J_{2''-3'} > 1.5$ Hz. Therefore, it is immediately clear that, for this hybrid duplex, the DNA sugar residues *are not* in the C2′-endo conformation; neither are they in the C3′-endo conformation ($P \approx 18°$). This latter conclusion was confirmed by the observation of (*n*)H6/H8 to (*n*)H2′ NOESY cross peaks that are stronger than the (*n*)H6/H8–(*n* − 1)H2′ cross peaks at short (90 mx) mixing times (data not shown).

Additional evidence for an intermediate conformation for the DNA sugars comes from H1′–H4′ and H2″–H4′ NOE data. Parts a and b of figure 5 show a typical H1′–H4′ NOE cross peak (residue C5) compared to the H1′–H2′ NOE cross peak of the same residue at 90 ms. Thus, considering that at this mixing time the H1′–H2′ NOE intensity may already be exaggerated by spin-diffusion (H1′–H2″–H2′), the H1′–H4′ NOEs are nevertheless stronger than the H1′–H2′ NOEs. Indeed, the H1′–H4′ distances for the DNA residues were found to be 2.6–2.8 Å, as determined from their NOESY buildup rates at 34-, 90-, 150-, and 220-ms mixing times. This is independent evidence that the DNA residues *are not* in the C2′-endo conformational range; neither are they in the C3′-endo range, since in these conformational ranges the H1′-H4′ distances are ca. 3.2–3.4 Å. Figure 5c shows two well-resolved H2″–H4′ NOE cross peaks at 90-ms mixing time corresponding to DNA residues A4 and A6. In the C2′-endo conformation, the H2″–H4′ distance is ca. 4.1 Å. Therefore, these NOEs should be very weak if the conformation of the DNA residues were S-type. Figure 5d shows a typical (*n*)H8 to (*n*)H1′ NOE cross peak (residue A6) at 90-ms mixing time (distance ca. 3.7–3.9 Å in the anti conformation). Clearly, the broader H2″–H4′ NOE intensities are greater than the H8 to H1′ NOE intensity. Indeed, the individual H2″–H4′ cross peaks shown in Figure 5c integrate to ca. twice the intensity of the A6 H8–H1′ cross peak in Figure 5d and correspond to distances of ca. 3.5 Å.

We next investigated the conformation of the RNA strand in a similar manner. The most diagnostic parameter by which to ascertain whether the RNA sugars are in the N or S (or E) conformation is the H1′–H2′ coupling, which is very strong (ca. 10 Hz) for the S conformation and very weak (<2 Hz) for the N conformation; see Figure 1. Figure 6a shows a 100-ms NOESY spectrum of the H1′–H2′ cross peaks for RNA residues 9–16 (the H1′–H2′ distance is chemically constrained to 2.70–3.05 Å, and this cross peak is *always* the strongest intrasugar NOE in this spectral region regardless of conformation). The corresponding region of the DQF-COSY spectrum is shown in Figure 6b, and it is immediately apparent that the RNA H1′–H2′ scalar couplings are undetectable. However, a note of caution is in order. It has been demonstrated (Wang et al., 1992) that in chimeric duplexes and RNA duplexes the T_1 relaxation times of RNA H1′ protons are ca. 2.5 times longer than those of the DNA H1′ protons; we have shown this to be tue also in hybrid chimeric duplexes (M. Salazar and B. R. Reid, submitted for publication). We therefore carried out nonselective T_1 measurements, and Table II shows that the T_1s of the RNA H1′ protons in the hybrid duplex are in the range 4 ± 1.4 s, whereas the DNA H1′ proton T_1s are in the range 1.65 ± 0.2 s. Thus, in spectra with typical (2–3 s) relaxation delays, weak H1′–H3′ COSY cross peaks might simply reflect incomplete recovery (partial saturation) of the slower-relaxing RNA H1′ proton. For these reasons, the DQF-COSY spectrum in Figure 6b was acquired with a 5.46-s duty cycle (5-s relaxation delay) to ensure that all ribose protons were at least 64% relaxed. Nevertheless, all RNA H1′–H3′ cross peaks remain undetectable, and we therefore conclude that all sugars in the RNA strand of the hybrid are in the general N-type conformational domain ($J_{1'-2'} < 2$ Hz), close to the

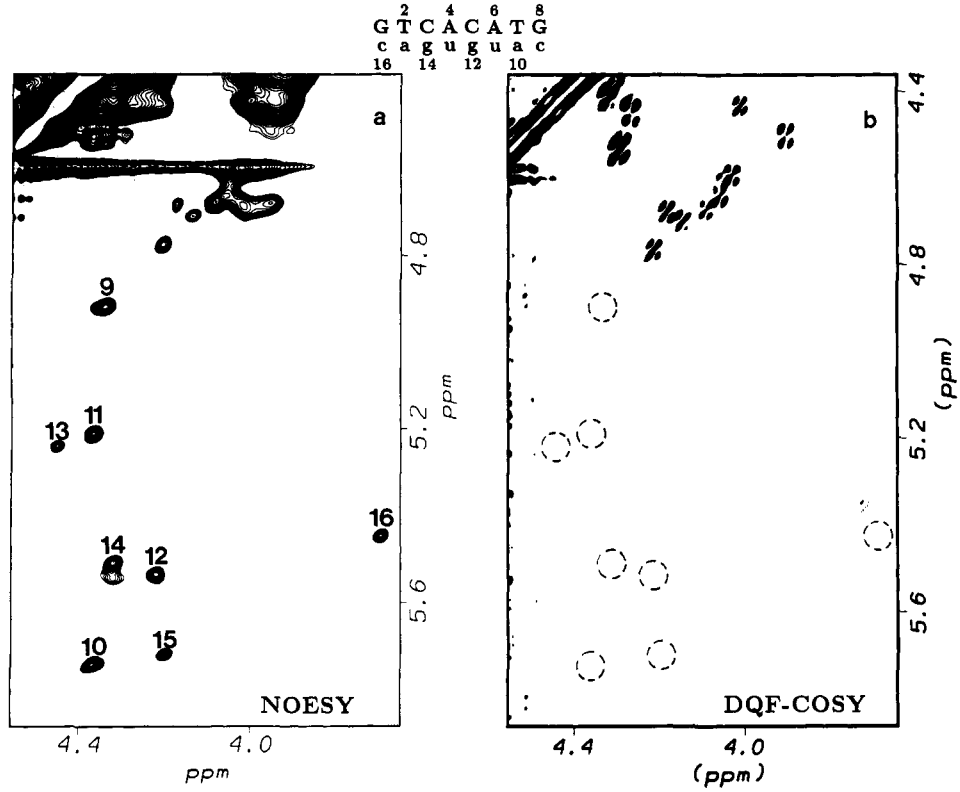

FIGURE 6: (a) 100-ms NOESY spectrum and (b) DQF-COSY spectrum of the RNA H1′–H2′/H3′/H4′/H5′/H5″ region of (GTCACATG)·(caugugac). The top region of the DQF-COSY spectrum is the same region as shown in Figure 3. The RNA H1′–H2′ NOESY cross peaks are labeled by residue number. The dashed circles in the DQF-COSY spectrum are the regions where the RNA H1′–H2′ cross peaks would appear if the coupling constant were >1.5 Hz. The duty cycle in the DQF-COSY spectrum was increased to 5.46 s by the use of a 5-s relaxation delay.

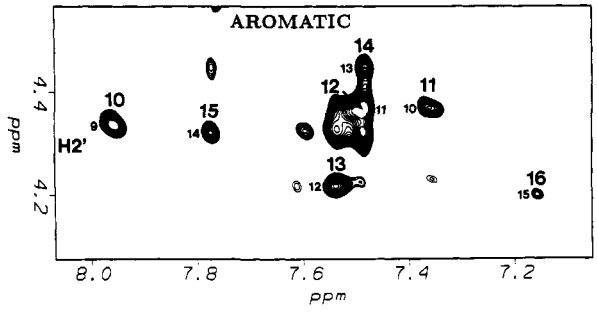

FIGURE 7: RNA H8/H6–H2′ and RNA/DNA H8/H6–H3′/H4′/H5′/H5″ regions of the NOESY spectrum of (GTCACATG)·(caugugac). At 100-ms mixing time, the RNA H8/H6–(n – 1)H2′ cross peaks (distance ca. 2.0–2.1 Å) are by far the strongest peaks and can be readily identified. The H8/H6 shifts are designated with the larger residue numerals, and the H2′ shifts, by smaller residue numerals.

C3′-endo conformation. Furthermore, the entire RNA strand exhibits A-form grometry in that the base H8/H6 protons exhibit strong NOEs to the (n – 1)H2′ but not to their own H2′, as shown in Figure 7. Such an NOE pattern is characteristic of A-form strand geometry, where the (n)H8/H6 to (n – 1)H2′ distance is ca. 2.0–2.1 Å.

The above NOE and J-coupling data were used to construct a model of the hybrid (GTCACATG)·(caugugac) duplex that satisfies these constraints. We used two starting structures: one with canonical A-form grometry and one with the coordinates of the Arnott hybrid model derived from fiber diffraction data (Arnott et al., 1986). The root mean square deviation (RMSD) between these two starting structures is ca. 1.3 Å. Both structures were then minimized using the AMBER force field (DISCOVER software package, BIOSYM Technologies, Inc., San Diego, CA) supplemented with our experimental constraints. For the DNA strand only, upper bounds of 3.65 Å for the H2″–H4′ intrasugar distances and 2.8 Å for the H1′–H4′ intrasugar distances were used to constrain the DNA sugars. No constraints were used for the RNA sugars, which retained a C3′-endo conformation throughout the minimization procedure. A cross-eye stereoview of the superimposed resulting structures is shown in Figure 8. Thus, a good convergence to the same structure was obtained from both starting models as indicated by an RMSD of 0.16 Å between the two final structures.

DISCUSSION

The present J-coupling/NOE study has been interpreted in terms of *single* nucleoside conformation rather than a model incorporating N ↔ S (C3′-endo to C2′-endo) interconversions for the DNA residues. These nucleoside dynamic equilibrium models (Rinkel & Altona, 1987; Schmitz et al., 1990; Stolarski et al., 1992) have not been documented experimentally and are frequently invoked merely to rationalize or fit the observed experimental couplings to the allegedly *precise* coupling constants, and sometimes this dynamic assumption is not required (Celda et al., 1989; Gochin et al., 1990). We note that recent ^{2}H solid-state NMR studies of DNA duplexes at high hydration levels indicate only *small-amplitude* motions in the DNA bases, sugars, and backbone linkers and argue against the existence of C2′-endo to C3′-endo ($\Delta P \approx 150°$) sugar ring interconversions (Kintanar et al., 1989; Huang et al., 1990; Alam et al., 1991). Thus, the inability to fit the observed experimental scalar couplings may simply be due to inadequately parameterized Karplus constants in the equation used to determine the sugar dihedral angles. We emphasize that we interpret our data in terms of a range of sugar conformations (albeit small) and *not* in terms of an absolutely

G T C A C A T G
c a g u g u a c

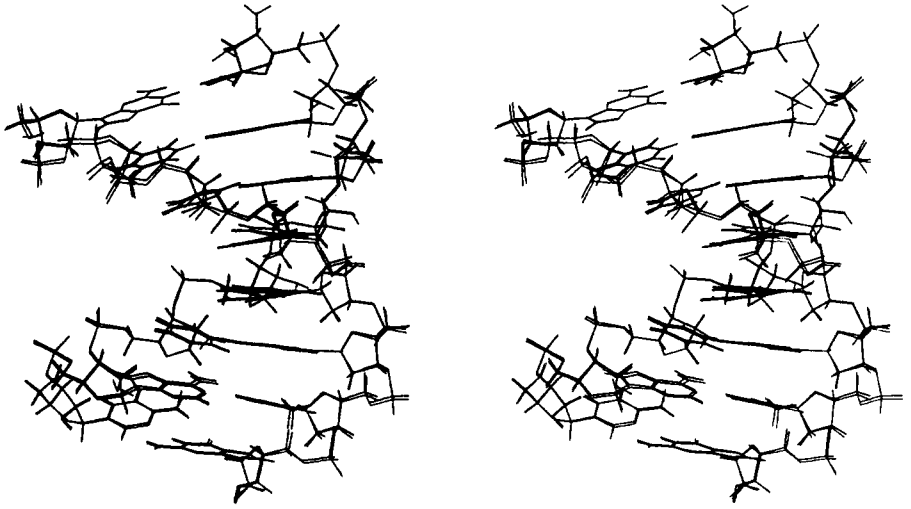

FIGURE 8: Cross-eye stereoview of the DNA·RNA hybrid (GTCACATG)·(caugugac) structure produced by distance-constrained energy minimization. The two structures were refined from a canonical A-form structure and a structure constructed from the coordinates of Arnott et al. (1986) and are superimposed; the RMSD is 0.16 Å.

rigid sugar ring conformer. A further point worth noting is that our observation of an H1′–H4′ distance of 2.6–2.8 Å in the DNA nucleosides is inconsistent with *any* admixture of N and S conformers since both of these conformers have H1′–H4′ distances > 3.2 Å.

Our motivation for studying the DNA·RNA hybrid (CTCACATG)·(caugugac) was the current widespread ambiguity and inconsistency surrounding the proposed structure for the polynucleotide strands in DNA·RNA hybrids. The central hexamer of this hybrid octomer we studied has previously been investigated by Reid et al. (1983) using circular dichroism and one-dimensional NOE methods at a proton frequency of 400 MHz. We have studied this octamer at a proton frequency of 500 MHz using well-established two-dimensional NMR methods. Our results and conclusions differ markedly from those of Reid and co-workers, and from those of Gupta et al. (1985), in two important aspects. Firstly, the DNA strand does not assume a B-type C2′-endo conformation, as they have claimed. Rather, the sugars are in a conformation that is neither S-type (C2′-endo) nor N-type (C3′-endo), but is instead in the general E-type (O4′-endo) domain. Secondly, and most importantly, the RNA strand is definitely not in the C2′-endo conformation as claimed by Reid et al. (1983) and by Gupta et al. (1985); this is easily and conclusively demonstrated by the absence of detectable H1′–H2′ COSY cross peaks for the ribose sugars and by strong H6/H8 NOESY cross peaks to the $(n-1)$H2′ (stronger than those to the (n)H2′) for the RNA strand at short mixing times.

The sugar residues in the DNA strand of the structure shown in Figure 8 have final P values of 106° ± 3°, while the sugar residues of the RNA strand retain a C3′-endo conformation with P values of 15° ± 3°. Interestingly, although the strands in this hybrid duplex are obviously heteromerous with respect to nucleoside conformation, they exhibit complementary base pairing and the residues of each strand are in register, with the same average rise of ca. 2.7 Å/residue on each strand. This rise per residue value is more typical of A-form structure than of B-form structure. Thus, the backbone of the DNA strand is condensed to this reduced rise via changes in the sugar and backbone conformations. However, the ^{31}P spectrum of this hybrid duplex is normal (data not shown), suggesting that these changes in the backbone phosphodiester torsion angles α and ζ are small and do not involve B_{II} (g^-,t) phosphates (Chou et al., 1992a,b; Cheng et al., 1992).

Our results are in good agreement with previous preliminary studies on the DNA·RNA dodecamer hybrid (CGTTATAATGCG)·(cgcauuauaacg) by Chou et al., (1989b) and are qualitatively similar to the strand conformations proposed for the hydrated fiber state of poly(dA)·poly(rU) by Arnott et al. (1986). From the published coordinates we examined the poly(dA)·poly(rU) hybrid structure of Arnott et al. (1986) and found, to our surprise, that the sugar conformations are actually C1′-exo (P = 108–144°) instead of C2′-endo. Thus, the hydrated fiber averaged structure proposed by Arnott et al. (1986) contains DNA sugars that are not in the C2′-endo (P = 162° ± 18°) conformation, as the abstract in their report suggests (Arnott et al., 1986), but actually have a significantly lower P value. In reality their sugar conformation is not too far from the conformation that we find in solution for the DNA sugars in the hybrid (GTCACATG)·(caugugac) duplex.

Reid and co-workers (1983) concluded that the DNA and RNA strand in a hybrid duplex have similar C2′-endo conformations; this conclusion was based on their claim of similar NOE patterns for both strands. Gupta et al. (1985), despite the observation of very strong generic (unassigned) H8 to H2′ NOEs, concluded that both the RNA and the DNA strands in the hybrid duplex poly(rA)·poly(dT) are in the C2′-endo conformation; this conclusion was based mainly on the *absence* of H8 to H3′ NOEs in the rA strand. Reid et al. (1983) used an irradiation time of 500 ms in their one-dimensional NOE studies and reported using a relaxation delay ca. 5 times the *longest* proton T_1, although the proton T_1 values were not reported. Gupta and co-workers (1985) used reasonably short irradiation times (25 ms, 19 dB; and 40 ms, 20 dB) but made no mention of the relaxation delays they used. We used the same salt concentration (100 mM NaCl) as that used by Reid et al. (1993), and only the sample

temperature was different (5 °C in their studies and 25 °C in our studies). Apart from minor end-fraying effects reflected in the chemical shift redundancy of the H2′ and the H2″ protons for G1 and G9, we consider it unlikely that a temperature of only 25 °C would change the conformation of *all* the residues in this DNA·RNA hybrid. The observation of interstrand H2–H1′ NOESY cross peaks (data not shown) indicates that the RNA and DNA strands form a stable heteroduplex at this moderate temperature. Furthermore, the observation of NOEs from methyl to $(n - 1)$ aromatic protons (data not shown) indicates a relatively normal, right-handed helical conformation for the DNA strand.

The disagreement between our results and those described above may very well be due to second-order spin-diffusion effects involving the anomeric protons and/or the use of relaxation times that were too short, leading to serious partial saturation effects on the slowly recovering anomeric and aromatic RNA protons in the Reid et al. and Gupta et al. studies. Moreover, due to the poor resolution in their spectra, Reid et al. (1983) were unable to assign the RNA H2′ protons and were thus unable to document and interpret the strong H8/H6 $(n - 1)$H2′ NOEs that are diagnostic of A-form structure. Regarding the CD data presented by Reid et al. (1983) and Gupta et al. (1985), we have found that even when double-stranded RNA is present internally in a helix, as in a D–R–D chimeric duplex, the CD spectra exhibit maxima and minima and molar ellipticities that are only slightly different from pure double-stranded DNA but have some features similar to double-stranded RNA (Wang et al. 1992). In fact, in chimeric duplexes where the DNA segment retains a C2′-endo conformation (except for the junction base pair) and the RNA segment retains a C3′-endo conformation (Chou et al., 1991; M. Salazar and B. R. Reid, in preparation) the CD spectra appear to reflect a proportionate superposition of A-form and B-form spectra (Wang et al., 1992). Thus CD spectra alone cannot serve as good indicators of structure where DNA and RNA are either covalently linked or hydrogen-bonded in a duplex.

Our study conclusively rules out the possibility of a B·B structure or an A·A structure for hybrid duplexes in solution. Although we expect there to be sequence-dependent variations in local structure, our results, and those from previous studies in this laboratory (Chou et al., 1989b; Salazar et al., 1993), suggest that a heteromerous duplex, in which the RNA strand retains its C3′-endo conformation and the DNA strand assumes an intermediate conformation with E-type sugars, may well be a general feature of DNA·RNA hybrid duplexes in solution. This finding may be of particular importance in understanding the action of RNase H on DNA·RNA hybrids. We note that Nakamura et al. (1991), in modeling a proposed complex of RNase H with a hybrid duplex, assumed that both strands of the RNA·DNA hybrid duplex were in an A-form structure with N-type sugars. Similar assumptions were also made by Kohlstaedt et al. (1992) in modeling the binding of a hybrid duplex by HIV-1 reverse transcriptase. More recently Katayanagi et al. (1992) have used a heteromerous (A·B) RNA·DNA hybrid in modeling the proposed complex of RNase H from *Escherichia coli* with an RNA·DNA hybrid. Although in the above studies both A·A and A·B hybrid duplex models produced similar results, and hybrid duplexes appear to be more A-like than B-like, we believe that neither model is completely appropriate for the DNA strand.

Finally, we note that a completely independent immunochemical approach using anti-hybrid antibodies (Sanford et al., 1988) results in essentially the same conclusions as our structural studies. The immunochemical studies indicate that the hybrid sections of the 36-base-pair duplex $rA_{12}d(GC)_6d(T)_{12}$, while having "some features of the A-helix", adopt a unique conformation which is distinct from both B-DNA and A-RNA.

ACKNOWLEDGMENT

The authors would like to thank Drs. S.-G. Kim and P. F. Flynn for technical discussions.

REFERENCES

Alam, T. M., Orban, J., & Drobny, G. P. (1991) *Biochemistry 30*, 9229–9237.
Altona, C., & Sundaralingam, M. (1972) *J. Am. Chem. Soc. 94*, 8205–8212.
Arnott, S., Chandrasekaran, R., Hall, I. H., & Puigjaner, L. C. (1983) *Nucleic Acids Res. 11*, 4141–4155.
Arnott, S., Chandrasekaran, R., Millane, R. P., & Park, H.-S. (1986) *J. Mol. Biol. 188*, 631–640.
Bax, A., & Lerner, L. (1988) *J. Magn. Reson. 79*, 429–438.
Benevides, J. M., & Thomas, G. J., Jr. (1988) *Biochemistry 27*, 3868–3873.
Celda, B., Widmer, H., Leupin, W., Chazin, W. J., Denny, W. A., & Wüthrich, K. (1990) *Biochemistry 28*, 1462–1471.
Cheng, J.-W., Chou, S.-H., & Reid, B. R. (1992) *J. Mol. Biol. 228*, 1037–1041.
Chou, S.-H., Flynn, P., & Reid, B. R. (1989a) *Biochemistry 28*, 2422–2435.
Chou, S.-H., Flynn, P., & Reid, B. R. (1989b) *Biochemistry 28*, 2435–2443.
Chou, S.-H., Flynn, P., Wang, A., & Reid, B. R. (1991) *Biochemistry 30*, 5248–5257.
Chou, S.-H., Cheng, J.-Y., Fedoroff, O. Y., Chuprina, V. P., & Reid, B. R. (1992a) *J. Am. Chem. Soc. 114*, 3114–3115.
Chou, S.-H., Cheng, J.-W., & Reid, B. R. (1992b) *J. Mol. Biol. 228*, 138–155.
Drobny, G., Pines, A., Sinton, S., Weitekamp, D. P., & Wemmer, D. E. (1970) *Faraday Symp. Chem. Soc. No. 13*, 49–55.
Fujiwara, T., & Shindo, H. (1985) *Biochemistry 24*, 896–902.
Gochin, M., Zon, G., & James, T. L. (1990) *Biochemistry 29*, 11161–11171.
Gray, D. M., & Ratliff, R. L. (1975) *Biopolymers 14*, 487–498.
Griesinger, C., Sørensen, O. W., & Ernst, R. R. (1985) *J. Am. Chem. Soc. 107*, 6394–6396.
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.
Gupta, G., Sarma, M. H., & Sarma, R. H. (1985) *J. Mol. Biol. 186*, 463–469.
Gupta, G., Sarma, M. H., & Sarma, R. H. (1986) *Int. J. Quant. Chem., Quant. Biol. Symp. 12*, 183–199.
Haasnoot, C. A. G., de Leeuw, F. A. A. M., & Altona, C. (1980) *Tetrahedron 36*, 2783–2792.
Huang, W.-C., Orban, J., Kintanar, A., Reid, B. R., & Drobny, G. P. (1990) *J. Am. Chem. Soc. 112*, 9059–9068.
Katayanagi, K., Miyagawa, M., Matsushima, M., Ishikawa, M., Kanaya, S., Nakamura, H., Ikehara, M., Matsuzaki, T., & Morikawa, K. (1992) *J. Mol. Biol. 223*, 1029–1052.
Kim, S.-G., Lin, L.-J., & Reid, B. R. (1992) *Biochemistry 31*, 3564–3574.
Kintanar, A., Huang, W.-C., Schindele, D. C., Wemmer, D. E., & Drobny, G. (1989) *Biochemistry 28*, 282–293.
Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., & Steitz, T. A. (1992) *Science 256*, 1783–1790.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Milman, G., Langridge, R., & Chamberlin, M. J. (1967) *Proc. Natl. Acad. Sci. U.S.A. 57*, 1804–1810.

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.

O'Brien, E. J., & MacEwan, A. W. (1970) *J. Mol. Biol. 48*, 243–261.

Ogilvie, K. K., Usman, N., Nicoghosian, K., & Cedergren, R. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 5764–5768.

Pardi, A., Martin, F. H., & Tinoco, I., Jr. (1981) *Biochemistry 20*, 3986–3996.

Reid, D. G., Salisbury, S. A., Brown, T., Williams, D. H., Vasseur, J.-J., Rayner, B., & Imabach, J.-L. (1983) *Eur. J. Biochem. 135*, 307–314.

Rich, A. (1960) *Proc. Natl. Acad. Sci. U.S.A. 46*, 1044–1053.

Rinkel, L. J., & Altona, C. (1987) *J. Biomol. Struct. Dyn. 4*, 621–631.

Salazar, M., Champoux, J. J., & Reid, B. R. (1993) *Biochemistry 32*, 739–744.

Sanford, D. G., Kotkow, K. J., & Stollar, B. D. (1988) *Nucleic Acids Res. 16*, 10643–10655.

Schmitz, U., Zon, G., & James, T. L. (1990) *Biochemistry 29*, 2357–2368.

Shindo, H., & Matsumoto, U. (1984) *J. Biol. Chem. 259*, 8682–8684.

States, D. J., Haberkorn, R. A., & Reuben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.

Steely, H. T., Jr., Gray, D. M., & Ratliff, R. L. (1986) *Nucleic Acids Res. 24*, 10071–10090.

Stolarski, R., Egan, W., & James, T. L. (1992) *Biochemistry 31*, 7027–7042.

Tunis, M.-J. B., & Hearst, J. E. (1968) *Biopolymers 6*, 1218–1223.

van de Ven, F. J. M., & Hilbers, C. W. (1988) *Eur. J. Biochem. 178*, 1–38.

Wang, A. C., Kim, S.-G., Chou, S.-H., Orban, J., Flynn, P., & Reid, B. R. (1992) *Biochemistry 31*, 3940–3946.

Zimmerman, S. B., & Pheiffer, B. H. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 78–82.