Gerald, P. S., & Efron, M. L. (1961) *Proc. Natl. Acad. Sci. U.S.A. 47*, 1758-1765.
Greer, J. (1971) *J. Mol. Biol. 59*, 107-126.
Heller, P., Coleman, R. D., & Yakulis, R. D. (1966) *J. Clin. Invest. 45*, 1021.
Hendra, P. J., & Loader, E. J. (1968) *Chem. Ind. (London)*, 718-719.
Iizuka, T., & Yonetani, T. (1972) *Methods Enzymol. 26*, 682-700.
Kitagawa, T., & Ozaki, Y. (1987) *Struct. Bonding (Berlin) 64*, 71-114.
Kitagawa, T., Kyogoku, Y., Iizuka, T., & Ikeda-Saito, M. (1976) *J. Am. Chem. Soc. 98*, 5169-5173.
Nagai, M. (1985) *Acta Haematol. Jpn. 48*, 2015-2022.
Nagai, M., & Yoneyama, Y. (1983) *J. Biol. Chem. 258*, 14379-14384.
Nagai, M., Yubisui, T., & Yoneyama, Y. (1980) *J. Biol. Chem. 255*, 4599-4602.
Nagai, K., Kagimoto, T., Hayashi, A., Taketa, F., & Kitagawa, T. (1983) *Biochemistry 22*, 1305-1311.
Nishikura, K., Sugita, S., Nagai, M., & Yoneyama, Y. (1975) *J. Biol. Chem. 250*, 6679-6685.
Ozaki, Y., Kitagawa, T., & Kyogoku, Y. (1976) *FEBS Lett. 62*, 369-372.
Ozaki, Y., Iriyama, K., Ogoshi, H., Ochiai, T., & Kitagawa, T. (1986) *J. Phys. Chem. 90*, 6105-6112.
Pulsinelli, P. D., Perutz, M. F., & Nagel, R. L. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 3870-3874.
Pyrz, J. W., Roe, A. L., Stern, L. J., & Que, L. (1985) *J. Am. Chem. Soc. 107*, 614-620.
Que, L., Jr. (1988) in *Biological Applications of Raman Spectroscopy* (Spiro, T. G., Ed.) Vol. 3, pp 491-521, Wiley, New York.
Rousseau, D. L., & Ondrias, M. R. (1983) *Annu. Rev. Biophys. Bioeng. 12*, 357-380.
Shaanan, B. (1982) *Nature 296*, 683-684.
Spiro, T. G. (1983) in *Iron Porphyrins* (Lever, A. B. P., & Gray, H., Eds.) Vol. 2, pp 89-159, Addison-Wesley, Reading, MA.
Spiro, T. G., Ed. (1988) in *Biological Applications of Raman Spectroscopy*, Vol. 3, Wiley, New York.
Spiro, T. G., & Burke, J. M. (1976) *J. Am. Chem. Soc. 98*, 5482-5489.
Spiro, T. G., Stong, J. D., & Stein, P. (1979) *J. Am. Chem. Soc. 101*, 2648-2655.
Steadman, J. H., Yates, A., & Huehns, E. R. (1970) *Br. J. Haematol. 18*, 435-446.
Teraoka, J., & Kitagawa, T. (1980) *J. Phys. Chem. 84*, 1928-1935.
Tomimatsu, Y., Kint, S., & Scherer, J. R. (1976) *Biochemistry 15*, 4918-4924.
Yu, N. T. (1986) *Methods Enzymol. 130*, 350-409.

Solid-Phase Synthesis and High-Resolution NMR Studies of Two Synthetic Double-Helical RNA Dodecamers: r(CGCGAAUUCGCG) and r(CGCGUAUACGCG)†

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: Ten-micromole solid-phase RNA synthesis has been successfully performed on an automated nucleic acid synthesizer with coupling efficiencies up to 99%, using the *tert*-butyldimethylsilyl group to protect the 2′-hydroxyl. The *tert*-butyldimethylsilyl group was easily removed by tetrabutylammonium fluoride under conditions in which virtually no 2′- to 3′-isomerization was found to occur. By use of this approach, the self-complementary RNA dodecamers r(CGCGAAUUCGCG) and r(CGCGUAUACGCG) were synthesized on an automated nucleic acid synthesizer, purified by TLC, and studied by high-resolution NMR. Imino protons were assigned from one-dimensional nuclear Overhauser effects. The nonexchangeable base, H1′, and H2′ protons were assigned by the sequential NOESY connectivity method. The NOE data from these two oligomers were analyzed qualitatively and compared to the ideal A- and B-type helix models of Arnott et al. (1972a,b). The internucleotide H6/H8 NOEs to the preceding H1′ in r(CGCGUAUACGCG) were found to be sequence-dependent and probably reflect the roll angles between adjacent bases. The internucleotide H6/H8 to H2′ NOEs of these oligomers correspond very well to an A-type conformation, but the interstrand adenine H2 NOEs to the following H1′ were much stronger than those predicted from the fiber model. These srong interstrand NOEs can be rationalized by base pair slide to favor more interstrand base overlap, as predicted by Callidine and Drew (1984).

The number of structural studies on synthetic DNA fragments has increased markedly in the last few years, by both X-ray crystal analysis (Dickerson et al., 1982; Wang et al., 1981; Shakked et al., 1983; Nelson et al., 1987) and 2D NMR methods (Hare et al., 1983; Scheek et al., 1983; Feigon et al., 1983; Wemmer et al., 1984a,b; Nilsson et al., 1986; Nilges et al., 1987; Kintanar et al., 1987). Single-crystal X-ray diffraction studies have revealed A-type, B-type, and left-handed Z-type gross DNA conformations as well as local structure variations induced by particular base sequences (Dickerson, 1983). At a lower level of detail, X-ray fiber

† These preliminary studies were supported by NIH Grant PO1 GM32681.
‡ Howard Hughes Medical Institute.
§ Biochemistry Department.
‖ Chemistry Department.

0006-2960/89/0428-2422$01.50/0

diffraction has led to a gross description of the overall polymer structure (Millane et al., 1984). The solution structure of medium-size nucleic acid duplexes could not be studied by high-resolution NMR until 2D NMR methods for the sequential assignment of nonexchangeable protons were developed (Hare et al., 1983; Scheek et al., 1983, 1984; Feigon et al., 1983; Wemmer et al., 1984a). By use of 2D NMR, almost all the nonexchangeable protons of DNA oligomers up to 34 nucleotides can now be assigned from a single NOESY[1] spectrum and a single COSY spectrum (Wemmer et al., 1984b). Moreover, the NOE data at different mixing times can be used to calculate reasonably accurate proton–proton distances which can then be used to elucidate the three-dimensional structure in solution. By means of molecular dynamics refinement of idealized duplexes, the structure of a B-type DNA has been elucidated (Nilsson et al., 1986; Nilges et al., 1987), and distance geometry methods have been used to determine the structure of a hairpin (Hare & Reid, 1986), a duplex with a base pair mismatch (Hare et al., 1986a), and a duplex with an extrahelical residue (Hare et al., 1986b). Two reviews on 2D NMR and its applications in nucleic acid structure determination have been published recently (Reid, 1987; Patel et al., 1987).

Similar studies on RNA fragments are potentially even more interesting than these DNA studies since RNA molecules have a greater tendency to adopt secondary and tertiary structures such as those involved in tRNA folding. However, very few 2D NMR studies on RNA molecules have been reported due to the difficulty of synthesizing pure RNA samples in large amounts. A great deal of effort is currently being directed toward the development of reliable methods for solid-phase synthesis of oligoribonucleotides that approach the efficiency of those already in routine use for DNA synthesis, especially the solid-phase phosphite triester method (Beaucage & Caruthers, 1981). The key to success in these endeavors lies in finding suitable, and mutually compatible, protecting groups for both the 2′- and 5′-hydroxy groups. The 2′-protecting group must remain stable throughout the phosphoramidite preparation and the repeated cycles of chain assembly, yet be specifically and completely removed at the end of the synthesis. However, for synthesis in the 3′- to 5′-direction, the 5′-protecting groups must be completely removed at every cycle of nucleotide addition, without affecting the 2′-protecting group.

Of the several 2′-protecting groups that have been proposed (Kempe et al., 1982; Tanaka et al., 1986; Reese, 1985; Ogilvie & Entwistle, 1981), the acid-labile tetrahydropyranyl group has emerged as the most popular. Several small RNA oligomers (six to eight residues) have been manually synthesized with this protecting group (Kierzek et al., 1986) even though the overall yields were less than 50% due to step yields of only ca. 90%. In agreement with others (Christodoulou, 1986; Garegg et al., 1986; Seliger et al., 1986), we have confirmed that it is rather difficult to obtain the required high selectivity of deprotecting the dimethoxytrityl group from the 5′-position without also partially removing the tetrahydropyranyl group from the 2′-position. HPLC analysis of the reaction showed many more peaks than expected, and it is apparent that considerable loss of Thp groups takes place during repeated acidic treatments in several cycles of synthesis, leading to 2′- to 3′-isomerization and phosphodiester cleavage. In fact, the original developers of the Thp group for 2′-protection have acknowledged that extensive isomerization and cleavage occur if stringently anhydrous conditions are not used (Reese & Skone, 1985), and they have suggested that Mthp or Thp groups be replaced by a 2′-protecting group that is fully stable under the protic acid conditions required for the rapid and complete removal of the 5′-DMT group.

The use of the *tert*-butyldimethylsilyl group to protect the 2′-hydroxy (Sung & Narang, 1982; Usman et al., 1985; Garegg et al., 1986) in solid-phase RNA synthesis has been less popular than the use of Mthp of Thp groups due to conflicting data regarding isomerization with the TBDMS group (Jones & Reese, 1979; Ogilvie & Entwistle, 1981). However, the TBDMS group has been successfully used to synthesize RNA oligomers up to 19 residues (Ogilvie et al., 1984). Since 2D NMR experiments require large samples of high purity and since we were unable to obtain satisfactory yields using Thp to protect the 2′-hydroxyl, we decided to reinvestigate TBDMS as a 2′-protecting group in solid-phase RNA synthesis using ribonucleoside 3′-phosphoramidites. As part of our continuing studies on the structure and dynamics of oligonucleotides in solution, we have now carried out 500-MHz NMR studies on the double-stranded RNA dodecamer duplexes CGCGAAUUCGCG and CGCGUAUACGCG synthesized with an automated nucleic acid synthesizer and solid-phase phosphoramidite methods. The 2′-TBDMS protection group was found to be stable throughout the entire 11-cycle synthesis and gave very high coupling yields (~98–99%). Furthermore, virtually no isomerization or chain cleavage could be detected during the synthesis procedures. The purity and integrity of the synthesized RNA dodecamers were analyzed by NMR and found to be excellent. This material was eminently suitable for the detailed RNA structural studies which we now report.

Materials and Methods

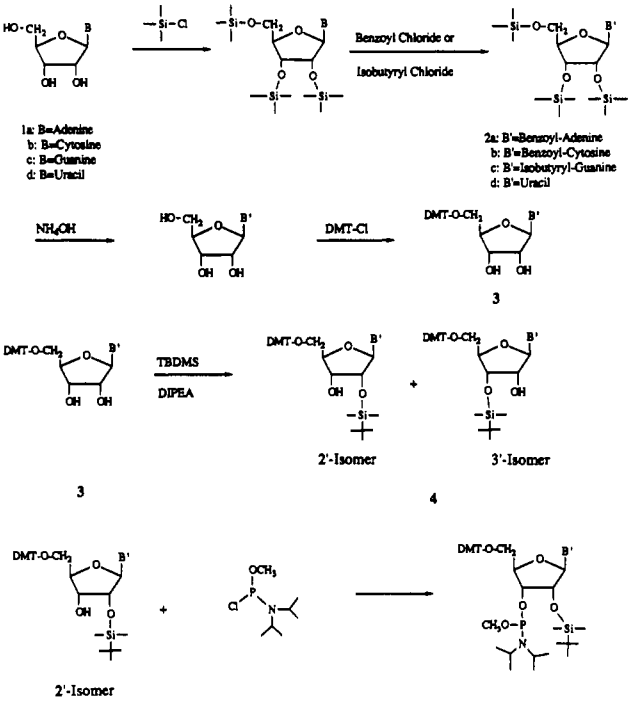
Reagents and Solvents. Silica gel for flash chromatography was purchased from Baker (~40-μm average particle diameter). Thin-layer chromatography (TLC) was performed on EM plastic-backed sheets (silica gel 60 F_{254}, 0.2 mm) in dichloromethane–methanol (9:1) (solvent A) or dichloromethane–acetone (11.5:1) (solvent B). Dowex 50W-X2 sodium form was prepared by neutralizing 200-mesh hydrogen form (Bio-Rad) with 1 N sodium hydroxide solution. Dichloromethane was purchased from EM and used as such without further drying. Adenosine, guanosine, cytidine, uridine, and RNase T2 were purchased from Sigma (St. Louis, MO). Dimethoxytrityl chloride, *tert*-butyldimethylsilyl chloride, tetrabutylammonium fluoride, trimethylchlorosilane, imidazole, benzoyl chloride, and isobutyryl chloride were purchased from Aldrich (Milwaukee, WI). Chloro(*N*,*N*-diisopropylamino)methoxyphosphine was purchased from American Biometics (Emeryville, CA).

The solid-phase support was synthesized following the procedure of Matteucci and Caruthers (1980), and its loading was measured with the trityl ion assay and found to be around 18–21 μmol/g.

Preparation of 5′-Dimethoxytrityl-2′-tert-butyldimethylsilyl N-Blocked Nucleoside Phosphoramidite. The steps in preparing the protected ribonucleoside phosphoramidites are shown in Scheme I. N^6-Benzoyladenosine, N^4-benzoylcytidine, N^2-isobutyrylguanosine, and N-blocked 5′-DMT nucleosides were prepared according to the procedure of Ti et al. (1982), except the amount of trimethylchlorosilane was increased to

[1] Abbreviations: COSY, two-dimensional autocorrelated spectroscopy; DIPEA, diisopropylethylamine; DMT, dimethoxytrityl; HPLC, high-pressure liquid chromatography; Mthp, methoxytetrahydropyranyl; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser effect spectroscopy; TBDMS, *tert*-butyldimethylsilyl; TEAA, tetraethylammonium acetate; Thp, tetrahydropyranyl; TLC, thin-layer chromatography.

Scheme I: Synthetic Pathway for the Preparation of 5′-DMT-2′-TBDMS N-Protected Ribonucleoside Phosphoramidite



a 7-fold molar excess. N-Blocked nucleosides were dried thoroughly under vacuum before being reacted with dimethoxytrityl chloride. 5′-DMT-2′-*tert*-butyldimethylsilyl N-blocked ribonucleosides were prepared and purified according to Sung et al. (1982). The final phosphoramidites were prepared by the following procedure: **4a–d** (the 2′-isomer, 3 mmol) were dissolved in 10 mL of dichloromethane (EM Omnisolv grade) and diisopropylethylamine (Applied Biosystems) (3 equiv, 1.58 mL) in a 10-mL reaction vessel preflushed with argon. Chloro(*N*,*N*-diisopropylamino)methoxyphosphine was added dropwise by syringe to the reaction vessel under argon at room temperature. The reaction was usually complete within 1 h. The solution was transferred with 35 mL of dichloromethane into a 100-mL separatory funnel and then extracted with 30 mL of an aqueous, saturated solution of sodium bicarbonate. The aqueous layer was further extracted with another portion of dichloromethane (30 mL), and the organic layer was combined and washed several times with a saturated solution of sodium bicarbonate. The final organic solution was dried over anhydrous sodium sulfate and evaporated to a foam under reduced pressure. The foam was dissolved in 5 mL of dichloromethane and purified by flash chromatography on a silica gel column. About 40 fractions of 15 mL each were collected, and each was monitored by TLC using solvent system B. The combined fractions were evaporated to produce a pure white powder which can be kept in a vacuum desiccator for over 6 months without serious decomposition. Isolated yields for the four phosphoramidites were around 75–85%.

RNA Synthesis. Ten-micromole RNA synthesis was carried out on an automated DNA synthesizer (Model 380B, Applied Biosystems, CA) using minor modifications of the recommended synthesis cycles. In a typical 10-μmol synthesis, the solid-phase support was first treated with 3% TCA in dichloromethane for 180 s to completely remove the 5′-DMT protection group. After several washes with acetonitrile, the appropriate nucleoside phosphoramidite solution (0.15 M) was premixed with activator (tetrazole) and then sent to the reaction column. The mole ratio of phosphoramidite to solid-phase support was 15 to 1. The coupling reaction was allowed to take place for 10 min, and then excess reagents were flushed out of the column. Capping reagent (acetic anhydride) was then sent to the column for 45 s and allowed to react for 180 s. The capping time was increased to 180 s to convert the guanine modification that can take place in the phosphite triester approach (Pon et al., 1986). Oxidation was then carried out for 30 s. The cycle was then repeated for the desired number of times. Each 10-μmol cycle takes about 50 min to complete and has a routine step yield of about 98–99%, on the basis of spectrophotometric determination of the released dimethoxytrityl cation.

Deprotection. After 30 min of thiophenol/triethylamine/dioxane (1:2:2) treatment to remove the methyl group from the phosphate, the partially protected RNA oligomer was eluted from the controlled pore glass resin with concentrated NH_4OH. Ethanol (1/3 volume) was added to the vial, which was then heated at 55 °C overnight. The solution was evaporated under reduced pressure to dryness and treated with 0.1 M TBAF for 4 h to remove the *tert*-butyldimethylsilyl group on the 2′-position. The extra TBAF was removed by exchanging with Dowex 50W-X2 (sodium form) and passing through Sephadex.

Purification. Oligonucleotides were purified on multiple TLC plates (Chou et al., 1983). After development in ammonia/*n*-propanol/water (35:55:10 by volume) for 4 h, the least mobile band was cut out and eluted twice with 1 mL of distilled water. Low but detectable levels of smaller oligomers (failure sequences) were observed under an ultraviolet lamp as bands preceding the predominant dodecamer. The duplex was further applied to a hydroxyapatite column and eluted with a gradient from 10 mM to 300 mM phosphate buffer (pH 7) to remove impurities leached from the TLC plates. The RNA dodecamers were then concentrated by evaporation, desalted on a G-15 column eluted with water, and lyophilized to dryness.

NMR Methods. All NMR experiments in this study were recorded on a Bruker WM-500 NMR spectrometer at ~30 °C. One-dimensional exchangeable proton spectra were acquired into 4096 complex points with a Redfield solvent-suppression pulse sequence. A spectral width of ~12 kHz was employed, with the carrier frequency set at ~12 ppm, and a total pulse length of 275 μs. NOE difference spectra were collected by coaveraging 16 scans with the saturating field directed off-resonance and subtracting an equal number of scans with the saturating field on-resonance.

COSY spectra were collected in the absolute-magnitude mode with 1024 complex points in t_2 and 400 points in t_1. NOESY spectra were recorded in the hypercomplex phase-sensitive mode (States et al., 1982) with 1024 complex t_2 points and 400 complex pairs in t_1.

The two-dimensional spectra were processed with software provided by Dennis Hare (Hare Research, Woodinville, WA). The COSY data were apodized with a sine-bell function. NOESY data were processed with a skewed phase-shifted sine-bell function. Noise ridges present in the 2D spectra were attenuated by multiplying the first row by 0.5 prior to transformation in t_1. Both COSY and NOESY spectra were subjected to diagonal low-point symmetrization (Baumann et al., 1981).

Results

Before 10-μmol syntheses of RNA dodecamers were tried, several preliminary 1-μmol syntheses of shorter sequences were first carried out and monitored by HPLC to determine the efficiency of the procedure. The results of the reverse-phase

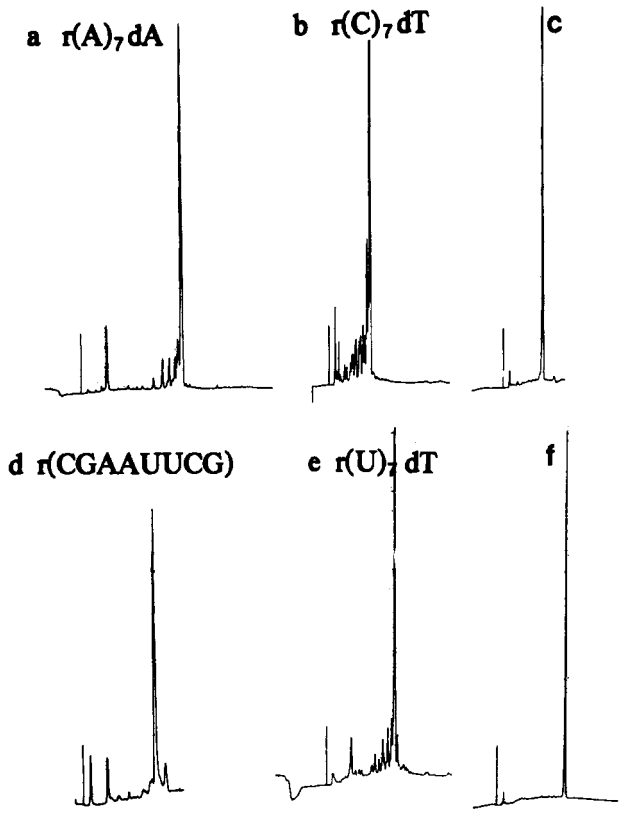


FIGURE 1: Analytical HPLC tracing of synthetic RNA oligomers: (a) crude $r(A)_7dA$, (b) crude $r(C)_6dT$, (c) purified $r(C)_6dT$, (d) self-complementary r(CGAAUUCG), (e) crude $r(U)_7dT$, and (f) purified $r(U)_7dT$. Solvent systems used were as follows: (A) 0.1 M TEAA in H_2O; (B) acetonitrile. Reverse-phase chromatography (C18 column) was carried out with a linear gradient from 98% A/2% B to 88% A/12% B in 30 min. The flow rate was 2 mL/min. In profiles c and f, the major fraction from runs b and e was collected and reinjected into the column. All the oligomers were also checked with anion-exchange chromatography and found to contain one major peak.

chromatography of single-stranded $r(A)_7dA$, $r(C_6dT$, $r(U)_7dT$, and self-complementary r(CGAAUUCG) are shown in Figure 1. The full-length $r(C)_6dT$ and $r(U)_7dT$ peaks were collected and reinjected into the HPLC column to assess their purity. The four crude RNA oligomers were also chromatographed on an anion-exchange column and found to contain one major component. The major peaks collected from reverse-phase column chromatography were further treated with RNase T2 to test for the presence of 2′–5′ phosphodiesters (migration); all were totally digested within 1 h. The self-complementary r(CGAAUUCG) was also purified by semipreparative C-18 column chromatography, and the major peak was checked by 500-MHz NMR to test its purity. However, the double-stranded r(CGCGAAUUCGCG) and r(CGCGUA-UACGCG) dodecamers could not be purified by C-18 column chromatography; the material eluted as broad and distorted peaks in either reverse-phase or anion-exchange chromatography, probably due to aggregation. These two RNA dodecamers were therefore purified on TLC plates; one major dark band was seen under 254-nm UV light in each case. About 12 TLC plates were used to purify the crude mixture from a 10-μmol synthesis; the major dark band was scraped from the plates and extracted with water several times (Chou et al., 1983). About 20 mg of pure RNA dodecamer was obtained from each 10-μmol synthesis.

Before discussing the experimental data, it is worthwhile to consider the idealized structure A-form RNA and B-form DNA and the qualitative NOE patterns expected from such structures. Double-helical segments of regular A-form RNA (UAU) and B-form DNA (TAT) were generated from the data of Arnott et al. (1972a,b) and are shown in Figure 2. Proton coordinates were generated by assuming ideal sp2 or sp3 bond angles and C–H bond lengths. Although some proton pairs are a little too close, these are reasonable models to start with.

The one-dimensional imino proton spectrum of r-(CGCGAAUUCGCG) in H_2O at 24 °C and the sequential NOE assignments of the imino proton resonances (Chou et al., 1983) are shown in Figure 3a. Five of the six imino protons are clearly seen at 24 °C. As is the case in DNA oligomers, the terminal base pair undergoes rapid exchange with bulk H_2O and cannot be observed at this temperature (Chou et al., 1983). Since the five imino protons are very well separated, their assignments by sequential NOE are straightforward and are shown in the upper part of the figure. The two resonances between 13.5 and 14.0 ppm were easily identified as AU imino protons from their strong NOEs to adenine H2 resonances between 7.0 and 8.0 ppm. When the lower field AU imino proton at 13.9 ppm was irradiated, only one NOE to a nearest-neighbor imino proton at 13.6 ppm was seen, indicating that the irradiated imino proton was from the base pair 6 position at the 2-fold symmetry axis. The imino protons were then irradiated consecutively to sequentially assign base pairs 5, 4, 3, and 2 (data not shown for base pairs 2 and 3). The corresponding data for r(CGCGUAUACGCG) are shown in Figure 3b together with the imino assignments. It is worth noting in Figure 3 that the NOEs from the imino protons to the adenine H2 of the adjacent base pair are markedly different and nonsymmetrical. This nonequivalent imino proton NOE behavior is consistent with the structures proposed for ideal A-form, B-form, or hybrid duplexes [see Figure 2 of this paper and Figure 1a,b of the accompanying paper (Chou et al., 1989)]. In a right-handed helical nucleic acid, a 5′-neighboring $n - 1$ adenine H2 on the same strand is always closer to the imino proton being irradiated (about 4.2 Å) than is a 3′-neighboring $n + 1$ adenine H2 on the same strand (about 5.9 Å), but the $m + 1$ adenine H2 on the opposite strand is closer (about 4 Å) than the $m - 1$ adenine H2 on the opposite strand (about 5.8 Å). Although the actual distance may well vary depending on the local conformation and sequence, we find this general rule derived from idealized coordinates to be a useful one for right-handed helices, where the adenine H2 is on the minor groove side. The rule can be useful in helping assign crowded imino proton spectra if the sequence contains AU or AT base pairs (see accompanying paper). The nonsymmetrical NOEs in Figure 3a can be interpreted as follows: in spectrum B of Figure 3a, the strong NOE at 7.7 ppm consists of NOEs to two adenine H2 resonances with the same chemical shift (the m and $n - 1$ H2 peaks), and the NOE from the 7U imino proton to the $m - 1$ A5 H2 (7.01 ppm) is very weak; however, in spectrum C of Figure 3a, considerable NOE intensity from the 8U imino proton to the $m + 1$ A6 H2 is also observed in addition to the NOE from the 8U imino proton to the m A5 H2. In Figure 3b, only weak NOEs from the uridine imino protons to the $n + 1$ adenosine H2 were observed.

The one-dimensional nonexchangeable proton spectrum of r(CGCGAAUUCGCG) in D_2O is shown in Figure 4. From the spectrum, it is obvious that no proton resonances occur upfield of ca. 4 ppm; i.e., the H2′ protons have moved downfield due to the electron-withdrawing effect of the 2′-hydroxyl group that replaces the H2″ proton in DNA, and furthermore, RNA contains no methyl groups. H2′, H3′, H4′, H5′, and H5″ have all moved into a crowded region between

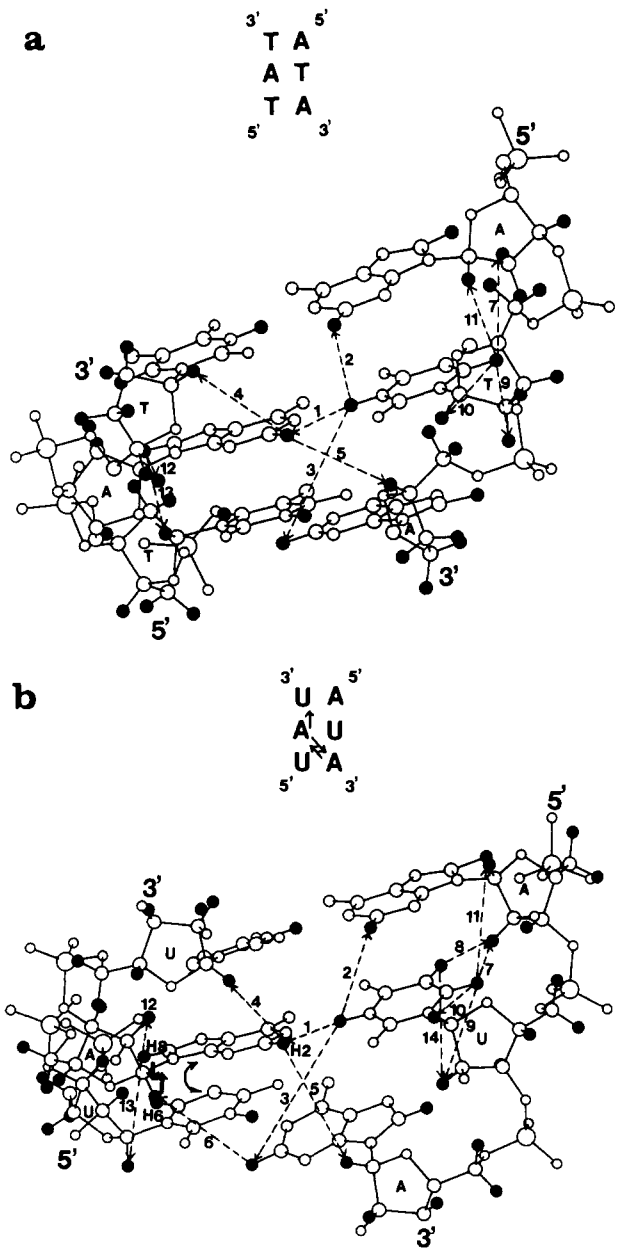


FIGURE 2: (a) Computer drawing of an idealized DNA duplex structure with the sequence 5'TAT3'. The coordinates were taken from Arnott et al. (1972a). Some hydrogen atoms relevant to NMR studies are shown as solid black circles. The proton pairs emphasized in the text are connected by dotted lines and are separated by the following distances: (1) imino proton of T to adenine H2 of the same base pair, 2.81 Å; (2) imino proton of T to the $n-1$ intrastrand adenine H2, 3.6 Å; (3) imino proton of T to the $n+1$ intrastrand adenine H2, 4.88 Å; (4) adenine H2 to the $n+1$ intrastrand H1', 4.65 Å; (5) adenine H2 n to the interstrand $m+1$ H1', 6.24 Å; (7) T H6 to the intrastrand $n-1$ H2', 3.91 Å; (9) T H6 to its own H2", 3.46 Å; (10) T H6 to its own H1', 3.72 Å; (11) T H6 to the $n-1$ H1', 3.5 Å; (12) adenine H8 to its own H1', 4.0 Å; (13) adenine H8 to the intrastrand $n-1$ H1', 3.6 Å. (b) Computer drawing of an idealized RNA duplex structure with the sequence 5'UAU3'. The coordinates were taken from Arnott et al. (1972b). Certain proton pairs are connected by dotted lines and have the following distances: (1) imino proton of U to adenine H2 of the same base pair, 2.87 Å; (2) imino proton of U to the intrastrand $n-1$ adenine H2, 4.2 Å; (3) imino proton of U to the intrastrand $n+1$ adenine H2, 4.66 Å; (4) adenine H2 to the intrastrand $n+1$ H1', 3.9 Å; (5) A H2 to the $m+1$ interstrand H1', 4.66 Å; (6) A H2 to the interstrand $m+1$ H1', 4.66 Å; (7) U H6 to the intrastrand $n-1$ H2', 1.57 Å; (8) U H5 to the intrastrand $n-1$ H2', 2.98 Å; (9) U H6 to its own H2', 3.93 Å; (10) U H6 to its own H1', 3.56 Å; (11) U H6 to the intrastrand $n-1$ H1', 4.2 Å; (12) A H8 to its own H1', 3.7 Å; (13) H8 to intrastrand $n-1$ H1', 4.0 Å; (14) U H5 to its own H2', 4.3 Å.

4 and 5 ppm, making them more difficult to assign. Another important difference worth mentioning compared to DNA spectra is the sharp singlest for the H1' resonances in RNA; this is in contrast to the H1' peaks in DNA oligomers which are split into multiplets and suggests that the sugar pucker has changed to a different conformation. The H1' singlet in RNA is only possible when the dihedral angle between the H1' and H2' is close to 90°, which is typical of an A-type or 3'-endo (N) sugar conformation.

A contour plot of the NOESY spectrum from 3.8 to 8.2 ppm of r(CGCGAAUUCGCG) at 100-ms mixing time is presented in Figure 5. The base H6/H8 to H2', H3', H4', H5", and H5' regions and the H1' to H2', H3', H4, H5", and H5' regions have become crowded and overlapped due to the similar chemical shifts of the H2', H3', H4', H5", and H5' protons (boxed regions b and c). This is unfortunate and makes the assignments of H2', H3', and H4' resonances more difficult. The base H6/H8 to H1' region (boxed region a) is shown expanded in Figure 6. At first glance the base to H1' connectivities look very similar to those for DNA, except for two strong cross-peaks at around 7 ppm and the absence of NOEs from cystosine H5 to the $n-1$ H8/H6. Also, in addition to the four cytidines, the two uridines generate H5–H6 cross-peaks, resulting in six very strong H5 to H6 cross-peaks which are easily identified. At longer mixing time (300 ms), weak NOEs from the three internal cytidine H5 to their preceding H1' were observed (data not shown), which are very helpful in assigning the H1' and H6/H8 protons. Since thymidine has been replaced by uridine in RNA molecules, one cannot use the interresidue methyl group to H6/H8 NOE as in DNA, and we have to assign all base and H1' resonances from this region only. In idealized double-helical RNA models, the base H8/H6 protons are virtually equidistant (~3.8 Å) from the H1' of their own sugar and the H1' of the preceding nucleotide, regardless of the sugar conformation (see Figure 2). In such a structure, two equal NOE spots should be detected for every H6/H8 except the 5'-terminal one, which has only a single NOE to its own H1'. Conversely, each H1' except the last residue at the 3'-end, should have two H6/H8 cross-peaks. In practice, we do find two cross-peaks for every H1' except the single spot at about 5.7 ppm, which was accordingly assigned to the H1' of 12G. The sequential assignments were then performed according to the helical nucleic acid connectivity method developed by several laboratories (Hare et al., 1983; Scheek et al., 1983; Feigon et al., 1983; Wemmer et al., 1984a). The strong H5 to H6 NOEs of all six pyrimidines and the weak H5 to $n-1$ H8/H6 NOEs in this region were used together with the sequence information to determine the assignments. The only anomaly compared to DNA asignments is the additional strong peaks from interstrand and intrastrand NOEs between H1' and adenine H2 resonances (see the cross-peaks with arrows in Figure 6). These cross-peaks are very useful for double checking the H6/H8 to H1' assignments, and we will discuss their structural implications later.

The expanded H1'/H5 to H2'/H3'/H4' NOESY region is shown in Figure 7. Every H1' should exhibit cross-peaks to its own H2'/H3'/H4' in this region, and these sets of three cross-peaks are connected by a horizontal line for each residue. Also, additional strong cross-peaks corresponding to NOEs between a pyrimidine H5 and its own H2' or the preceding H2' were detected and are indicated by arrows in the figure. These NOEs in the corresponding DNA sequence are very weak (Hare et al., 1983) and suggest that the sugars in the RNA dodecamer are adopting a C3'-endo conformation instead of a C2'-endo conformation (the H5 to preceding H2'

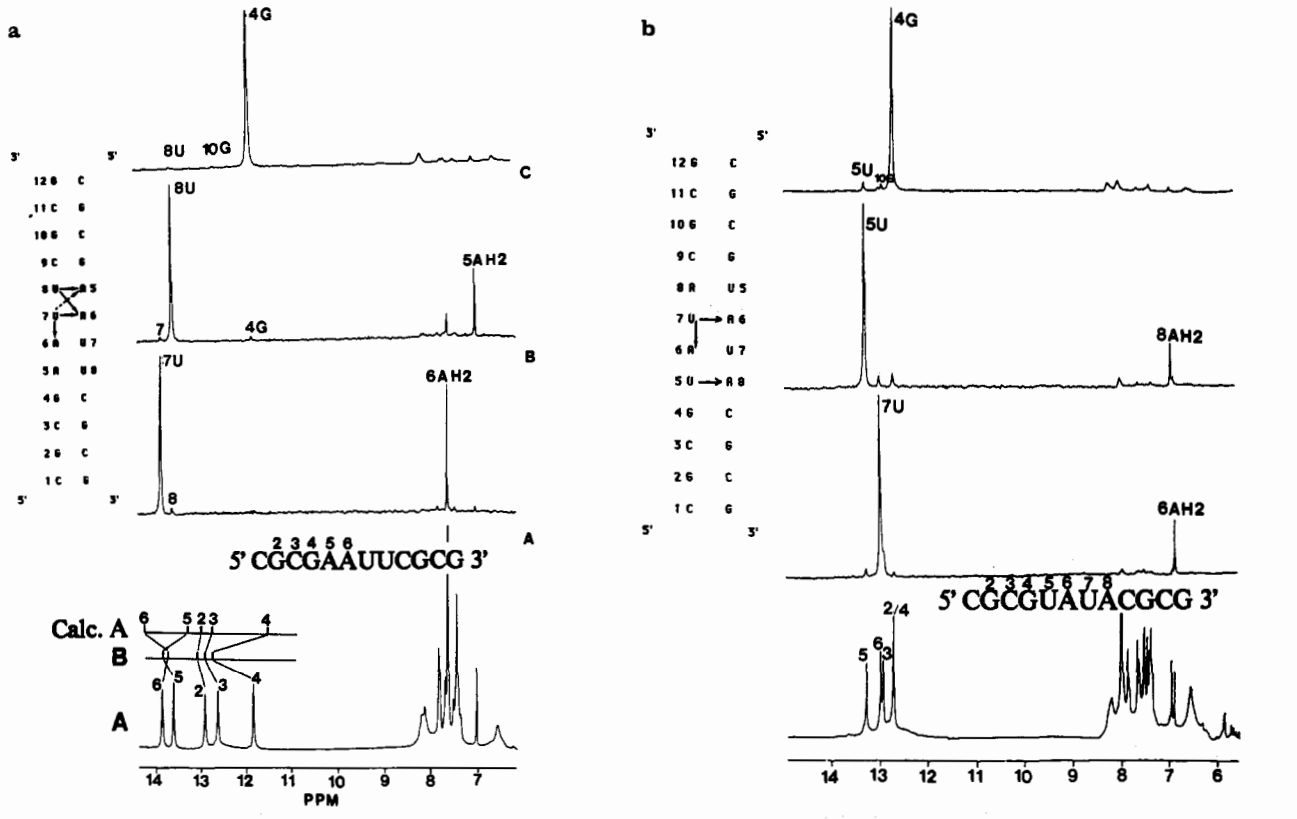


FIGURE 3: (a) Spectrum of the imino and aromatic protons of r(CGCGAAUUCGCG) in H_2O and difference imino NOE spectra. The chemical shifts of the imino protons of the B-form d(CGCGTTAACGCG) and the A-form r(CGCGUUAACGCG) and those calculated on the basis of ring current shifts for A form are also shown at the bottom. The DNA chemical shifts are from Patel et al. (1983). The calculated values are from empirically calibrated ring current shifts based on the spectra of t-RNAs (Shulman et al., 1973); second-order effects were ignored. In the calculation of the chemical shift value of the penultimate base pair, only half of the terminal base pair contribution was counted since it is partially opened at this temperature. The arrows indicate imino to H2 NOEs, with the dotted arrows indicating weak NOEs. (b) Equivalent spectra as in (a) for the sequence r(CGCGUAUACGCG).

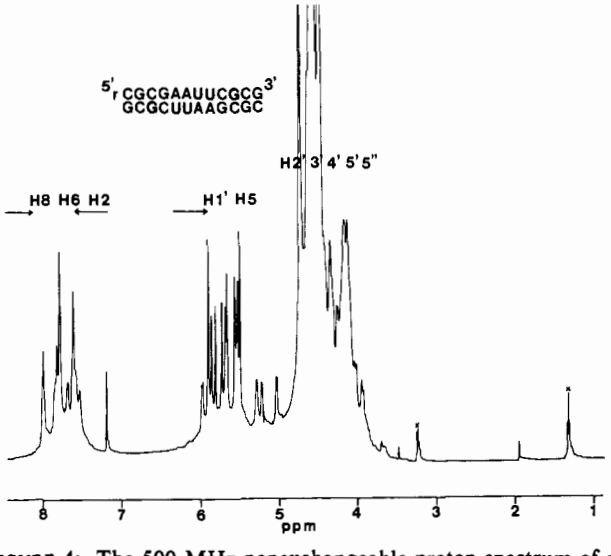


FIGURE 4: The 500-MHz nonexchangeable proton spectrum of r(CGCGAAUUCGCG) in D_2O. The arrows in the base H8, H6, H2 region show the upfield shift of purine H8 protons and the downfield shift of pyrimidine H6 protons, and the arrow in the H1′ region indicates the upfield shift of the H1′ protons compared to their chemical shifts in the DNA analogue.

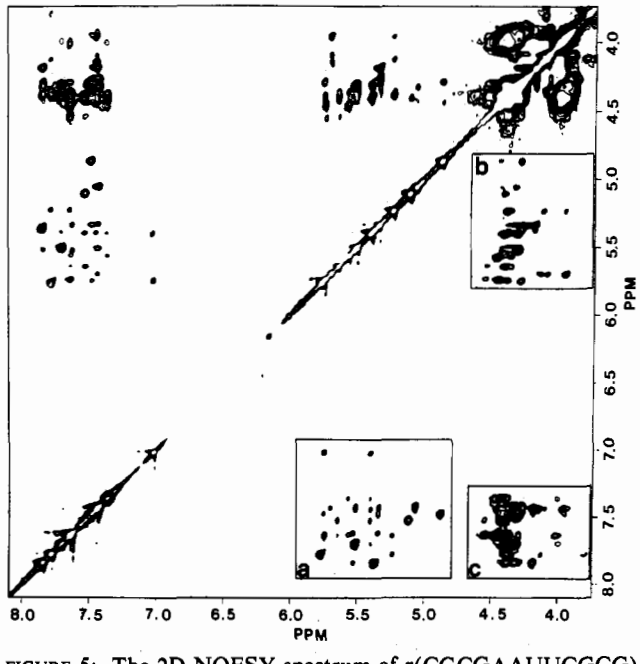


FIGURE 5: The 2D NOESY spectrum of r(CGCGAAUUCGCG) collected at 100-ms mixing time at 32 °C. The aromatic to H1′ region, H1′ to H2′ region, and aromatic to H2′ region are surrounded by boxes a, b, and c and expanded in Figures 6, 7, and 8, respectively.

distance in an ideal A-form structure is around 3 Å while in an ideal B-form structure this distance is about 3.6 Å). Among the H1′ to H2′/H3′/H4′ NOEs, the H1′ to H2′ will always be the strongest of the three since this distance is the shortest H1′ intrasugar distance (ca. 2.7 Å) regardless of the sugar pucker conformation. Thus the H1′–H2′ NOEs were easily identified at higher contour levels where only 12 strong spots were seen (data not shown). The H2′ resonances of 1C, 11C, 8U, 4G, 2G, and 12G were assigned directly from their H1′. The 10G-7U, 3C-9C, and 5A-6A H2′ pairs could not be directly assigned in this way due to overlap of their H1′ chemical shifts. Fortunately, the strong 7U H5 and 8U H5 cross-peaks

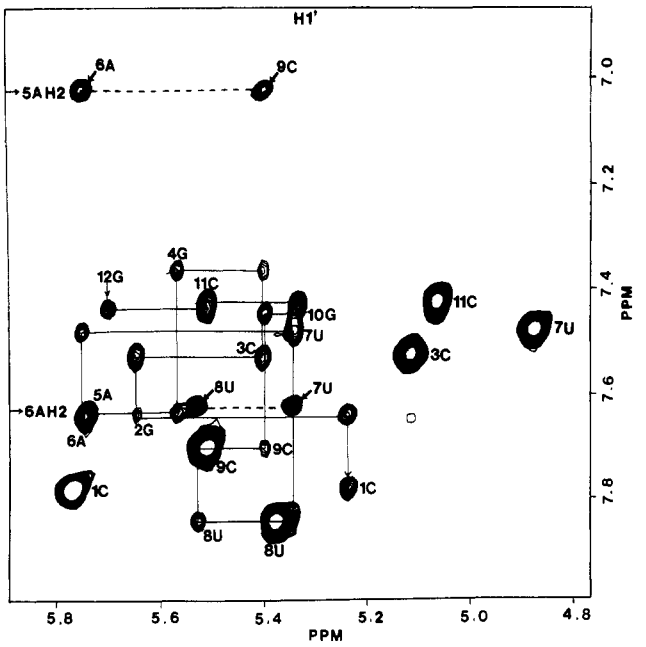


FIGURE 6: Expanded spectrum of the aromatic to H1′ region of r(CGCGAAUUCGCG). The six strong pyrimidine H5 to H6 cross-talks are the most intense peaks. The sequential connectivity from the 3′-end (12G) to the 5′-end (1C) is connected by solid lines. Only the intraresidue cross-peaks are labeled. The horizontal dotted lines connect the intrastrand and interstrand NOEs between A H2 and the $n + 1$ and $m + 1$ H1′s, which are marked with arrows.

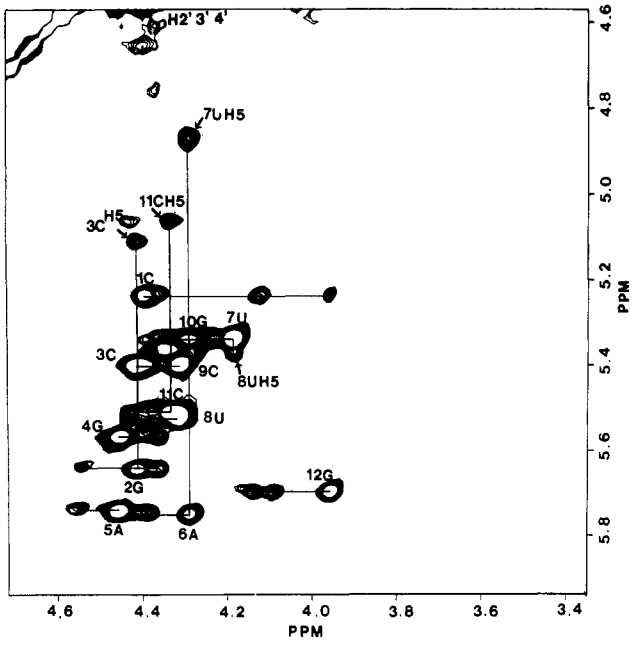


FIGURE 7: Expanded 2D NOESY of the H1′ to H2′ region of r(CGCGAAUUCGCG). Twelve strong H1′ to H2′ cross-peaks were easily detected and are labeled with their residue number. At higher counter levels, these are the only peaks that can be seen (data not shown). Horizontal lines connect the cross-peaks from H1′ to H2′, H3′, and H4′ of the same residue. Since the differentiation between H3′ and H4′ is not reliable based only on the small difference in distance to H1′, they were not assigned at this point. The cross-peaks from the H5 of pyrimidine to the $n - 1$ H2′ or n H2′ are labeled as 7U H5 etc. and are connected to the H1′–H2′ by vertical lines.

to their preceding H2′ clearly identified the H2′ resonances of 6A and 7U, respectively. Thus the H2′ of 6A could be differentiated from that of 5A, and the H2′ of 7U could be differentiated from that of 10G in this way. These assignments were also confirmed from the internucleotide H8/H6 to H2′

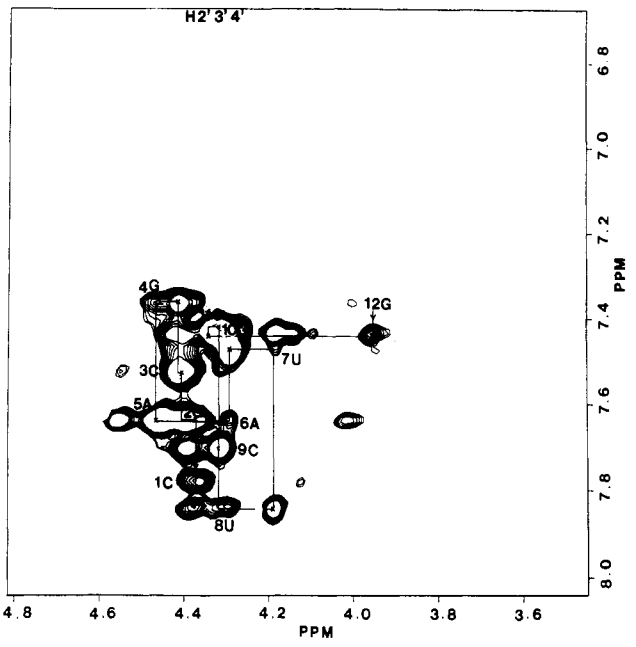


FIGURE 8: Expanded 2D NOESY spectrum of the aromatic to H2′ region of r(CGCGAAUUCGCG). Most of the strong peaks in this region are those from the interresidue aromatic to $n - 1$ H2′ cross-peaks which are an important criterion for the C3′-endo sugar pucker conformation. The sequential connectivity can be connected as in the aromatic to H1′ region, with some help from the H1′ to H2′ region since some of the peaks are overlapped. The intraresidue cross-peaks are labeled as 7U etc., while those of the interresidue NOEs are marked with a ×.

cross-peaks (see Figure 8). It is interesting to note that 11C H5 exhibits an NOE to its own H2′ proton instead of to the H2′ of its 5′-neighbor, which is in sharp contrast to the 7U/8U case where their H5 protons exhibit an NOE only to the 5′-neighbor H2′ and not to their own H2′. We were not able to tell if the 3C H5 has an NOE to 3C H2′ or to 2G H2′ since the 3C H2′ and 2G H2′ have the same chemical shift. It is also noteworthy that, in the same region of the r(CGCGUAUACGCG) spectrum (Figure 12), all pyrimidine H5 resonances show an NOE only to the H2′ of the 5′-neighbors. This different behavior for the cytidine and uridine in the two sequences must be related to differences in local structure. We are, however, unable to say anything definitive about the detailed structure at this point; the solution structure of these RNA duplexes must await accurate distance measurements to be reported at a later date.

The expanded H8/H6 to H2′/H3′/H4′ NOESY region is shown in Figure 8. It is complicated due to extensive overlap of the intraresidue H8/H6 to H2′, H3′, and H4′ NOEs and interresidue H8/H6 to H2′ NOEs. At higher contour levels, we were able to simplify the spectra by sampling only the more intense cross-peaks and were able to extract some important information. The most intense cross-peaks were due to internucleotide aromatic to H2′ NOEs (data not shown). We were therefore able to sequentially connect the H8/H6 to H2′ cross-peaks, as was the case for the H8/H6 to H1′ cross-peaks, and these are connected by solid lines in Figure 8. By themselves, these connectivities were not unambiguous, but since the very strong interresidue H8/H6 to H2′ and intraresidue H1′ to H2′ NOEs were complementary to each other, we were able to reliably connect these peaks. The interresidue NOEs are denoted with an "x" in the figure. We have resisted the temptation to specifically assign the H3′ and H4′ resonances at present since this would involve a structural assumption that may not be warranted. The assignment of these additional

Table I: Comparison of A-Type and B-Type Chemical Shifts (ppm) of r(CGCGAAUUCGCG)[a]

				1 C	2 G	3 C	4 G	5 A	6 A	7 U	8 U	9 C	10 G	11 C	12 G
	H8 and C 6H			C 1'H			C 5H			C 2H			C 2'H
	B	A	Δ (B − A)	B	A	Δ (B − A)	B	A	Δ (B − A)	B	A	Δ (B − A)	A
1C	7.6	4.8	***−0.2***	5.7	5.3	**0.4**	5.69	5.8	−0.11				4.39
2G	7.9	7.62	**0.28**	5.83	5.64	**0.19**							4.42
3C	7.24	7.5	***−0.26***	5.55	5.39	**0.16**	5.35	5.1	0.25				4.42
4G	7.82	7.35	**0.47**	5.4	5.55	**−0.15**							4.45
5A	8.1	7.61	**0.49**	5.94	5.74	**0.2**				7.16	7.01	0.15	4.46
6A	8.1	7.61	**0.49**	6.1	5.74	**0.36**				7.58	7.61	−0.03	4.25
7U (T)	7.1	7.46	***−0.36***	5.85	5.34	**0.51**		4.86					4.17
8U (T)	7.35	7.8	***−0.45***	6.08	5.52	**0.56**		5.36					4.28
9C	7.43	7.68	***−0.25***	5.68	5.39	**0.29**	5.6	5.5	0.1				4.28
10G	7.88	7.43	**0.45**	5.8	5.35	**0.45**							4.25
11C	7.3	7.41	***−0.11***	5.73	5.5	**0.23**	5.4	5.05	0.35				4.32
12G	7.9	7.42	**0.48**	6.09	5.68	**0.41**							3.94

[a] The data for the B-DNA duplex is from Hare et al. (1983). The Δ value is the chemical shift difference between the B-DNA and A-RNA.

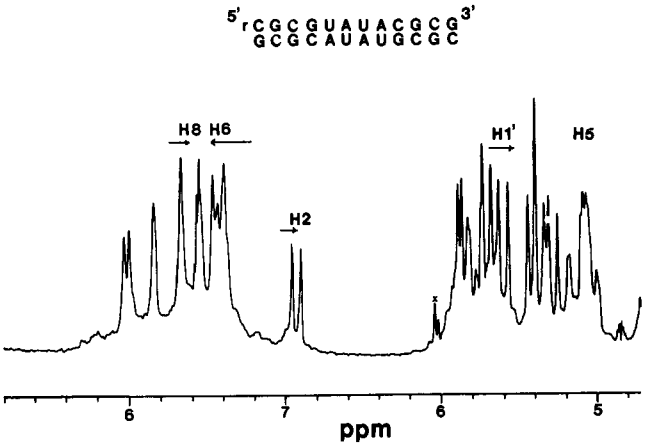


FIGURE 9: The 500-MHz nonexchangeable proton spectrum of the r(CGCGUAUACGCG) in D_2O at 32 °C. The arrows in the base H8, H6, H2 region show the upfield shift of purine H8 protons and adenosine H2 protons and downfield shift of pyrimidine H6 protons. The arrow in the H1′ region indicates the upfield shift of the H1′ protons compared to their chemical shifts in the DNA analogue.

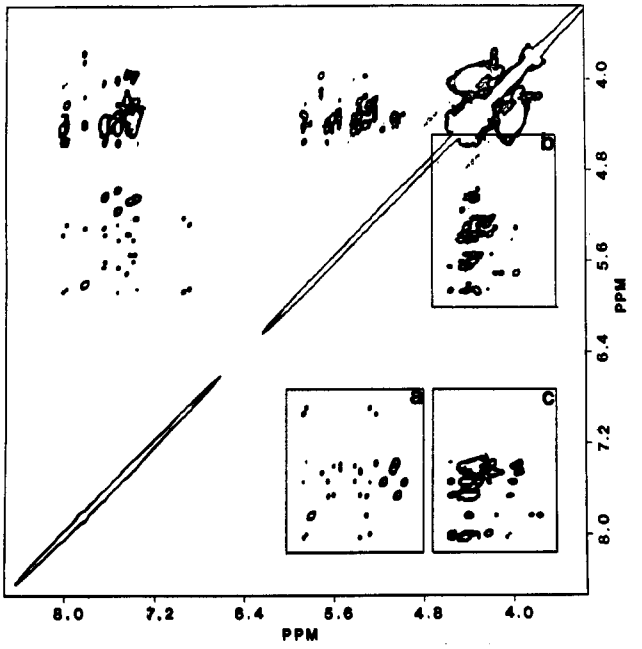


FIGURE 10: The 2D NOESY spectrum of r(CGCGUAUACGCG) collected at 200-ms mixing time at 31 °C. The aromatic to H1′, H1′ to H2′, and aromatic to H2′ regions are surrounded by boxes, a, b, and c and expanded in Figures 11, 12, and 13, respectively.

sugar resonances must await more reliable analysis of through-bond coupling (COSY, RELAY, TOCSY) to the assigned H1′ and H2′ resonances. The assignments and their comparison to the analogous DNA sequence are listed in Table I.

When the assigned chemical shifts of the RNA dodecamer r(CGCGAAUUCGCG) were examined and compared to those for d(CGCGAATTCGCG), several interesting patterns emerged. The chemical shifts of the purine H8 resonances move upfield in the ribo duplex by about 0.35 ppm, while those of the pyrimidine H6 resonances move downfield by about 0.25 ppm compared to those of the DNA analogue. Also, almost all the H1′ and H5 resonances move upfield in the RNA (see Table I). In order to test whether this pattern of chemical shift change is a general RNA phenomenon and to see if the interstrand A H2 to H1′ NOE is reciprocal, we also synthesized the related double-helical RNA oligomer r(CGCGUAUACGCG) and qualitatively analyzed its structure using the same techniques. The DNA analogue of this sequence has been studied previously (Wemmer et al., 1985), and this sequence also constituted a second RNA duplex in which to monitor the cross-strand A H2 to H1′ NOE pattern as well as the sequence-dependent effects of repeated pyrimidine–purine steps.

The 1D nonexchangeable proton spectrum of the H8, H6, H5, and H1′ resonances of r(CGCGUAUACGCG) is shown in Figure 9. The spectral resolution of this sequence is better than that of r(CGCGAAUUCGCG), and at least 10 well-resolved H1′ singlets can be seen. As was the case in r(CGCGAAUUCGCG), the singlet fine structure for the H1′ resonances indicates a H1′–C1′–C2′–H2′ dihedral angle close to 90° and suggests that the sugar conformations are near the C3′-endo (N) domain. Two sharp adenine H2 singlets are also seen at the high-field (~7 ppm) end of the aromatic region.

The complete NOESY spectrum at 200 ms is shown in Figure 10, and the H6/H8 to H1′ region of this RNA dodecamer is shown expanded in Figure 11. The well-resolved cross-peaks make the base and H1′ assignments very straightforward. The two A H2 resonances are far upfield from the other aromatic protons, and their interstrand and intrastrand cross-peaks to H1′ are as strong as those in r(CGCGAAUUCGCG). Furthermore, the H6/H8 to H1′ cross-peaks showed more intensity variations than those in (CGCGAAUUCGCG). The C9 H6 to A8 H1′ and U7 H6 to A6 H1′ cross-peaks are quite weak while the 10G–9C, 8A–7U, and 6A–5U cross-peaks are much stronger. These different NOE intensities indicated sequence-dependent

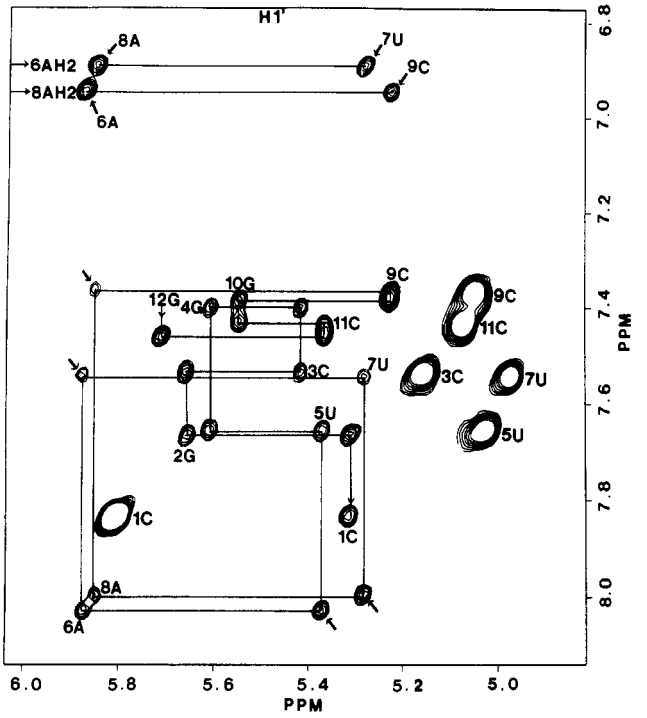


FIGURE 11: Expanded NOESY spectrum of the aromatic to H1′ region of r(CGCGUAUACGCG). The sequential connectivity starts at 12G and ends at 1C. Only the intraresidue cross-peaks are labeled. The two A H2 to H1′ connectivities were connected by horizontal lines in the upper part of the spectrum and marked with arrows. The NOEs between 6A H2 and 8A H1′ as well as 8A H2 to 6A H1′ are clearly seen and are of special significance since in the DNA–RNA hybrid (see accompanying paper) these NOEs are observed only in the direction from the C2′-endo strand to the C3′-endo strand. Also note the different NOE intensities of the aromatic to H1′ peaks in this region, which indicates structural irregularities in this A-form RNA dodecamer.

structure variation and can be rationalized by increased roll angles toward the minor groove between these adjacent bases (see Discussion). However, these effects are not seen in the CG dinucleotide steps, which exhibit almost equal interresidue H6/H8 to H1′ NOEs, indicating a different response to steric clash by UA compared to CG dinucleotide steps. This difference may be due to more rigid, less distortable, hydrogen bonding in CG base pairs. It is not unreasonable to assume that UA or TA dinucleotide steps may take on larger propeller twist angles and cause more local structure perturbations than CG dinucleotide steps since the former contain only two hydrogen bonds and have more freedom to assume more stable base stacking interactions. Also the corresponding NOE intensities in r(CGCGAAUUCGCG) are more or less equal because the 5′GAAUUC3′ sequence in the central region does not contain any pyrimidine–purine (5′–3′) dinucleotide steps.

The expanded NOESY spectrum of the H1′/H5 to H2′/H3′/H4′ region of this dodecamer is shown in Figure 12. As in the r(CGCGAAUUCGCG) sequence, the 12 H1′ to H2′ cross-peaks were the strongest and were accordingly assigned without any problem, except for the 5U and 11C pair which overlap but can be assigned by referring to the H6/H8 to H2′ region. The cross-peaks between pyrimidine H5 resonances and the H2's of the 5′-neighbor differ markedly for these two sequences. In r(CGCGUAUACGCG), the H5s of 3C, 11C, 5U, and 9C have cross-peaks to the preceding H2′ that are of similar intensity, while 7U has a weaker NOE to 6A H2′; in r(CGCGAAUUCGCG) however, the 7U H5 to 6A H2′ NOE is one of the strongest cross-peaks. Also the 11C H5 now exhibits an NOE to 10G H2′ instead of to its own H2′.

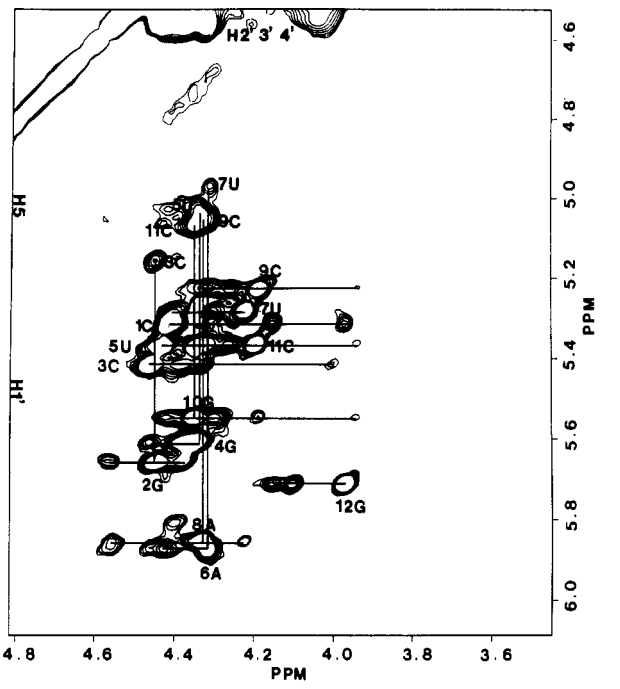


FIGURE 12: Expanded spectrum of the H1′/H5 to H2′/H3′/H4′ region of r(CGCGUAUACGCG). Twelve strong H1′ to H2′ cross-peaks were easily detected and are labeled 1C etc. Horizontal lines connect the NOEs from H1′ to H2′, H3′, and H4′ of the same residues. As in r(CGCGAAUUCGCG), the H3′ and H4′ pairs were not assigned at this point. The cross-peaks from the H5 of pyrimidine to the $n-1$ H2′ are labeled as 7U H5 etc. and are connected to the H1′–H2′ by vertical lines. The NOE between 7U H5 and 6A H1′ is very weak while in r(CGCGAAUUCGCG) it is very strong (see Figure 6), and this indicates structural differences between r(CGCGUAUACGCG) and r(CGCGAAUUC gCG).

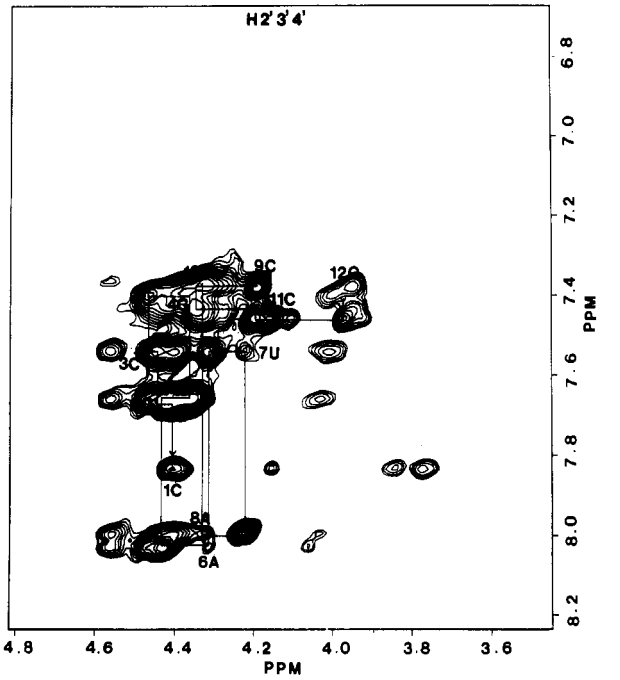


FIGURE 13: Expanded 2D NOESY spectrum of the aromatic to H2′ region of r(CGCGUAUACGCG). The sequential connectivity was also traced as in the aromatic to H1′ region, with the help of the H1′ to H2′ region. The intraresidue cross-peaks are labeled 7U etc., while the interesidue ones are marked with a ×. The peaks in this region were more resolved that those of the r(CGCGAAUUCGCG), and the interresidue cross-peaks were more easily identified.

Such intensity variations must reflect local sequence-dependent structural differences. The expanded NOESY spectrum of

Table II: Comparison of A-Type and B-Type Chemical Shifts (ppm) of r(CGCGUAUACGCG)[a]

	1 C	2 G	3 C	4 G	5 U	6 A	7 U	8 A	9 C	10 G	11 C	12 G
	H8 and C 6H			C 1'H			C 5H			C 2H			C 2'H
	B	A	Δ (B − A)	B	A	Δ (B − A)	B	A	Δ (B − A)	B	A	Δ (B − A)	A
1C	7.48	7.86	***−0.38***	5.61	5.34	**0.27**	5.74	5.82	−0.08				
2G	7.78	7.7	**0.08**	5.73	5.64	**0.09**							4.45
3C	7.15	7.55	***−0.4***	5.55	5.42	**0.13**	5.55	5.16	0.39				4.46
4G	7.71	7.4	**0.31**	5.79	5.61	**0.18**							4.14
5U (T)	7.01	7.65	***−0.64***	5.54	5.37	**0.17**		5.02					4.43
6A	8.09	8.02	**0.07**	6.03	5.87	**0.16**				7.14	6.94	0.2	4.31
7U (T)	6.96	7.55	***−0.59***	5.43	5.28	**0.15**		4.95					4.22
8A	8.02	7.98	**0.04**	5.96	5.84	**0.12**				7.21	6.98	0.23	4.33
9C	7.06	7.38	***−0.32***	5.4	5.22	**0.18**	5.05	5.04	0.01				4.20
10G	7.66	7.39	**0.27**	5.67	5.55	**0.12**							4.14
11C	7.13	7.43	***−0.3***	5.6	5.36	**0.24**	5.22	5.07	0.15				4.20
12G	7.75	7.46	**0.29**	5.97	5.71	**0.26**							3.96

[a] The data for the B-DNA duplex is from Wemmer et al. (1985). The Δ value is the chemical shift difference between the B-DNA and A-RNA.

the H6/H8 to H2'/H3'/H4' region is shown in Figure 13, with sequential connectivities along the sequence. All the assigned base proton, H1', and H2' chemical shifts of this dodecamer are listed in Table II.

Discussion

RNA is generally believed to assume a C3'-endo conformation because of steric constraints involving the 2'-OH group. However, recent data on DNA–RNA hybrids suggested that RNA can also take on a C2'-endo conformation in solution (Reid et al., 1983a; Gupta et al., 1985); in the fully solvated fiber state the hybrid duplex has many similarities to B-DNA and is unlike A-DNA, yet the ribo strand sugar pucker is C3'-endo (Zimmerman & Pheiffer, 1981). The data presented here strongly indicate that in ribo–ribo duplexes both RNA dodecamers adopt a C3'-endo A-type conformation. First, the chemical shifts of the imino protons of r(CGCGAAUUCGCG) were compared to those of the analogous DNA sequences and to the calculated values based on t-RNA data (Schulman et al., 1973) and found to agree well with the A-type conformation. Second, in both RNA duplexes the anomeric H1' resonances appear as sharp singlets in the one-dimensional nonexchangeable spectra (Figures 4 and 9), indicating weak coupling to H2'. Thus the dihedral angle between H1' and H2' is close to 90°, which is typical of a C3'-endo sugar pucker, in contrast to the C2'-endo B conformation where this angle is 180° and the H1'–H2' coupling is maximal. This conclusion is corroborated by the lack of H1'–H2' cross-peaks in the 2D COSY spectrum, except for a very weak 1'–2' cross-peak for residue 12G (data not shown). Third, computer models of idealized A- and B-form duplexes show very different distances between base H8/H6 protons and the H2' and H3' sugar protons (see Figure 2). In A-form duplexes, the base protons are closer to the preceding H2' (~1.8 Å) than to their own H2' (~3.9 Å) whereas in B-form duplexes this is reversed and the base protons are now closer to their own H2' (~1.9 Å) than to the preceding H2' (~4.0 Å). Experimentally, we found that most of the strong H8/H6 to H2' peaks were indeed due to internucleotide NOEs and the intranucleotide H8/H6 to H2' cross-peaks were very weak in both dodecamers. In the well-resolved cases of 8U–7U, 7U–6A, 6A–5A, 4G–3C, and 2G–1C in r(CGCGAAUUCGCG) and 12G–11C, 11C–10G, 10G–9C, 8A–7U, 7U–6A, 6A–5U, and 4G–3C in r(CGCGUAUACGCG) the internucleotide H8/H6 to H2' cross-peaks are very strong, and their intensity is as strong or stronger than the H5 to H6 cross-peaks which correspond to a distance of ca 2.4 Å (see the whole plots in Figures 5 and 10). Thus the majority of internucleotide H8/H6 to H2' distances are probably 2.4 Å or less and are in the distance range expected for A-type double-helical structure but not for B-type structure. The NOE data on B-type DNA dodecamers d(CGCGAATTCGCG) (Hare et al., 1983) and d(CGCGTATACGCG) (Wemmer et al., 1985) indicate that the distances between base protons and their own H2' protons are much shorter than the internucleotide H8/H6 to H2' distances. On the basis of the above observations, it appears that these two RNA dodecamers exhibit all the major characteristics of a normal A-type duplex with C3'-endo sugar conformation. The H8/H6 to H2' NOE has previously been used by others as a criterion to identify the C3'-endo sugar conformation (Reid et al., 1983b; Haasnoot et al., 1984). The C3'-endo sugar conformation was also found in a small RNA pentamer and in two RNA hexamers (Clore et al., 1985; Westerink et al., 1984; Hasnoot et al., 1984). In addition to the H8/H6 to H2' cross-peaks, the pyrimidine H5 to H2' cross-peaks also revealed interesting features. Most of these medium cross-peaks are due to interresidue NOEs, and the ca. 3-Å intensities are close to those expected for idealized A-form geometry. However, exceptions were found; i.e., the H5 of 11C in the CGCGAAUUCGCG sequence exhibits an NOE to its own H2' instead of to its 5'-neighbor and 7U H5 has only a weak NOE to 6A H2'. These effects presumably reflect fine-structure variations in local conformation.

The NOEs between H8/H6 and the H1' of the 5'-neighbor also exhibit sequence-dependent behavior in r(CGCGUAUACGCG). For purines preceded by pyrimidines the NOEs are very strong (8A–7U, 6A–5U), while for pyrimidines preceded by purines they are very weak (9C–8A, 7U–6A). Interestingly, when compared to the DNA analogue d(CGCGTATACGCG) (Wemmer et al., 1985), the reverse was found in that the base to $n-1$ H1' NOEs for 8–7 and 6–5 are weak and the 9–8 and 7–6 NOEs have become stronger. This interesting pattern reversal is most likely due to opposite relative roll angles toward the minor groove. In the A-type DNA crystal data on d(GGTATACC), the relative roll angles were found to be around 12°–15° in TA steps but closer to 0° in AT steps. Since our RNA dodecamers are also in the A-type conformation, we believe these same roll angle differences may well exist in the UA and AU steps. The structure obtained from poly[d(AT)] fiber data (Millane et al., 1984) indicates opposite relative roll angles of TA and AT to those found in regular B-type DNA (Dickerson, 1983). The roll angles for TA in A-type poly[d(AT)] have opened toward the major groove whereas those of the AT steps open toward the minor groove; such a phenomenon would nicely explain

our NOE results. By inspecting the regular A-type structure (Figure 2), we see that opening the angle into the minor groove will decrease the distance between the n H8 and the $n-1$ H1′ and increase the distance between the $n-1$ H6 and the $n-2$ H1′; conversely, opening the angle toward the major groove will increase the n H8 to $n-1$ H1′ distance and decrease the $n-1$ H6 to $n-2$ H1′ distance. This would nicely rationalize the different NOE patterns observed in d(CGCGTATACGCG) and r(CGCGUAUACGCG). This opening up of the pyrimidine–purine (5′–3′) steps toward the minor groove also increases the distance between the H6 and its own H1′ and offers an explanation for the very weak 7U H6 to 7U H1′ cross-peak. The 9C H6 to 9C H1′ cross-peak may also be weak, but it is unfortunately overlapped with the strong 10G–9C cross-peak and cannot be examined. This pattern is not observed in r(CGCGAAUUCGCG) since their is no pyrimidine–purine (5′–3′) step in the central GAAUU sequence.

Having confirmed that both of these RNA dodecamers are in the A-type conformation, we next sought a structural explanation of why the interstrand or intrastrand NOEs between the adenine H2 and the H1′ of the following base pair are so strong. In the following discussion we will designate the nucleotides of a base pair as residues n and m. In order to avoid ambiguity in such terms as "interstrand NOEs to the following base pair", we will designate such cross-strand NOEs as n to m, n to $m-1$, or n to $m+1$, where $m-1$ precedes, i.e., is 5′ to m and is paired to $n+1$, and conversely $m+1$ follows: i.e., is 3′ to m and is paired to $n-1$. As shown in Figure 2, in ideal A-form or B-form duplexes there are no H1′ protons within 4.6 Å of the adenine H2. In agreement with this, the experimental data on B-form DNA indicate that most of the A H2 to H1′ NOEs are either very weak or nonexistent at all reasonable mixing times. However, the strong A H2 to H1′ NOEs in the present RNA studies can be explained by invoking the B–A transition model proposed by Callidine and Drew (1984). In their model, the interstrand steric clash between purine bases in a pyrimidine-purine (5′–3′) step can be relieved either by sliding back the propeller-twisted base pairs by about 0.5 Å in a B-form DNA to favor intrastrand base overlap or by sliding forward the propeller-twisted base pairs by about 1.5 Å to permit more extensive cross-strand overlap in A-form DNA. In B-form DNA, the roll angles remain near zero, while in A-form DNA the roll angle opens up by about 15° toward the minor groove (see Figure 14). Since our NOE data clearly indicate the A-type conformation for our RNA dodecamers, the "sliding in" of the propeller-twisted base pairs by about 1.5 Å would account for about 1.1-Å distance reduction in the interstrand distance between an adenine H2 on residue n and the H1′ of residue $m+1$. The resulting 3.5-Å distance (4.6 Å − 1.1 Å) would then explain the observed NOE intensity for these types of cross-peaks which are somewhat stronger than those between the H6/H8 and H1′ protons (an average distance of about 3.8 Å). The idealized A-form model cannot account for these strong cross-strand NOEs, probably because the fiber diffraction data from which the model was constructed are of too low resolution to reveal the fine details of sequence-dependent structure variations (Callidine & Drew, 1984; Dickerson, 1983). When the A-type crystal structure of d(GGTATACC) (Shakked et al., 1983) is examined carefully, one does observe that the interstrand and intrastrand adenine H2 to $m+1$ and $n+1$ H1′ distances are equal to or smaller than those of the H8/H6 to H1′ distances. Furthermore, the C3′-endo and C2′-endo hybrid model of dAAA·dTTT proposed by Arnott et al. (1983)

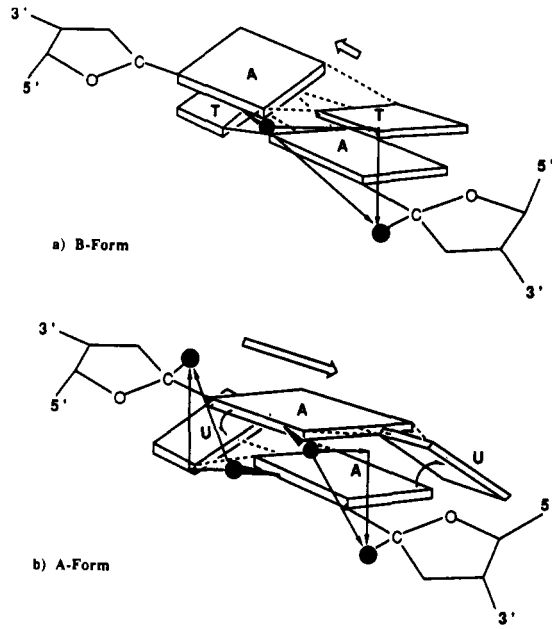


FIGURE 14: A dinucleotide base pair in A-form and B-form structure. The pyrimidine (Y) and purine (R) bases are drawn as small and large rectangles. A H2 and H1′ are shown as solid black circles. The interstrand distances between the A H2 and the $m+1$ adenine H1′ in both structures were connected through triangles. The roll angles in the A form are denoted by curved arrows. The base sliding in both structures is represented by white arrows. For simplicity, the helical twisting is not shown.

reveals a 3.9-Å distance for the interstrand adenine H2 to $m+1$ H1′ (see Figure 1b of accompanying paper).

The above types of NOEs can serve as criteria for evaluating A- or B-form structures in nucleic acid duplexes since the B-type A H2 to H1′ NOEs should be very weak according to the Callidine and Drew model. As shown in Figure 2, one would not expect to see any A H2 to H1′ cross-peaks in perfect B-form DNA since there are no H1′ protons within 4.6 Å of any adenine H2 protons. Although extensive local deviations of the structure from idealized B-form DNA were observed in single crystals of d(CGCGAATTCGCG) (Dickerson & Drew, 1981), the shortest distance of any adenine H2 to any H1′ proton is still above 4 Å; such distances would give only very weak NOEs, and that is precisely what we observed in solution (Hare et al., 1983; Wemmer et al., 1985); even at the 300-ms mixing time used in the DNA experiments, only very weak adenine H2 to H1′ NOEs were seen. It is, however, interesting to note that more intense adenine H2 to H1′ NOEs do appear in the NOESY spectra of bent DNA sequences containing oligo(A) runs (Kintanar et al., 1987), of DNA fragments containing modified bases (Patel et al., 1986), and of poly(dA-dT) (Assay-Munt & Kearns, 1984), but they are still weaker than most of the H8/H6 to H1′ NOEs (3.5–3.9 Å). Equally strong H2–H1′ and H8/H6–H1′ cross-peaks have never been observed in DNA molecules, and such cross-peaks may be uniquely characteristic of A-form RNA molecules. This type of cross-peak has also been detected in a pentameric RNA by Clore et al. (1985).

We have carried out a comparison of the chemical shifts in A-type and B-type duplexes of the same sequence, since the B-form DNA of the present A-form RNAs have been assigned previously (Hare et al., 1983; Wemmer et al., 1985). Some interesting chemical difference patterns can be extracted from the data (see Tables I and II). First, in r(CGCGAAUUCGCG) (Table I), all the H1′ protons move upfield by 0.16–0.56 ppm, with the exception of 4G whose H1′ moves

downfield by about 0.15 ppm; the H1′s of U7 and U8 are the most upfield shifted. The quite large upfield shifts of the H1′ protons must be due to conformational changes because the substitution of a hydrogen atom by a hydroxy group at C2′ should induce a downfield shift rather than an upfield shift. Second, all purine H8 protons move upfield by 0.3–0.5 ppm while cytosine H6 protons and uridine H6 protons move downfield by about 0.2 and 0.4 ppm, respectively. Third, all the H5 protons (except for 1C) move upfield, even though their H6 protons move downfield. These different patterns of chemical shift change clearly indicate different stacking geometries of the bases in the two types of duplex. To test the generality of this chemical shift change pattern, the variant dodecamer r(CGCGUAUACGCG) was synthesized and subjected to a similar analysis. Although specific chemical shifts depend on the base sequence, we find a similar trend in the chemical shift changes for this dodecamer, but to a somewhat different extent. As with the chemical shift change in r(CGCGAAUUCGCG), all the H1′ of r(CGCGUAUACGCG) move upfield compared to those of d(CGCGTATACGCG) by about 0.09–0.27 ppm; although still substantial, the change is not as drastic as for the AAUU duplex. The purine H8 resonances also move upfield but to a lesser extent, especially the two adenine H8s which only move about 0.04 and 0.07 ppm; in r(CGCGAAUUCGCG) both adenine H8 peaks move upfield by 0.49 ppm compared to those their DNA counterpart. The pyrimidine H6 resonances in r(CGCGUAUACGCG) move downfield, as they do in r(CGCGAAUUCGCG), but to a larger extent. 5U–H6 and 7U–H6 move downfield by 0.64 and 0.59 ppm, compared to 0.36 and 0.45 ppm for the r(CGCGAAUUCGCG) duplex. Again, as in the former duplex, all H5 peaks move upfield while all H6 peaks move downfield. By comparison of these two sequences with their DNA analogues, the following trends were found as a consequence of changing the sugar pucker conformation from B form to A form: (1) all H1′s move upfield; (2) all purine H8 protons moved upfield while pyrimidine H6 protons move downfield; (3) except for the 3′-terminal residue, all cytosine H5 protons move upfield while their H6 resonances moved downfield. Although it is still too early to generalize these patterns of chemical shift change as definitive criteria for assessing B-form or A-form duplex structures, it is certainly an attractive possibility, and we hope to extend the correlation to additional sequences. Since the chemical shift of the nucleotide is very dependent on the geometry of the surrounding bases, the proposed 1.5-Å sliding in of the base pairs should induce large local ring current mediated chemical shift changes, and the base pair reorientation may be such as to shield the purine H8 and pyrimidine H5 and deshield the pyrimidine H6 protons.

As shown in this paper, the C 3′H, 4′H, 5′H, and 5″H are too overlapped in dodecamer duplexes to be easily assigned by simple COSY/NOESY methods. Work is in progress to assign such resonances in shorter RNA duplexes with simpler NMR spectra. Even at the dodecamer level, we are optimistic that many of the 3′H and 4′H resonances can be assigned from suitably optimized RELAY and TOCSY spectra (Flynn et al., 1988).

In terms of purifying synthetic oligonucleotides in preparative (20 mg) quantities, we found TLC to be an effective, reasonably efficient, and economic method. The resolution is good, and up to 400 A_{260} of pure RNA or DNA can be purified from ten analytical plates in a short time. The most important advantage of TLC over the more expensive HPLC approach is that it can be used for almost any type of oligonucleotide sequence without suffering aggregation problems. It has been shown that oligonucleotides containing self-complementary sequences or G-rich sequences are rather difficult to separate by HPLC unless very strong denaturing agents are used in the solvent system (Newton et al., 1983). The weakly alkaline developing solvents used in TLC prevent oligonucleotides from self-aggregation and facilitate separation. Although cleavage of RNA chains under alkaline conditions is possible, we found no evidence for cleavage under the mild developing conditions described above.

With respect to the chemical strategy of RNA synthesis, we tried several different 2′-OH protection groups and found that *tert*-butyldimethylsilyl is the only group that produces excellent yields in a reasonable time in automated 10-μmol syntheses. Although the tetrahydropyranyl group (or methoxytetrahydropyranyl) is compatible with small-scale manual synthesis under very anhydrous conditions, it is extremely difficult to scale up this approach for an automated synthesizer since the acid deprotection of the 5′-DMT group takes about 3 min and this causes quite significant 2′-deprotection. The absence of 3′–2′ migration and the presence of only the correct 3′- to 5′-phosphodiester linkage in the RNA produced with 2′-*tert*-butyldimethylsilyl protection was shown by the complete digestion of the synthetic RNA by RNase T2 within 1 h (data not shown); RNase T2 is completely specific for 3′- to 5′-phosphodiesters and does not cleave 2′-5′ phosphodiesters (Kierzek et al., 1986). Although migration is possible under alkaline and protonic conditions, it is definitely not a problem in the phosphoramidite synthesis approach. In designing our RNA synthesis strategy, we also tested ribose β-cyanoethyl phosphoramidites as well as methyl phosphoramidites. Although the β-cyanoethyl triester is quite popular in DNA synthesis, we found that the methoxy phosphoramidite gave much better coupling yields than the β-cyanoethyl analogue. It is possible that steric hindrance by the more bulky 2′-*tert*-butyldimethylsilyl protecting group prevented the β-cyanoethyl phosphoramidite from reacting with the 5′-OH group in the solid phase.

Recently, two papers involving enzymatic RNA synthesis were reported (Sharmeen & Taylor, 1987; Milligan et al., 1987). Although such methods are promising for making long synthetic RNA, the optimum conditions have to be worked out for each DNA template. A further disadvantage of the enzymatic approach is that there is no easy way to put deoxynucleosides or modified ribonucleoside at a specific position in the RNA. For shorter RNA oligomers (<12), the efficiency of the enzymatic method might not be good enough for practical preparation of large-scale (20–30 mg) 2D NMR samples. Depending on the uses to which the product will be put, it is apparent that the methods are complementary, and both methods for RNA synthesis will be useful for different applications.

We have now been able, for the first time, to study the structure of a full-turn RNA duplex by 2D NMR. Over the past few years, several X-ray and 2D NMR studies have been carried out on synthetic DNA duplexes that have revealed many important aspects about DNA structure, both in the solid state and in the aqueous state. The present synthetic and NMR studies now pave the way for analogous systematic studies on RNA structure as a function of sequence. In the following paper we present the results of a 2D NMR study of a DNA–RNA hybrid duplex.

ACKNOWLEDGMENTS

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

continuous encouragement during this project.

Registry No. **1a**, 58-61-7; **1a** [2′,3′,5′-tris(*O*-trimethylsilyl)], 10457-15-5; **1b**, 65-46-3; **1b** [2′,3′,5′-tris(*O*-trimethylsilyl)], 51432-41-8; **1c**, 118-00-3; **1c** [2′,3′,5′-tris(*O*-trimethylsilyl)], 53274-35-4; **1d**, 58-96-8; **2a**, 118684-36-9; **2a** (O-deprotected), 4546-55-8; **2b**, 118684-37-0; **2b** (O-deprotected), 13089-48-0; **2c**, 118684-38-1; **2c** (O-deprotected), 64350-24-9; **2d**, 10457-16-6; **3a**, 81246-82-4; **3b**, 81246-76-6; **3c**, 81246-83-5; **3d**, 81246-79-9; **4a** (2′-*O*-TBDMS), 81265-93-2; **4a** (3′-*O*-TBDMS), 81256-88-4; **4a** [2′-*O*-TBDMS, 3′-*O*-P(OMe)N(*i*-Pr)$_2$], 118684-39-2; **4b** (2′-*O*-TBDMS), 81256-87-3; **4b** (3′-*O*-TBDMS), 81246-78-8; **4b** [2′-*O*-TBDMS, 3′-*O*-P(OMe)N-(*i*-Pr)$_2$], 118684-40-5; **4c** (2′-*O*-TBDMS), 81279-39-2; **4c** (3′-*O*-TBDMS), 81256-89-5; **4c** [2′-*O*-TBDMS, 3′-*O*-P(OMe)N(*i*-Pr)$_2$], 118684-41-6; **4d** (2′-*O*-TBDMS), 81246-80-2; **4d** (3′-*O*-TBDMS), 81246-81-3; **4d** [2′-*O*-TBDMS, 3′-*O*-P(OMe)N(*i*-Pr)$_2$], 114207-67-9; DMT-Cl, 40615-36-9; TBDMS-Cl, 18162-48-6; r(CGCGAAU-UCGCG), 118681-51-9; r(CGCGUAUACGCG), 118684-42-7; Me_3SiCl, 75-77-4; PhCOCl, 98-88-4; $(CH_3)_2CHCOCl$, 79-30-1; ClP(OMe)N(*i*-Pr)$_2$, 86030-43-5.

References

Arnott, S., & Hukins, D. W. L. (1972a) *Biochem. Biophys. Res. Commun. 47*, 1504.

Arnott, S., & Hukins, D. W. L. (1972b) *Biochem. Biophys. Res. Commun. 48*, 1392.

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

Assa-Munt, M., & Kearns, D. R. (1984) *Biochemistry 23*, 791.

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

Beaucage, S. L., & Caruthers, M. H. (1981) *Tetrahedron Lett. 22*, 1859.

Callidine, C. R., & Drew, H. R. (1984) *J. Mol. Biol. 178*, 773.

Chou, S. H., Hare, D. R., Wemmer, D. E., & Reid, B. R. (1983) *Biochemistry 22*, 3037.

Chou, S. H., Wemmer, D. E., Hare, D. R., & Reid, B. R. (1984) *Biochemistry 23*, 2257.

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

Christodoulou, C., Agrawal, S., & Gait, M. J. (1986) *Tetrahedron Lett. 27*, 1521.

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

Dickerson, R. E. (1983) *J. Mol. Biol. 166*, 419.

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

Dickerson, R. E., Drew, H. R., Conner, B. N., Wing, R. M., Fratini, A. V., & Kopka, M. L. (1982) *Science 216*, 475.

Drew, H. R., Wing, R. M., Takano, T., Broka, C., Tanaka, S., Itakura, K., & Dickerson, R. E. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 2179.

Feigon, J., Leupin, W., Denny, W. A., & Kearns, D. R. (1983) *Biochemistry 22*, 5943.

Flynn, P. F., Kintanar, A., Reid, B. R., & Drobny, G. P. (1988) *Biochemistry 27*, 1191.

Garegg, P. J., Lindh, I., Regberg, T., Stawinski, J., & Stromberg, R. (1986) *Tetrahedron Lett. 34*, 4055.

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

Haasnoot, C. A. G., Westerink, H. P., van der Marel, G. A., & van Boom, J. H. (1984) *J. Biomol. Struct. Dyn. 2*, 345.

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.

Hare, D. R., Shapiro, L., & Patel, D. J. (1986a) *Biochemistry 25*, 7445.

Hare, D. R., Shapiro, L., & Patel, D. J. (1986b) *Biochemistry 25*, 7456.

Jones, S. S., & Reese, C. B. (1979) *J. Chem. Soc., Perkin Trans. 1*, 2762.

Kempe, T., Chow, F., Sundquist, W. I., Nardi, T. J., Paulson, B., & Peterson, S. M. (1982) *Nucleic Acids Res. 10*, 6695.

Kierzek, R., Caruthers, M. H., Longfellow, C. E., Swinton, D., Turner, D. H., & Freier, S. M. (1986) *Biochemistry 25*, 7840.

Kintanar A., Klevit, R. E., & Reid, B. R. (1987) *Nucleic Acids Res. 15*, 5845.

Matteucci, M. D., & Caruthers, M. H. (1980) *Tetrahedron Lett. 21*, 719.

Millane, R. P., Walker, J. K., Arnott, S., Chandrasekaran, R., Birdsall, D. L., & Ratliff, R. L. (1984) *Nucleic Acid Res. 12*, 5475.

Milligan, J. F., Groebe, D. R., Witherell, G. W., & Uhlenbeck, O. C. (1987) *Nucleic Acids Res. 15*, 8783.

Nelson, H. C. M., Finch, J. T., Luisi, B. F., & Klug, A. (1987) *Nature 330*, 221.

Newton, C. R., Greene, A. R., Heathcliffe, G. R., Atkinson, T. C., Holland, D., Markham, A. F., & Ege, M. D. (1983) *Anal. Biochem. 129*, 22.

Nilges, M., Clore, G. M., Gronenborn, A. M., Piel, N., & McLaughlin, L. W. (1987) *Biochemistry 26*, 3734.

Nilsson, L., Clore, G. M., Gronenborn, A. M., Brunger, A. T., & Karplus, M. (1986) *J. Mol. Biol. 188*, 455.

Ogilvie, K. K., & Entwistle, D. W. (1981) *Carbohydr. Res. 89*, 203.

Ogilvie, K. K., Nemer, M. J., & Gillen, M. F. (1984) *Tetrahedron Lett. 25*, 1669.

Patel, D. J., Kozlowski, S. A., Ikuta, S., Itakura, K., Bhatt, R., & Hare, D. R. (1983) *Cold Spring Harbor Symp. Quant. Biol. 47*, 197.

Patel, D. J., Shapiro, L., Kozlowski, S. A., Gaffney, B. L., & Jones, R. A. (1986) *J. Mol. Biol. 188*, 677.

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

Pon, R. T., Usman, N., Damha, M. J., & Ogilvie, K. K. (1986) *Nucleic Acids Res. 14*, 6453.

Reese, C. B. (1985) *Nucleosides Nucleotides 4*, 117.

Reese, C. B., & Skone, P. A. (1985) *Nucleic Acids Res. 14*, 5219.

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., & Imbach, J.-L. (1983a) *Eur. J. Biochem. 135*, 307.

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

Scheek, R. M., Russo, N., Boelens, R., & Kaptein, R. (1983) *J. Am. Chem. Soc. 105*, 2914.

Seliger, H., Gupta, K. C., Kotschi, U., Spaney, T., & Zeh, D. (1986) *Chem. Scr. 26*, 561.

Shakked, A., Rabinovich, D., Kennard, O., Cruse, W. B. T., Salisbury, S. A., & Viswamitra, M. A. (1983) *J. Mol. Biol. 166*, 183.

Sharmeen, L., & Taylor, J. (1987) *Nucleic Acids Res. 15*, 6705.

Shulman, R. G., Hilbers, C. W., Kearns, D. R., Reid, B. R., & Wong, Y. P. (1973) *J. Mol. Biol. 78*, 57.

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

Sung, W. L., & Narang, S. A. (1982) *Can. J. Chem. 60*, 111.

Tanaka, T., Tamatsukuri, S., & Ikehara, M. (1986) *Nucleic Acids Res. 14*, 6265.

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