Ti, G. S., Gaffney, B. L., & Jones, R. A. (1982) *J. Am. Chem. Soc. 104*, 1316.
Usman, N., Pon, R. T., & Ogilvie, K. K. (1985) *Tetrahedron Lett. 26*, 4567.
Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., van der Marel, G., van Boom, J. H., & Rich, A. (1981) *Science 211*, 171.
Wemmer, D. E., Chou, S. H., Hare, D. R., & Reid, B. R. (1984a) *Biochemistry 23*, 2262.
Wemmer, D. E., Chou, S. H., & Reid, B. R. (1984b) *J. Mol. Biol. 180*, 41.
Wemmer, D. E., Chou, S. H., & Reid, B. R. (1985) *Nucleic Acids Res. 13*, 3755.
Westerink, H. P., van der Marel, G. A., van Boom, J. H., & Haasnoot, C. A. G. (1984) *Nucleic Acids Res. 12*, 4323.
Zimmerman, S. B., & Pheiffer, B. H. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 78.

High-Resolution NMR Study of a Synthetic DNA–RNA Hybrid Dodecamer Containing the Consensus Pribnow Promoter Sequence: d(CGTTATAATGCG)·r(CGCAUUAUAACG)†

Shan-Ho Chou,‡,§ Peter Flynn,‖ and Brian Reid*,§,‖

Howard Hughes Medical Institute, Biochemistry Department, and Chemistry Department, University of Washington, Seattle, Washington 98195

Received July 5, 1988; Revised Manuscript Received November 3, 1988

ABSTRACT: The nonsymmetrical double-helical hybrid dodecamer d(CGTTATAATGCG)·r(CGCAUUAUAACG) was synthesized with solid-phase phosphoramidite methods and studied by high-resolution 2D NMR. The imino protons were assigned by one-dimensional nuclear Overhauser methods. All the base protons and H1′, H2′, H2″, H3′, and H4′ sugar protons of the DNA strand and the base protons, H1′, H2′, and most of the H3′–H4′ protons of the RNA strand were assigned by 2D NMR techniques. The well-resolved spectra allowed a qualitative analysis of relative proton–proton distances in both strands of the dodecamer. The chemical shifts of the hybrid duplex were compared to those of the pure DNA double helix with the same sequence (Wemmer et al., 1984). The intrastrand and cross-strand NOEs from adenine H2 to H1′ resonances of neighboring base pairs exhibited characteristic patterns that were very useful for checking the spectral assignments, and their highly nonsymmetric nature reveals that the conformations of the two strands are quite different. Detailed analysis of the NOESY and COSY spectra, as well as the chemical shift data, indicate that the RNA strand assumes a normal A-type conformation (C3′-endo) whereas the DNA strand is in the general S domain but not exactly in the normal C2′-endo conformation. The overall structure of this RNA–DNA duplex is different from that reported for hybrid duplexes in solution by other groups (Reid et al., 1983a; Gupta et al., 1985) and is closer to the C3′-endo–C2′-endo hybrid found in poly(dA)·poly(dT) and poly(rU)·poly(dA) in the fiber state (Arnott et al., 1983, 1986).

DNA–RNA hybrids play an important role in biological information transfer in such processes as transcription of DNA into messenger RNA, transfer RNA, and ribosomal RNA, reverse transcription of viral RNA sequences into DNA sequences, replication of DNA via Okazaki fragments in which DNA–RNA primer duplexes are extended into long DNA chains, etc. Despite their biological importance, relatively little detailed information on the conformation of DNA–RNA hybrids is available at the present time; this is mostly due to the difficulty of synthesizing RNA chains. The synthetic DNA–RNA hybrids poly(dA)·poly(rU) and poly(dI)·poly(rC) were found to adopt a structure in which the DNA chains are in the C2′-endo conformation and the RNA chains are in the C3′-endo conformation (Arnott et al., 1986); the structure of poly(rA)·poly(dT) was found to be able to adopt either an all C3′-endo conformation at low humidity or a C3′-endo conformation for the RNA chain and a B-like C3′-exo conformation for the DNA chain at high solvation in the fiber state (Zimmerman & Pheiffer, 1981). In solution poly(rA)·poly(dT) and the small hexameric hybrid d(TCACAT)·r(AUGUGA) were also tentatively assigned as being in an all C2′-endo (B-type) conformation from one-dimensional NOE data (Gupta et al., 1985; Reid et al., 1983a). The differences may be due to sequence effects or to the different conditions used in each study, but it is also possible that synthetic homopolymers or small oligomer duplexes may not be good models for true RNA–DNA hybrid structures. We therefore decided to synthesize a full-turn DNA–RNA hybrid helix and to use now well-established 2D NMR techniques to study the structure of this important molecule. In the preceding paper (Chou et al., 1989) we described new developments in solid-phase RNA synthesis that now make it possible to produce large amounts of synthetic RNA strands for 2D NMR studies on RNA–RNA and RNA–DNA duplexes. We now report a qualitative structural study on a synthetic DNA–RNA promotor duplex and compare it to the analogous DNA duplex which has been studied previously by NMR (Wemmer et al., 1984). The duplex d(CGTTATAATGCG)·r(CGCAUUAUAACG) contains seven internal base paired adenine residues, and the seven adenine H2 protons were important in assigning the hybrid spectra as well as in revealing critical

† B.R. acknowledges the support of NIH Grant PO1 GM32681.
‡ Howard Hughes Medical Institute.
§ Biochemistry Department.
‖ Chemistry Department.

0006-2960/89/0428-2435$01.50/0

clues to the hybrid duplex conformation.

Materials and Methods

The procedure for rapid and efficient solid-phase RNA synthesis using 5′-DMT 3′-(*N*,*N*-diisopropylmethoxyphosphoramidites) protected with TBDMS at the 2′-position was described in the preceding paper. Ten-micromole RNA and DNA syntheses were carried out separately on an automated DNA synthesizer (Applied Biosystems Model 380B). The d(CGTTATAATGCG) and r(CGCAUUAUAACG) products were purified by analytical TLC, as described previously [see preceding paper (Chou et al., 1989)]. Equal amounts (about 10 mg) of each strand were combined in 50 mL of 10 mM phosphate buffer, heated to 80 °C, slowly cooled to room temperature, and then chromatographed on a hydroxyapatite column with a phosphate gradient to remove single-stranded material, as described in the preceding paper.

NMR Spectroscopy. All NMR experiments used in this study were performed on a Bruker WM-500 NMR spectrometer at a temperature of 31 °C. Exchangeable proton spectra were obtained with a Redfield solvent-suppression scheme with the carrier frequency set at approximately 12 ppm and a total pulse length of 275 μs. NOE difference spectra were collected by coaveraging 16 scans with applied radio-frequency (rf) irradiation at a desired resonance position and subtracting that sum from an equal number of scans with the rf off-resonance; a total of 3200 scans were accumulated.

Absolute-magnitude COSY data were collected from 1024 complex points in t_2 and 400 t_1 points. NOESY spectra with mixing times of 100 and 300 ms were acquired in the phase-sensitive mode (States et al., 1982), from 1024 complex t_2 points and 400 pairs of real and imaginary t_1 experiments.

The COSY and NOESY data were processed with the FTNMR software package written by Dennis Hare (Hare Research, Woodinville WA.). COSY data were apodized with an unshifted sine-bell function. NOESY data were apodized with a skewed phase-shifted sine-bell function. Noise ridges in the t_1 dimension were attenuated by multiplying the first row by half prior to transformation. Finally, after Fourier transformation, COSY and NOESY spectra were subjected to diagonal low-point symmetrization (Baumann et al., 1981).

Results

Although the NOE data to be presented can be described in the abstract without a structural model, we found it useful to use the coordinate data for the homopolymer hybrid poly(dA)·poly(dT), representing a C3′-endo/C2′-endo structure, and poly(dA)·poly(rU), representing a C2′-endo/C3′-endo structure (Arnott et al., 1983, 1986), as reference models for qualitative analysis and comparison of our NOE data. Although the data do not agree with all aspects of these structures, they nevertheless serve as valuable base-line models. The three base pairs (dAAA·rUUU and dAAA·dTTT) were therefore generated with these model coordinates (including all hydrogen atoms) and are shown in Figure 1a,b.

Imino Proton Studies. The imino proton and base proton spectrum and the imino NOE data are shown in Figure 2. Ten imino resonances are seen at 30 °C accounting for 10 of the 12 base pairs, with the imino protons of the two terminal base pairs exchanging with solvent too fast to be detected. Assignment of the imino protons was complicated by the fact that six of ten resonances are very close in chemical shift. However, the imino assignements were simplified by previous observations on adenine H2 proximities. For intrastrand NOEs, in a right-handed helical nucleic acid, the $n-1$ adenine H2 is always closer to the imino proton being irradiated (about

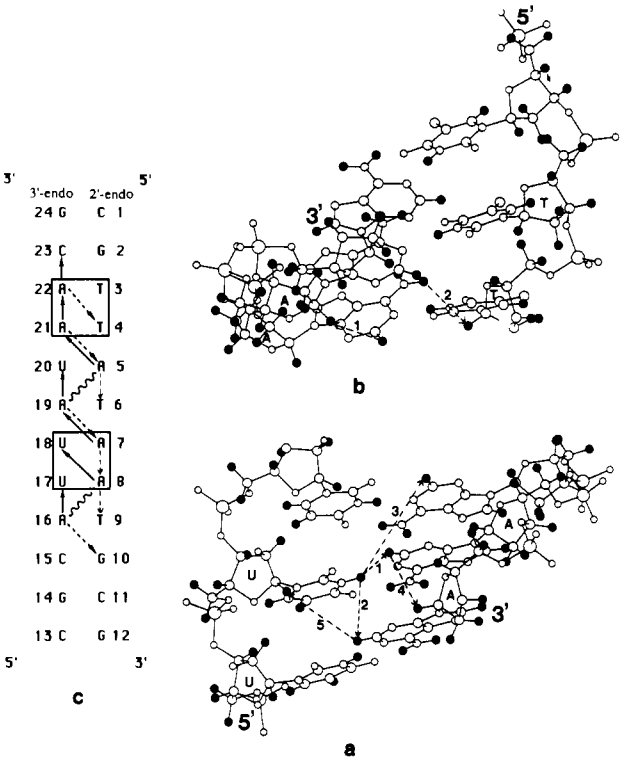

FIGURE 1: (a) Computer drawing of a hybrid model of r(UUU)·d(AAA) using the coordinates from Arnott et al. (1986). The r(UUU) is in the C3′-endo conformation, and the d(AAA) is in the C2′-endo conformation. Some proton pairs discussed in the text are connected by dotted lines. (Path 1) Imino proton to the H2 of the complementary adenine (2.8 Å); (path 2) imino proton to the cross-strand $m+1$ A H2 (3.99 Å); (path 3) imino proton to the cross-strand $m-1$ A H2 (5.8 Å); (path 4) intra-stranded A H2 to the $n+1$ H1′ (4.5 Å); (path 5) cross-stranded A H2 to $m+1$ H1′ (4.19 Å). (b) Computer drawing of a hybrid model of d(AAA)·d(TTT) using the coordinates from Arnott et al. (1983). The d(AAA) is in the C3′-endo conformation, and the d(TTT) is in the C2′-endo conformation. The large propeller twisting between base pairs is clearly seen. Some proton pairs discussed in the text are connected by dotted lines. (Path 1) Intrastranded A H2 to $n+1$ H1′ (3.16 Å); (path 2) cross-stranded A H2 to $m+1$ H1′ (3.92 Å). (c) Summary of the experimentally observed connectivities from A H2 to $n+1$ and $m+1$ H1′. The solid arrows and dotted arrows represented strong and weak NOEs, respectively. The lower box includes the corresponding duplex of (a) (rU·dA) and the upper box includes the duplex of (b) (rA·dT). Also shown are wiggly lines which indicate the *absence* of the A H2 to A H2 NOEs.

4.2 Å) than is the $n+1$ adenine H2 (about 5.9 Å) whereas in cross-strand NOEs the $m+1$ adenine H2 is closer (about 4 Å) to the (n) imino proton than is the $m-1$ adenine H2 (about 5.8 Å) (Chou et al., 1989). This relationship to neighboring H2 protons allowed us to assign all the imino resonances. As shown in Figure 2a, the four well-resolved imino proton peaks were easily assigned to T3, U17, T9, and G10. When the imino proton of T3 is irradiated, only one sharp H2 NOE to 22A appears at 7.6 ppm, and no NOE was found to the 21A H2. When the imino proton of U17 was irradiated, a strong NOE to 8A H2 and a small NOE to 16A H2 were found, but no NOE to 7A H2 was detected. Similarly, irradiation of the imino proton of 9T gave a strong NOE to 16A H2 and a weaker NOE to 8A H2. Thus the n imino NOEs to the $n-1$ and $m+1$ H2 resonances observed in RNA–RNA duplexes (see preceding paper), as well as in DNA–DNA duplexes with the same sequence as the present case (Wemmer et al., 1984), appear to be a general phenomenon in right-handed duplexes. We have found no exceptions to this rule in studying over 100 context effects in several duplexes, and hence, we feel reasonably confident in using this

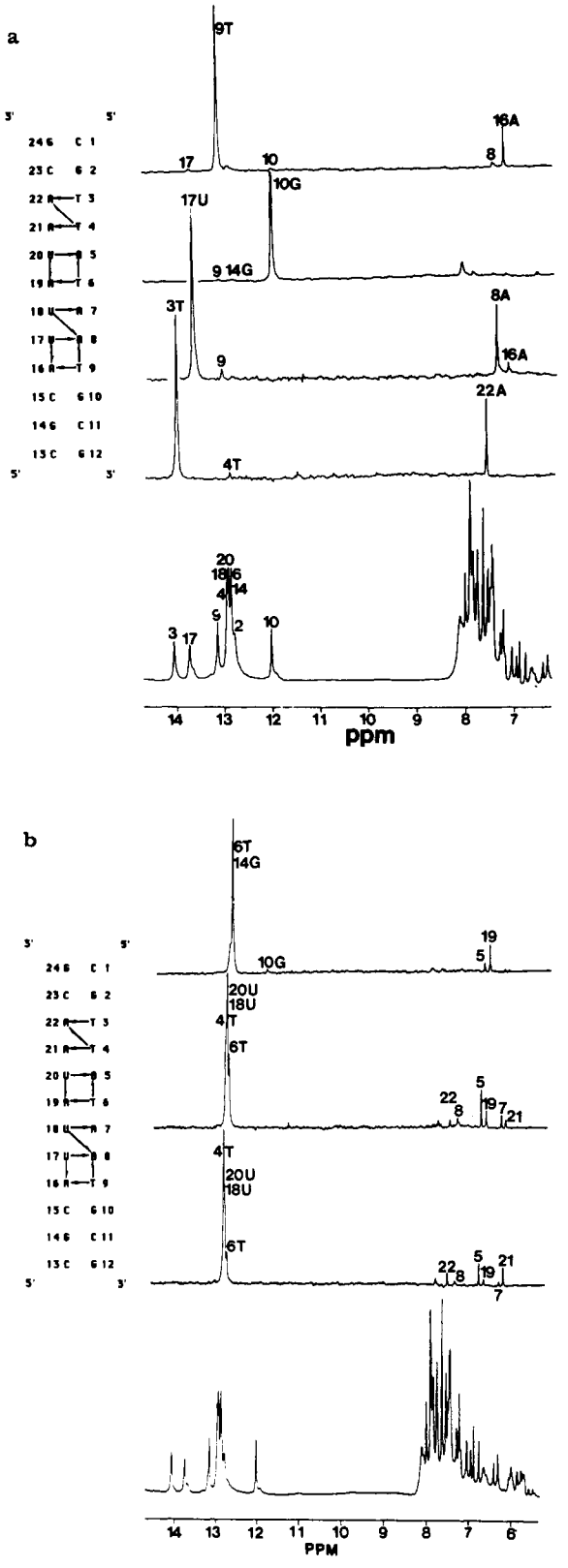

FIGURE 2: (a) 500-MHz one-dimensional imino and aromatic proton spectrum (lower) and NOEs from the four resolved imino resonances of d(CGTTATAATGCG)·r(CGCAUUAUAACG) in 20 mM sodium phosphate buffer (pH 6.8) and 200 mM sodium chloride solution at 30° C. The horizontal arrows in the caption at the left show the NOEs from the imino proton to its complementary adenine H2, and the vertical or diagonal arrows show the NOEs to the $n - 1$ and $m + 1$ neighbor H2 resonances. (b) 500-MHz one-dimensional imino and aromatic proton spectrum (lower) and the NOEs from the poorly resolved imino proton peaks of d(CGTTATAATGCG)·r(CGCAUUAUAACG) under the same condition as above.

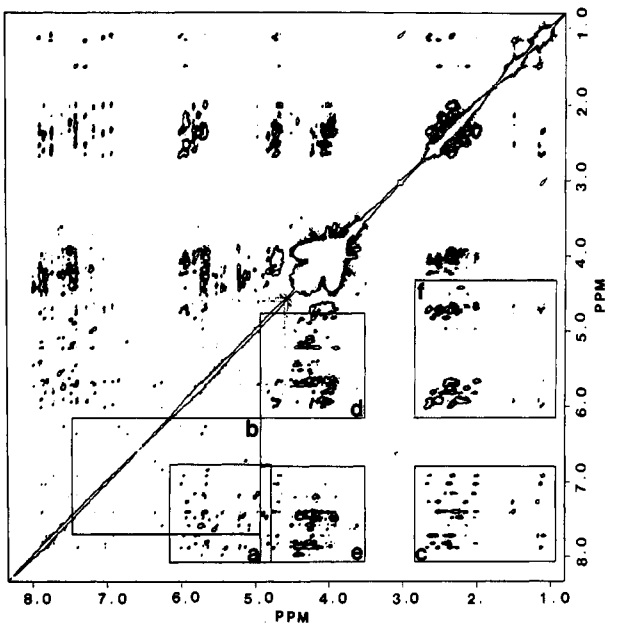

FIGURE 3: Whole NOESY spectrum of d(CGTTATAATGCG)·r(CGCAUUAUAACG). The spectrum was obtained at 30 °C with a mixing time of 200 ms. The base to H1′ region is boxed in (a) and expanded in Figure 4 for the DNA strand; the same region for the RNA strand is expanded in Figure 5; the A H2 to A H2 and the A H2 to H1′ cross-peaks are boxed in (b) and expanded in Figure 6; the aromatic to H2″, H2′, and methyl groups of the DNA strand are boxed in (c) and shown in Figure 7; the H1′ to H3′ and H1′ to H4′ of the DNA strand and the H1′ to H2′/H3′/H4′ of the RNA strand are boxed in (d) and shown in Figure 9, while the base to H3′ of the DNA strand and the base to H2′ of the RNA strand are boxed in (e) and shown in Figure 10.

adenine H2 NOE pattern to identify the six partially resolved imino protons between 12.7 and 13.0 ppm that are too overlapped to be irradiated completely selectively. The results are shown in Figure 2b, and the assignments are indicated at the bottom of Figure 2a.

NOESY Studies. The complete NOESY spectrum of the hybrid duplex is shown in Figure 3. The spectrum differs from the pure DNA–DNA Pribnow duplex NOESY spectrum in that it contains seven extra A H2 to H1′ cross-peaks in the base to H1′ region and is also quite complicated in the 3.8–4.6 ppm region due to the overlapped chemical shifts of the 2′-, 3′-, 4′-, 5″-, and 5′-protons of the RNA strand and the 4′-, 5″-, and 5′-protons of the DNA strand. However, the base to H2″ and H2′ region of the DNA strand is simplified somewhat since only 12 pairs of H2″ and H2′ occur in the 2.0–2.8 ppm region. The principal NOEs used to connect adjacent residues in DNA and RNA strands are still those between base protons and H1′ protons. Additional NOEs from the base protons to the H2″ and H2′ protons of the $n - 1$ residues are also valuable for assigning peaks in the DNA strand and determining the sugar pucker conformations. Unambiguous H2′ assignments in the RNA strand were made possible by the fairly strong intrasugar H1′ to H2′ and interresidue H2/H6/H8 to $n - 1$ H2′ NOEs.

The NOESY spectrum of the H6/H8 to H1′ and H5 region of the DNA strand is shown expanded in Figure 4. Some cross-peaks (9T, 6T, 3T, 4T, 8A, 7A, and 5A) in this region were identified as belonging to the DNA strand by means of methyl to $n - 1$ H8/H6 NOEs (see Figure 7). The NOEs belonging to the DNA strand will therefore be assigned first. Furthermore, the eight strong H5 to H6 cross-peaks of the three uridines in the RNA strand and the five cytidines were easily identified from the COSY spectrum (data not shown) and serve as a guide to correctly assigning this region. Since

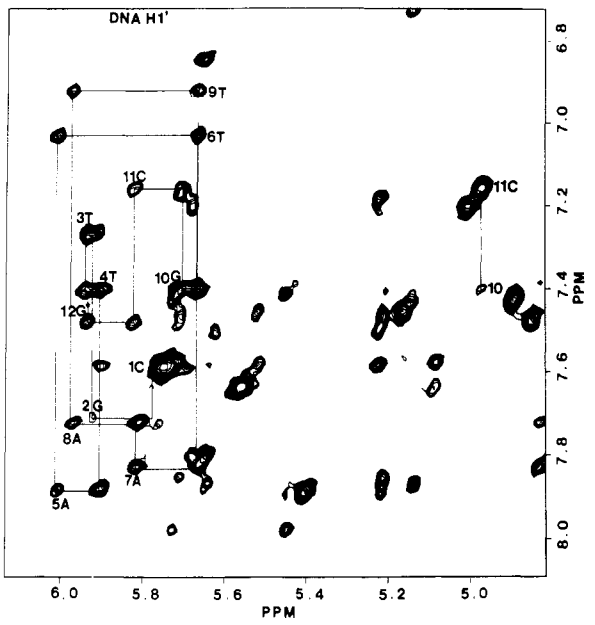

FIGURE 4: Expanded plot of the base to H1′ region of the DNA strand. Only the intraresidue H8/H6 to H1′ peaks are labeled. The NOE from 11C H5 to 10G H8 is shown as a vertical line at right. Almost equal-intensity internucleotide and intranucleotide NOEs are seen in the DNA strand.

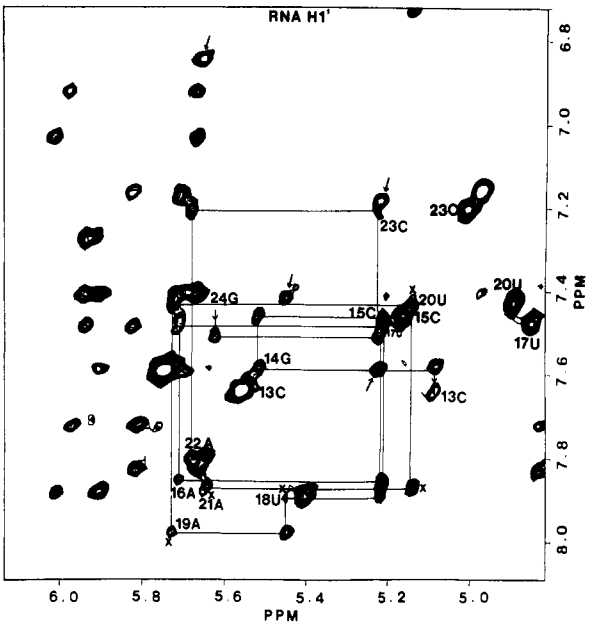

FIGURE 5: Expanded plot of the base to H1′ region of the RNA strand. The six strong H5 to H6 cross-peaks belonging to the RNA strand are labeled with large characters. The extra peaks due to the adenine H2 to $n + 1$ H1′ are indicated by arrows. As in the DNA strand, only the intraresidue aromatic to H1′ peaks are labeled. The cross-peaks in the region from 21A to 18U are further marked with an asterisk.

the methyl of the thymine in position 4 is the only methyl in the entire sequence that can generate NOEs to two thymine H6 resonances (4T H6 and 3T H6), the 4T and 3T methyls were easily identified (Figure 7), and the 2G, 3T, and 4T base protons were therefore immediately assigned with no ambiguity. These assignments constitute a convenient starting point for sequential assignment of the base and H1′ region. Apart from the identical chemical shifts of 9T H1′ and 6T H1′, and a minor problem with 10G H8 (which was assigned from the 11C H5 to 10G H8 NOE), the H8/H6 to H1′ connectivities were straightforward and are shown in Figure 4. The 1C H1′ was determined from the H1′ to H2′ cross-peak since the 2G H8 to 1C H1′ cross-peak was missing due to terminal fraying (Hare et al., 1983).

After the H8/H6 and H1′ resonances of the DNA strand were assigned, the remaining half of the cross-peaks in this region belong to the RNA strand, and their assignment is shown in Figure 5. The assignment of this region was more difficult than for the DNA strand since the ability to corroborate base assignments via methyl to H8/H6 NOEs is no longer possible. Fortunately, most of the base to H1′ cross-peaks of the RNA strand are well separated from those of the DNA strand, and the 24G and 13C terminal peaks can also be easily identified. The six remaining pyrimidine H5 to H6 cross-peaks that do not belong in the DNA strand are also useful guides. The only complications are the extra peaks introduced into this region by the new A H2 cross-peaks to the $n + 1$ and $m + 1$ H1′ protons. These are the 16A H2 to 17U H1′, 8A H2 to 18U H1′, and 22A H2 to 23C H1′ cross-peaks and are marked with arrows to set them apart from the normal peaks. Incorrectly linking these cross-peaks into the H8/H6 to H1′ network leads to rapid termination of the sequential assignment connectivities. After some trial and error, the correct linkage was established without much difficulty.

Comparison of the cross-peaks in this region showed some characteristic differences between the DNA and RNA strands. In the DNA strand, the intra- and interresidue H6/H8 to $n - 1$ H1′ NOEs are of relatively similar intensities, and no

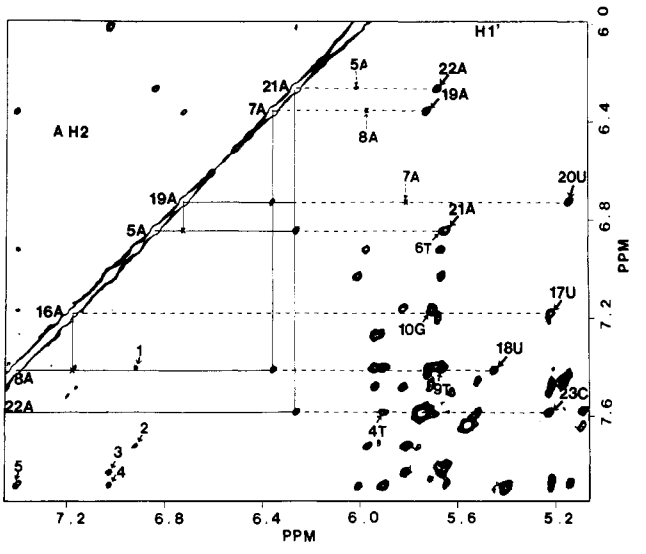

FIGURE 6: Expanded plot of the A H2 to A H2 and the A H2 to H1′ cross-peaks. Connectivities between A H2 resonances are shown as solid lines while those between A H2 and H1′ resonances are shown as dotted lines. Missing cross-peaks too weak to be detected at 200 ms are indicated by a (×). Some of the H8/H6 to H8/H6 cross-peaks are also shown and labeled with small numbers: 1, 10G–9T; 2, 9T–8A; 3, 7A–6T; 4, 6T–5A; 5, 5A–4T.

marked sequence-dependent variations were found; this same feature was also observed in the pure DNA–DNA Pribnow sequence (Wemmer et al., 1984). In the RNA strand, however, one does observe some sequence-dependent features; the NOE pattern of the 18U-19A-20U-21A sequence is similar to that of the 5U-6A-7U-8A sequence in the double-stranded RNA dodecamer r(CGCGUAUACGCG) (see preceding paper) except in the hybrid case the NOE intensities vary even more.

The A H2 to A H2 NOEs and the A H2 to $n + 1$ and $m + 1$ H1′ NOEs are shown in Figure 6. Some of the weak

base to base cross-peaks in the DNA strand are also shown and, at lower contour level, can be traced from 12G to 3T without any problem. However, no such NOEs can be detected in the corresponding RNA strand, and this indicates that the base stacking in the two strands is not the same. The seven adenine H2 resonances were assigned from A H2 to H1′ cross-peaks but were further confirmed by A H2 to A H2 NOEs and by NOEs from the imino protons to the A H2 resonances. The absence of H2–H2 NOEs between 5A and 19A and between 8A and 16A is probably simply due to the right-handed nature of the duplex. We also observed markedly inequivalent NOEs from A H2 resonances to the $n + 1$ H1′ and $m + 1$ H1′ protons in the hybrid duplex. These are especially noteworthy since in double-stranded RNA these NOEs are of almost equal intensity in the two strands due to the symmetrical nature of the base pair dyad axes (Chou et al., 1989). As shown in Figure 6 and summarized in Figure 1c, the intrastrand H2 to $n + 1$ H1′ NOEs in the DNA strand are weak or absent (e.g., 7A to 8A), but the interstrand A H2 NOES to H1′ protons on the RNA strand are very strong (e.g., 7A to 19A, 8A to 18U). Conversely, the intrastrand H2 to H1′ NOEs in the RNA strand are very strong whereas the interstrand NOEs to the H1′ protons on the DNA strand are weak to medium. Unfortunately, the only available atomic coordinates are for homopolymers, and we were unable to model the reciprocal interstrand NOEs between an adenine H2 and an adenine $m + 1$ H1′. Nevertheless we constructed two computer models corresponding to the 17U-18U/7A-8A and the 21A-22A/3T-4T regions of the Pribnow hybrid, and they are shown in structures a and b of Figure 1, respectively. These models give a rough estimate of the range of expected relative distances. As shown in Figure 1a (paths 4 and 5), the interstrand adenine H2 to $m + 1$ H1′ distance (about 4.19 Å) is shorter than the intrastrand H2 to $n + 1$ H1′ distance (about 4.45 Å) in rUUU·dAAA while the reverse is true in dAAA·dTTT where, in C3′-endo adenosines, the interstrand H2 to $m + 1$ H1′ distance (about 3.92 Å) is now longer than that of the intrastrand distance to the $n + 1$ H1′ (about 3.16 Å) (Figure 1b, paths 1 and 2). Thus, the model distances agree qualitatively with the observed NOEs in terms of relative distance information. We also find the distances in dAAA·dTTT to be qualitatively closer to the actual NOE intensities we observe. Although the present data are inadequate to be interpreted quantitatively in terms of precise distances, it is safe to say that the two strands in this hybrid are certainly in quite different conformations and the duplex does not have the normal strand equivalence of classical A or B duplexes.

The NOEs from base protons to H2′/H2″ protons and methyl groups on the DNA strand are shown in Figure 7. Compared to the same region of the pure DNA–DNA sequence (Wemmer et al., 1984), this region is now much simpler due to the shift of the complementary RNA H2′ resonances to lower field. The assignments are quite straightforward and present no problem. Since this region contains important information concerning the sugar pucker, the assignments were further confirmed from the H1′ to H2″ and H1′ to H2′ cross-peaks (data not shown). It is interesting to note that the intranucleotide and internucleotide H6/H8 to $n - 1$ H2′ NOEs are more or less equal in this hybrid, in contrast to what we observed in the DNA–DNA duplexes d(CGCGAATTCGCG) and d(CGTTATAATGCG)·d(CGCATTATAACG) (Hare et al., 1983; Wemmer et al., 1984). In the normal B-form DNA model (see Figure 2 of previous paper), the intranucleotide H6/H8 to n H2′ ranges

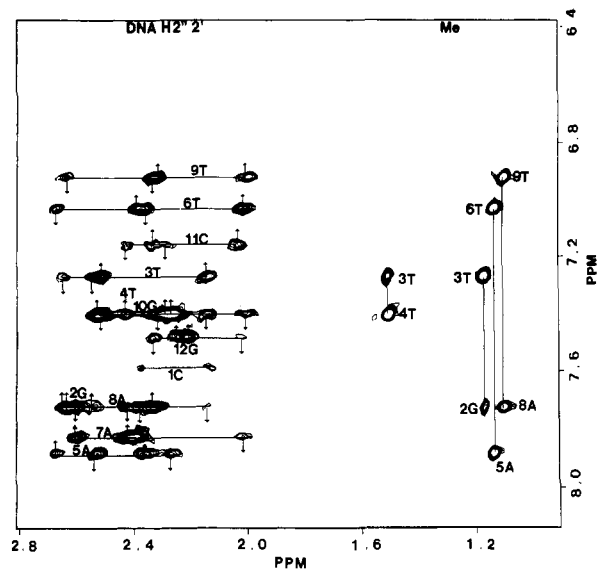

FIGURE 7: Expanded NOESY spectrum of the H8/H6 to H2″, H2′, and methyl groups of the DNA strand. The cross-peaks between the methyl group and $n - 1$ base protons are shown as vertical lines. The cross-peaks from H8/H6 to H2″ and H2′ are shown as horizontal solid lines and further labeled with an up arrow to designate intraresidue NOEs and a down arrow to designate the interresidue NOEs. Since all of the H2″ and H2′ of purines have moved upfield by about 0.2 ppm and those of the pyrimidines have moved slightly downfield, the H2″ and H2′ of the purines and pyrimidines have begun to overlap. Larger H8/H6 to $n - 1$ H2′ NOEs than those in DNA–DNA duplexes are observed, and the base NOEs to the n H2′ and to the $n - 1$ H2′ are of similar intensity.

from 1.8 to 2.1 Å while the internucleotide H8/H6 to $n - 1$ H2′ varies from 3.6 to 3.9 Å; in the A-form model, the intraresidue H6/H8 to n H2′ ranges from 3.9 to 4.2 Å while the internucleotide H8/H6 to $n - 1$ H2′ is much shorter. Thus considerable differences should be observed in the H6/H8 to $n - 1$ H2′ and n H2′ NOE intensities in B-DNA, and that is more or less what we observed for every nucleotide in those two B-form double-helical DNA sequences. The quite similar intensities of the NOEs from the H6/H8 to the $n - 1$ H2′ and to the n H2′ in most nucleotides of the DNA strand in the hybrid duplex is an important clue leading us to tentatively suggest that the sugar pucker in the DNA strand is in neither the C2′-endo nor the C3′-endo conformation, but in an intermediate form.

The COSY spectrum of the H1′ to H2″/H2′ and H6 to methyl regions of the DNA strand is shown in Figure 8. Although the 1′–2′ and 1′–2″ cross-peaks are not well resolved, there is no evidence of strong 1′–2′ coupling and weak 1′–2″ coupling as was the case in DNA–DNA duplexes, and the data suggest the coupling constants between H1′, H2″ and H1′, H2′ are roughly equal in each residue of the DNA strand in the hybrid. These combined data suggest that the DNA strand adopts neither the C2′-endo conformation nor the C3′-endo conformation. A C3′-endo configuration for the DNA strand would change the dihedral angles of H1′–C1–C2–H2′ from 150° (in C2′-endo) to 99° and decrease the 1′–2′ coupling constant to almost zero, which is not the case here.

The NOESY spectrum of the H1′ to H2′, H3′, and H4′ of the RNA strand and the H1′ to H3′ and H4′ of the DNA strand is shown in Figure 9. The DNA H3′ and H4′ cross-peaks are enclosed in boxes and were assigned from the corresponding assigned H1′. The H3′ resonances of 6T and 9T were distinguished via H6 to H3′ NOEs on the basis of their different H6 chemical shifts (see Figure 10). The H1′ to H4′ cross-peaks of 6T, 9T, and 10G in the DNA strand overlapped

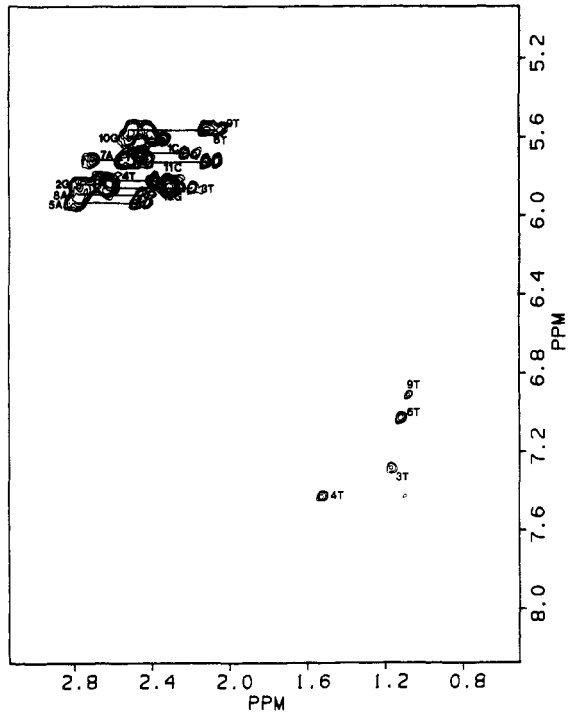

FIGURE 8: Expanded COSY spectrum of the H6 to methyl and H1′ to H2″ and H2′ region of the DNA strand. The H1′, H2″ and H2′ cross-peaks are connected by solid lines and labeled with the residue number. Four small cross-peaks corresponding to the four-bond coupling between the methyl group and the thymine H6 are shown in the lower right region.

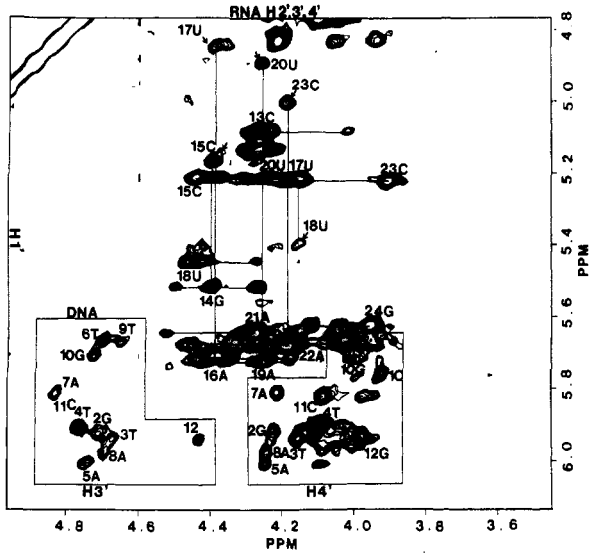

FIGURE 9: Expanded plot of the H1′ to H3′ and H1′ to H4′ of the DNA strand and the H1′ to H2′ of the RNA strand. The NOEs belonging to the DNA strand are boxed. In the RNA strand the five cross-peaks between pyrimidine H5 resonances and their $n-1$ H2′ are labeled with arrows and connected to the $n-1$ H1′ to $n-1$ H2′ cross-peak with solid lines. In the RNA region, the H1′ to H2′, H3′, H4′ are connected by a horizontal line, and only the H1′ to H2′ cross-peak is labeled.

with the RNA H1′ to H2′ peaks but were unambiguously assigned with the help of the cross-peaks from base to H4′ in the DNA strand. In the RNA strand, horizontal lines have been drawn to connect the H1′ to H2′, H3′, and H4′ cross-peaks of the same residues. The strongest peaks are the H1′ to H2′ cross-peaks since those are always the shortest distances regardless of the sugar pucker conformation (about 2.8 Å), and the H2′ chemical shifts were further confirmed from the base to H2′ NOEs. The H2′s of 13C, 20U, 18U, 14G, 24G,

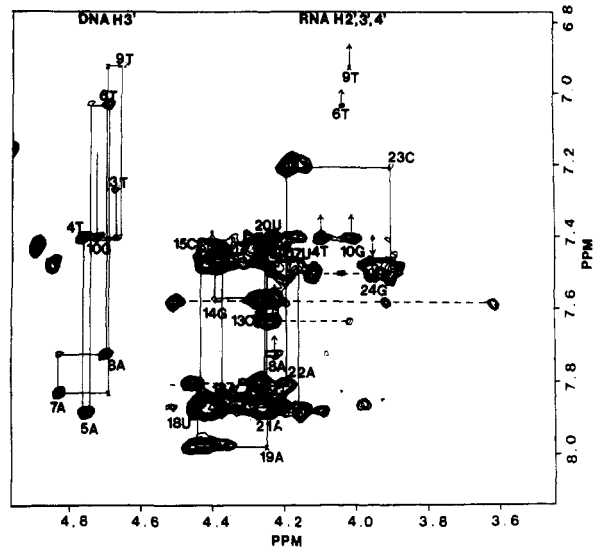

FIGURE 10: Expanded plot of the base to H3′ cross-peaks of the DNA strand and the base to H2′, H3′, and H4′ of the RNA strand. The sequential connectivities in both DNA and RNA strands are also shown. Some H6/H8 to H4′ peaks in the DNA strand are labeled as up arrows. Very weak peaks that can only be seen at lower contour levels are designated by a (×).

and 21A were well resolved and present no particular problem. The H2′s of 14G, 17U, 16A, 19A, and 22A were identified via NOEs to the $n+1$ H5 resonance (Chou et al., 1989). These five pyrimidine H5 to $n-1$ H2′ NOEs of medium intensity again indicate that the RNA chain is very probably in a A-type conformation. The H2′s of 15C and 23C were assigned indirectly from the base to $n-1$ H2′ (Figure 10). The H3′ and H4′ of the RNA strand cannot be assigned unambiguously at this time and are listed as pairs of values in Table I.

The NOESY spectrum of the base to H2′, H3′, and H4′ of the RNA strand and the base to H3′ of the DNA strand is shown in Figure 10. The sequential assignment of the DNA base to H3′ cross-peaks was derived from the base to H1′ connectivity and is connected by solid lines from 3T to 10G in the left part of the figure. The H3′ assignments were confirmed by H1′–H3′ cross-peaks. The 10G, 9T, 8A, 7A, 6T, 5A, 4T, 3T, and 2G H3′ resonances were clearly identified. In the RNA strand, some of the base to H2′, H3′, and H4′ cross-peaks of the same residue are connected with dotted lines, and the interresidue base to H2′ sequential connectivities are shown with solid lines. The assignments in this region make no assumptions about the sugar pucker conformation; they are made on the basis of base proton and independent H2′ assignments made elsewhere. The strong H1′ to H2′ peaks in Figure 9 dovetail with the internucleotide base to $(n-1)$ H2′ connectivities, and combined with the chemical shifts of the base protons, they were sufficient to establish the connectivities. It is important to note that no assumptions were made about the sugar conformation during these assignments.

Discussion

For DNA–RNA hybrids, several different combinations of conformations for the two strands have been proposed under different conditions. In fiber diffraction studies, the synthetic DNA–RNA duplexes poly(dA)·poly(rU) and poly(dI)·poly(rC) were found to adopt a hybrid structure with the DNA chains in the C2′-endo conformation and the RNA chains in the C3′-endo conformation (Arnott et al., 1986), while the poly(rA)·poly(dT) duplex was found to be able to adopt either an all C3′-endo structure or a C3′-endo structure for the RNA

Table I: Comparison of the Chemical Shifts (ppm) of Hybrid DNA–RNA Dodecamer and Its Pure DNA Analogue[a]

	1	2	3	4	5	6	7	8	9	10	11	12
	dC	G	T	T	A	T	A	A	T	G	C	G
	rG	C	A	A	U	A	U	U	A	C	G	C
	24	23	22	21	20	19	18	17	16	15	14	13

	H8/H6			H1′			H5			H2		
	B	H	Δ	B	H	Δ	B	H	Δ	B	H	Δ
1C	7.63	7.58	0.05	5.75	5.78	−0.08	5.84	5.74	0.10			
2G	**7.98**	**7.68**	**0.30**	**6.03**	**5.92**	**0.11**						
3T	7.29	7.27	0.02	6.04	5.94	0.10						
4T	7.39	7.39	0.00	5.75	5.90	−0.15						
5A	**8.33**	**7.87**	**0.46**	**6.18**	**6.00**	**0.18**				**7.19**	**6.84**	**0.35**
6T	7.13	7.04	0.09	5.49	5.67	−0.18						
7A	**8.18**	**7.82**	**0.36**	**5.94**	**5.80**	**0.14**				**6.60**	**6.36**	**0.24**
8A	**8.10**	**7.71**	**0.39**	**6.10**	**5.97**	**0.13**				**7.50**	**7.40**	**0.10**
9T	6.98	6.92	0.06	5.69	5.66	0.03						
10G	**7.81**	**7.39**	**0.42**	**5.81**	**5.70**	**0.11**						
11C	7.31	7.16	0.15	5.78	5.82	−0.04	5.37	5.58	0.27			
12G	**7.90**	**7.48**	**0.42**	**6.11**	**5.94**	**0.17**						
13C	7.58	7.65	−0.07	5.74	5.09	0.65	5.85	5.58	0.27			
14G	**7.98**	**7.39**	**0.54**	**5.88**	**5.54**	**0.34**						
15C	7.38	7.46	−0.08	5.61	5.22	0.39	5.44	6.16	0.28			
16A	**8.32**	**7.84**	**0.48**	**6.28**	**5.72**	**0.56**				**7.58**	**7.16**	**0.42**
17U(T)	7.16	7.48	−0.32	5.88	5.23	0.65		4.84				
18U(T)	7.37	7.91	−0.54	5.80	5.46	0.34		5.42				
19A	**8.29**	**7.99**	**0.30**	**6.17**	**5.73**	**0.44**				**7.08**	**6.72**	**0.36**
20U(T)	7.10	7.44	−0.34	5.47	5.15	0.32		4.89				
21A	**8.19**	**7.86**	**0.33**	**5.86**	**5.66**	**0.20**				**6.84**	**6.28**	**0.56**
22A	**8.09**	**7.78**	**0.31**	**6.04**	**5.68**	**0.36**				**7.63**	**7.56**	**0.07**
23C	7.16	7.21	−0.05	5.63	5.23	0.40	5.21	5.00	0.21			
24G	**7.83**	**7.51**	**0.32**	**6.10**	**5.63**	**0.47**						

	H2″			H2′			H3′			H4′		
	B	H	Δ	B	H	Δ	B	H	Δ	B	H	Δ
1C	2.37	2.38	−0.01	1.88	2.14	−0.26				4.06	3.92	0.14
2G	**2.84**	**2.65**	**0.19**	**2.67**	**2.55**	**0.12**	**4.97**	**4.72**	**0.25**	**4.37**	**4.22**	**0.15**
3T	2.55	2.54	0.01	2.10	2.16	−0.06	4.84	4.68	0.16	4.24	4.15	0.09
4T	2.43	2.53	−0.10	2.16	2.28	−0.12	4.84	4.76	0.08	4.16	4.10	0.06
5A	**2.91**	**2.68**	**0.23**	**2.65**	**2.36**	**0.29**	**5.02**	**4.74**	**0.28**	**4.42**	**4.25**	**0.17**
6T	2.37	2.40	−0.03	2.01	2.03	−0.02	4.84	4.70	0.14	4.14	4.06	0.08
7A	**2.91**	**2.60**	**0.31**	**2.69**	**2.40**	**0.29**	**5.03**	**4.84**	**0.19**	**4.37**	**4.23**	**0.15**
8A	**2.85**	**2.66**	**0.19**	**2.53**	**2.33**	**0.20**	**4.99**	**4.70**	**0.29**	**4.43**	**4.25**	**0.18**
9T	2.31	2.34	−0.03	1.90	2.01	−0.11	4.84	4.67	0.17	4.13	4.01	0.12
10G	**2.64**	**2.31**	**0.33**	**2.57**	**2.26**	**0.31**	**4.96**	**4.73**	**0.23**	**4.34**	**4.00**	**0.34**
11C	2.31	2.34	−0.03	1.88	2.04	−0.16	4.79	4.85	−0.06	4.18	4.08	0.10
12G	**2.57**	**2.26**	**0.31**	**2.37**	**2.22**	**0.15**		**4.43**			**3.97**	
13C					4.25			4.23, 4.02[b]				
14G					4.40			4.50, 4.26				
15C					4.43			4.39, 4.26				
16A					4.40			4.38[c]				
17U(T)					4.14			4.43, 4.24				
18U(T)					4.46			4.43, 4.28				
19A					4.26			4.42, 4.35				
20U(T)					4.26			4.31, 4.22				
21A					4.28			4.53[c]				
22A					4.19			4.46, 4.30				
23C					3.91			4.19, 4.14				
24G					3.97			4.12, 4.04				

[a] Pure DNA shifts are listed in the B column, with hybrid duplex chemical shifts in the H column and the chemical shift difference between B and H in the Δ column. [b] All the chemical shifts of the H3′ and H4′ of the RNA strand were not differentiated and listed as a pair of values. [c] Only one of the value were identified.

chain and a B-like C3′-exo structure for the DNA chain, depending on the level of hydration (Zimmerman & Pheiffer, 1981) in the fiber state. In solution, poly(rA)·poly(dT) and the small hexameric hybrid duplex d(TCACAT)·r(AUGUGA) were tentatively assigned as having both strands in the C2′-endo (B-type) conformation from one-dimensional NOE data (Gupta et al., 1985; Reid et al., 1983). The more detailed high-resolution 2D NMR studies reported indicate that in the full-turn hybrid duplex d(CGTTATAATGCG)·r(CGCAUUAUAACG) the RNA strand adopts the familiar A-type C3′-endo conformation; the DNA strand, while difinitely not in the C3′-endo conformation, is also not perfectly in the C2′ endo state, and appears to assume a C2′-endo variant structure intermediate between the two states.

The RNA Strands. (*A*) *Chemical Shifts.* The chemical shifts of the base and H1′ resonances of the RNA strand in the hybrid dodecamer show a very similar pattern to that observed in a purely RNA double helix (Chou et al., 1989). That is, the H8s of purines move upfield by 0.3–0.5 ppm while the H6s of pyrimidines move downfield by 0.1–0.5 ppm compared to the analogous DNA–DNA duplex. In contrast to the downfield H6 shifts, the H5 resonances of cytosine move upfield by about 0.2 ppm. Furthermore, all the H1′s move upfield by 0.2–0.6 ppm. These observations indicate a con-

formation for the hybrid RNA strand that is very similar to the A-type conformation found in pure RNA duplexes.

(*B*) *NOE Data.* The best evidence that the RNA strand of the hybrid duplex is in a C3′-endo sugar conformation comes from the following NOE data:

(i) The H6/H8 to $n - 1$ H1′ NOE pattern in the 18U-19A-20U-21A sequence in th hybrid is very similar to that of the UAUA sequence in the pure r(CGCGUAUACGCG) duplex.

(ii) The interresidue H6/H8 to $n - 1$ H2′ NOEs are uniformly quite strong. As we and several other laboratories have noted, this is a characteristic of the A-type C3′-endo conformation (Chou et al., 1989; Reid et al., 1983b; Clore et al., 1985; Mellema et al., 1983). Although it is not necessary to make any assumptions about the sugar pucker conformation in order to make sequential base to H2′ connectivities, it turned out that the strongest peak in each step was from the interresidue base–H2′ NOE. The weaker intraresidue base–H2′ intensities qualitatively correspond quite well with the standard A-form model. Five strong pyrimidine H5 to $n - 1$ H2′ NOE were also seen, further corroborating a C3′-endo conformation for the RNA strand. In the DNA strand of the hybrid, or in double-helical DNA, these corresponding peaks are either absent or very weak.

(iii) The strong intrastrand A H2 to $n + 1$ H1′ NOEs of every adenosine (Figures 1c and 5) indicate that the RNA strand in the hybrid adopts a similar conformation to that in a pure RNA duplex (see preceding paper).

In summary, both the chemical shift data and NOE information clearly indicate that the RNA strand is in the A-type C3′-endo conformation.

The DNA Strand. (*A*) *Chemical Shifts.* The chemical shifts of the DNA strand show a quite different pattern to those of the RNA strand. The H8 resonances of purines still move upfield by 0.3–0.45 ppm, but the pyrimidine H6 peaks do not move downfield compared to their B-DNA duplex chemical shift, as they do in the RNA strand. Instead, they also move upfield, but to a lesser extent (ranging from 0 to 0.15 ppm). Also, the H1′ shifts do not change as drastically as they do in the RNA strand, and they also become residue specific; purine H1′s move upfield by 0.1–0.18 ppm while the pyrimidine H1′s move downfield by 0.03–0.18 ppm. The purine H2″ and H2′ also move upfield by 0.12–0.31 ppm. Thus the chemical shift considerations alone tend to suggest that the DNA strand adopts a conformation that is markedly different from the RNA strand and is probably not the same as in pure DNA–DNA duplexes.

(*B*) *Coupling Constants.* The COSY spectrum of the H1′ to H2″ and H2′ region (Figure 8) indicates that the H1′ coupling to both H2″ and H2′ is quite similar in most residues. This immediately eliminates the possibility of a C3′-endo conformation for the DNA strand because the C3′-endo sugar pucker causes a 90° dihedral angle between H1′ and H2′ and would produce only a single H1′ to H2″ cross-peak in the COSY spectrum instead of two cross-peaks. The precise coupling constants are difficult to estimate without performing a *J*-resolved experiment, but it is possible that they lie in the range corresponding to O4′-endo to C1′-exo to C2′-endo (Chary et al., 1987).

(*C*) *NOE Data.* The NOE data for the DNA strand, as for the RNA strand, are perhaps the most important criteria for monitoring the sugar conformation. The 10G, 9T, 8A, 6T, 5A, 4T, and 3T base to $n - 1$ H2′ NOEs are stronger than in the pure DNA–DNA analogs (Wemmer et al., 1984) but weaker than would be expected for a pure C3′-endo conformation. The data indicate that most of the residues in the DNA strand are in neither a perfect C2′-endo nor perfect C3′-endo conformation; instead, they appear to adopt a form intermediate between these two extremes. The sugar pucker in DNA duplexes has been found to be quite variable, ranging from O1′-endo to C3′-exo in some sequences (Chary et al., 1987; Dickerson & Drew, 1981), and it is not surprising that the sugars of the DNA strand have adjusted to fit into the hybrid environment.

Combining the chemical shift, COSY, and NOESY data, it is clear that the DNA strand in the hybrid duplex is in the general C2′-endo domain and could be either O1′-endo or C1′-exo; such intermediate sugar conformations have been found in the crystal state in d(CGCGAAUUCGCG) (Dickerson & Drew, 1981) and in solution (Chary et al., 1987). With the present qualitative NOE data, we are not able to determine the precise sugar conformation at this point; such finer detail will be explored by quantitative distance measurements at a later date.

Generally speaking, our NOESY data fit the X-ray models of poly(dA)·poly(rU) and poly(dA)·poly(dT) (Arnott et al., 1986, 1983) reasonably well. The present preliminary data are insufficient to reveal the propeller -twist angles of the base pairs or the precise sugar pucker in the DNA strand. The RNA strand, is, however, very close to the classical C3′-endo conformation. The inequivalence of the RNA strand H2 to $m + 1$ H1′ NOEs and the DNA strand H2 to $m + 1$ H1′ NOEs (compare 5A to 21A versus 21A to 5A or 7A to 19A versus 19A to 7A in Figure 6) clearly indicates that the two chains are different. Since the NOESY spectrum of this hybrid is quite well resolved, it should be possible to use recently developed distance geometry methods to calculate the three-dimensional structure in solution (Hare & Reid, 1986; Reid, 1987; Patel et al., 1987). However, before such studies can begin, the H3′ and H4′ resonances of the RNA strand must be unambiguously assigned in order to generate distance constraints from these protons. Assigning the H3′ and H4′ protons only on the basis of differential NOE intensity to the H1′, as suggested by others (Clore et al., 1985), is potentially dangerous since the H1′–H3′ and H1′–H4′ distances differ by only about 0.4 Å, and thus small differential NOE intensity differences might be due to local structure irregularities as found in this hybrid. It is equally dangerous to attempt to differentiate the H2′, H3′, and H4′ resonances on the basis of chemical shifts since they all resonate between 3.9 and 4.5 ppm. We are currently synthesizing smaller oligomers with simpler NMR spectra and propose to study them by RELAY and TOCSY methods in order to assign the sugar protons unambiguously.

Finally, it is worthwhile to compare the conclusions from the present study to those from other investigations of hybrid DNA–RNA duplexes. Exclusively on the basis of the *absence* of a detectable NOE from the adenine H8 to a tentatively assigned H3′ in a poorly defined poly(rA)·poly(dT) sample of chain length 200 ± 50 base pairs, together with a great deal of unconstrained molecular model building, Sarma and coworkers generated a B-form model for poly(rA)·poly(dT) in solution that was "fully consistent with the observed NOE data" in which the rA and dT strands are conformationaly equivalent and in which both strands are in the B-form conformation (Gupta et al., 1985). While this is perhaps the first molecular structure to be determined on the basis of the lack of NOE distance constraints, the resulting structure is at variance with the results of Raman spectroscopy and fiber diffraction studies on poly(rA)·poly(dT) and should be treated

with a grain of caution; at high humidity, X-ray fiber diffraction indicates a hybrid duplex with the poly(rA) strand in the C3′-endo A conformation (Zimmermann & Phieffer, 1981), while Raman studies on poly(rA)·poly(dT) in solution indicate nonequivalent strands with a C2′-endo B-form dT strand and a C3′-endo A-form rA strand (Benevides & Thomas, 1988). Homopolymers constitute a poor choice for NMR studies because intraresidue and interresidue NOEs cannot be distinguished, and the use of a weak or absent H8 to H3′ NOE by Gupta et al. (1985) as a criterion for eliminating structures in the A-form family is dubious in light of our finding that many H8 to H3′ NOEs are rather weak for residues that are unambiguously in the A-form conformation by several other independent NOE criteria. Our results are also in conflict with the conclusions of Reid et al. (1983a); in their study of the short hybrid duplex d(TCACAT)·r(AUGUGA) they tentatively concluded that *both* strands adopt similar C2′-endo conformations. As they admit in their paper, the spectra were extensively overlapped, precluding selective irradiation of many resonances in their one-dimensional NOE approach. Furthermore, perhaps the most diagnostic conformational criterion, namely, the comparison of intranucleotide and internucleotide base to H2′ NOEs, could not be tested because they were unable to assign any of the H2′ resonances of the RNA strand. Hence, their tentative conclusion that the DNA–RNA hybrid has the same B-form conformation as the DNA–DNA duplex is based on ambiguous data. Although the discrepancy may be due to their studying a shorter half-turn duplex, we feel that their comparative studies were not sufficiently detailed to support any firm conclusions regarding the conformation of the RNA strand. Our own studies on d(CGTTATATGCG)·r(CGCATATAACG) unambigously indicate that the r(CGCATATAACG) strand is in the A-type C3′-endo conformation while the d(CGTTATAATGCG) strand is close to, but slightly different from, the C2′-endo conformation.

Acknowledgments

S.-H.C. thanks Roger Perlmutter, Head of the HHMI synthesis facility at the University of Washington, for his encouragement.

Registry No. Hybrid dodecamer duplex, 118513-72-7; DNA strand, 89144-57-0; RNA strand, 118495-50-4.

References

Arnott, S.,Chandrasekaran, R., Hall, I. H., & Puigjaner, L. C. (1983) *Nucleic Acids Res. 11*, 4141.

Arnott, S., Chandrasekaran, R., Millane, R. P., & Park, H.-S. (1986) *J. Mol. Biol. 188*, 631.

Bauman, R., Kumar, A., Ernst, R. R., & Wüthrich, K. (1981) *J. Magn. Reson. 44*, 76.

Benevides, J. M., & Thomas, G. J. (1988) *Biochemistry 27*, 3868.

Chary, K. V. R., Hosur, R. V., Govil, G., Tan, Z. K., & Miles, H. T. (1987) *Biochemistry 26*, 1315.

Chou, S.-H., Flynn, P., & Reid, B. (1989) *Biochemistry* (preceding paper in this issue).

Clore, M., Gronenborn, A. M., & McLaughlin, L. W. (1985) *Eur. J. Biochem. 151*, 153.

Dickerson, R. E., & Drew, H. R. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 7318.

Gupta, G., Sarma, M. H., & Sarma, R. H. (1985) *J. Mol. Biol. 186*, 463.

Hare, D. R., & Reid, B. R. (1986) *Biochemistry 25*, 5341.

Hare. D. R., Wemmer, D. E., Chou, S. H., Drobny, G., & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319.

Mellema, J.-R., Haasnoot, C. A. G., van der Marel, G. A., Wille, G., van Boeckel, C. A. A., van Boom, J. H., & Altona, C. (1983) *Nucleic Acids Res. 11*, 5717.

Patel, D. J., Shapiro, L., & Hare, D. (1987) *Q. Rev. Biophys. 20*, 35.

Reid, B. R. (1987) *Q. Rev. Biophys. 20*, 1.

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

Reid, D. G., Salisbury, S. A., Bellard, S., Shakked, A., & Williams, D. H. (1983b) *Biochemistry 22*, 2019.

States, D. J., Habekorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286.

Wemmer, D. E., Chou, S. H., Hare, D. R., & Reid, B. R. (1984) *Biochemistry 23*, 2262.

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