Brossi, A., Yeh, H. J. C., Chrzanowska, M., Wolff, J., Hamel, E., Lin, C. M., Quinn, F., Suffness, M., & Silverton, J. (1988) *Med. Res. Rev. 8*, 77-94.
Cortese, F., Bhattacharyya, B., & Wolff, J. (1977) *J. Biol. Chem. 252*, 1134-1140.
Deinum, J., Lincoln, P., Larsson, T., Lagercrantz, C., & Erkell, L. J. (1981) *Acta Chem. Scand. B35*, 667-681.
Demas, J. N., & Crosby, G. A. (1971) *J. Phys. Chem. 75*, 991-1024.
Detrich, H. W., III, & Williams, R. C., Jr. (1978) *Biochemistry 17*, 3900-3907.
Detrich, H. W., III, Williams, R. C., Jr., Macdonald, T. L., Wilson, L., & Puett, D. (1981) *Biochemistry 20*, 5999-6005.
Dimroth, K., Reichardt, C., Seipmann, F., & Bohlmann, F. (1963) *Justus Liebigs Ann. Chem. 661*, 1-37.
Dustin, P. (1984) *Microtubules*, 2nd ed., Springer-Verlag, New York.
Engelborghs, Y., & Fitzgerald, T. J. (1987) *J. Biol. Chem. 262*, 5204-5209.
Fernholz, H. (1950) *Justus Liebigs Ann. Chem. 568*, 63-72.
Fitzgerald, T. J. (1976) *Biochem. Pharmacol. 25*, 1383-1387.
Garland, D. L. (1978) *Biochemistry 17*, 4266-4271.
Gratton, E., Jameson, D. M., & Hall, R. D. (1984) *Annu. Rev. Biophys. Bioeng. 13*, 105-124.
Hastie, S. B., & Rava, R. P. (1989) *J. Am. Chem. Soc. 111*, 6993-7001.
Hastie, S. B., Williams, R. C., Jr., Puett, D., & Macdonald, T. L. (1989) *J. Biol. Chem. 264*, 6682-6688.
Head, J., Lee, L. L.-Y., Field, D. J., & Lee, J. C. (1985) *J. Biol. Chem. 260*, 11060-11066.
Hrbek, J., Jr., Hruban, L., Simanek, V., Santavy, F., Snatzke, G., & Yemul, S. S. (1982) *Collect. Czech. Chem. Commun. 47*, 2258-2279.
Lakowicz, J. R. (1983) *Principles of Fluorescence Spectroscopy*, Plenum Press, New York.
Lakowicz, J. R., Laczko, G., Chereck, H., Gratton, E., & Limkeman, M. (1984) *Biophys. J. 46*, 463-477.
Lambeir, A., & Engelborghs, Y. (1981) *J. Biol. Chem. 256*, 3279-3282.
Margulis, T. N., & Lessinger, L. (1978) *Biochem. Biophys. Res. Commun. 83*, 472-478.
McRae, E. G., & Kasha, M. (1958) *J. Chem. Phys. 28*, 721-742.
Medrano, F. J., Andreu, J. M., Gorbunoff, M. J., & Timasheff, S. N. (1989) *Biochemistry 28*, 5589-5599.
Penefsky, H. A. (1971) *J. Biol. Chem. 252*, 2891-2899.
Rao, C. N. R. (1975) *Ultraviolet and Visible Spectroscopy*, 3rd ed., pp 60-79, 118-127, Butterworths, London.
Rava, R. P., Hastie, S. B., & Myslik, J. C. (1987) *J. Am. Chem. Soc. 109*, 2202-2203.
Ray, K., Bhattacharyya, B., & Biswas, B. B. (1981) *J. Biol. Chem. 256*, 6241-6244.
Stryer, L. (1978) *Annu. Rev. Biochem. 47*, 819-846.
Ward, L. D., & Timasheff, S. N. (1988) *Biochemistry 27*, 1508-1514.
Williams, R. C., Jr., & Lee, J. C. (1982) *Methods Enzymol. 85*, 376-385.
Yeh, H. J. C., Chrzanowska, M., & Brossi, A. (1988) *FEBS Lett. 229*, 82-86.
Zavala, F., Guenard, D., Robin, J. P., & Brown, E. (1978) *J. Med. Chem. 23*, 546-549.

Conformation and Dynamics of an RNA Internal Loop†

Gabriele Varani, Brian Wimberly, and Ignacio Tinoco, Jr.*
Department of Chemistry and Laboratory of Chemical Biodynamics, University of California, Berkeley, Berkeley, California 94720

Received April 10, 1989; Revised Manuscript Received June 1, 1989

ABSTRACT: The conformation and the dynamics of an RNA oligonucleotide (26 nucleotides) which is a model for loop E in eukaryotic 5S RNA have been investigated by one- and two-dimensional NMR. The central portion of the oligonucleotide contains two G A oppositions, a common feature of ribosomal RNAs. The exchangeable proton spectrum indicates that an internal loop separates two stems of four and five base pairs. This observation is not consistent with structures for loop E containing mismatched G·A base pairs proposed from chemical and enzymatic studies on *Xenopus laevis* 5S RNA. The nonexchangeable proton spectrum has been assigned by two-dimensional NMR. Scalar couplings from correlated experiments and interproton distances from NOESY experiments at short mixing times have been used to determine glycosidic angles, sugar puckers, and other conformational features. The conformation of the stems is very close to standard A-form RNA, and extensive base stacking continues into the internal loop. This result provides a structural basis for the large favorable enthalpy of duplex formation determined in thermodynamic studies. Unusual structural and dynamic features are localized in the nucleotides connecting the loop to the stems.

The three-dimensional structure of RNA consists of a few fundamental motifs and their interactions. Secondary structure motifs are double-helical stems, hairpin and internal loops, bulges, and mismatched base pairs. Tertiary motifs include multistem junctions, base triples, and pseudoknots. Understanding the structure, the dynamics, and the thermodynamic stability of the individual motifs will help in understanding the principles of RNA folding, the specificity of RNA–protein interactions, and the catalytic abilities of RNA.

Most information on RNA structure at the atomic detail was provided by crystallographic studies on tRNA (Saenger, 1983). However, some of the motifs of RNA structure, for example, internal loops, are not present in tRNA. Thus,

† This work was supported in part by National Institutes of Health Grant GM 10840 and by Department of Energy, Office of Energy Research, Office of Health and Environmental Research, Grant DE-FG03-86ER60406.

0006-2960/89/0428-7760$01.50/0

relatively little is known about them. The absence of detailed structural information about internal loops prevents their realistic incorporation in models of RNA. For example, double-helical tracts are simply connected by straight lines in models for the three-dimensional folding of 16S ribosomal RNA (Stern et al., 1988). However, some ribosomal proteins specifically recognize internal loops (Leffers et al., 1988; Gregory et al., 1988), and the nucleotide sequence of internal loops in ribosomal RNAs is frequently conserved (Noller, 1984). Most internal loops in ribosomal RNA are purine rich, and mismatched base pairs can be formed between opposing strands; G A oppositions are particularly common (Noller, 1984).

To date, most of the structural information on RNA internal loops and mismatches comes from chemical and enzymatic studies. The best known is loop E in 5S RNA, part of the binding site for both ribosomal protein L5 and transcription factor IIIA in eukaryotes. The structure of both prokaryotic and eukaryotic loop E has been probed enzymatically (Andersen et al., 1987a,b), by a combination of both chemical and enzymatic probes (Romaniuk et al., 1987, 1988; Romby et al., 1988) and by exchangeable proton nuclear magnetic resonance (NMR)[1] (Gewirth et al, 1987; Li et al., 1987). The general features of the secondary structure model for 5S RNA have been confirmed, but details of the local structure remain largely obscure: three different models have been proposed for loop E of *Xenopus laevis* 5S RNA (Andersen et al., 1984a,b; Christiansen et al., 1987; Romaniuk et al., 1987, 1988). As both the rate and the extent of reaction of an RNA site to a particular probe depend on the RNA conformation and on the experimental conditions, the conclusions may depend on the conditions and the probe. Furthermore, enzymatic probes such as S_1 and V_1 nucleases may bind preferentially to one of several possible alternate conformations of similar free energy and thus induce conformational transitions when they bind to RNA. In general, chemical and enzymatic probes give very useful measures of differences in reactivity (and thus differences in conformation) of an RNA, but they cannot provide a detailed picture of the local geometry of the RNA.

Detailed information on the local conformation of nucleic acids can be provided by nonexchangeable proton NMR using two-dimensional techniques. The major limit of this technique is the size of the oligonucleotides amenable to NMR studies (less than about 30 nucleotides or approximately 10000 molecular weight). Thus, it is impossible to study intact RNA molecules, not even tRNA or 5S RNA; only selected structural motifs can be studied. NMR is an excellent technique for the characterization of the basic structural motifs in nucleic acids.

So far, the overwhelming majority of nucleic acid NMR studies have been devoted to deoxyoligonucleotides; this has been mainly due to the more difficult synthesis of ribooligonucleotides. New enzymatic and chemical techniques are now available for efficiently synthesizing large amounts (more than 10 mg) of RNA oligonucleotides. Furthermore, it has been generally believed that only very small ribooligonucleotides could be analyzed by nonexchangeable proton NMR, because of the crowded ribose proton spectrum. Here we show that the proton spectrum of large RNA oligonucleotides can be assigned and analyzed despite the overlap of the sugar proton resonances. We think that even RNAs with 40 nucleotides can be studied by nonexchangeable proton NMR in favorable cases. The geometry and the dynamics of RNA structural motifs in solution can then be studied at the atomic detail. The results of chemical and enzymatic probing can be further interpreted with the help of the NMR studies, and a better understanding of the structural target of these probes can be gained.

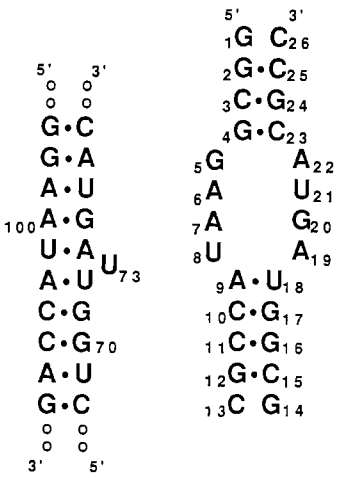


FIGURE 1: Nucleotide sequence of the duplex oligonucleotide investigated here (right) is compared with loop E in *X. laevis* 5S RNA (left). Base pairing and numbering for loop E are from Andersen et al. (1984a,b).

In the present paper, the structure and the dynamics of an RNA oligonucleotide with a sequence similar to that of loop E of eukaryotic 5S RNA have been studied by one- and two-dimensional NMR. Two oligonucleotides (13-mers) were synthesized enzymatically by using T7 RNA polymerase. The resulting duplex is shown in Figure 1 and compared with loop E of *X. laevis* 5S RNA. The sequence was chosen to maximize the possibility of forming G·A base pairs; therefore, the U that would thus result in a bulge (U_{73} in *X. laevis*) was omitted. The base pairs at each end of the internal loop are the same as in *X. laevis*, but other pairs in the stems were chosen to improve duplex stability and efficiency of synthesis. The purpose of this study is to investigate the details of the structure and the dynamics of a model purine-rich internal loop, and to study the stability of mismatched G·A base pairs in RNA. Structure and dynamics of the stems are also of interest. Since long double-helical tracts are rare in RNA, comparison with long A-form RNA helices studied by fiber diffraction is important.

Materials and Methods

RNA Synthesis and Purification. Oligoribonucleotides were synthesized by using T7 RNA polymerase and synthetic DNA templates following the procedure of Milligan et al. (1987). DNA templates were synthesized on an Applied Biosystem 381A and purified by 20% polyacrylamide gel electrophoresis under denaturing conditions (7 M urea). The following DNA templates were used:

5′-T-A-A-T-A-C-G-A-C-T-C-A-C-T-A-T-A-G-3′

3′-A-T-T-A-T-G-C-T-G-A-G-T-G-A-T-A-T-C-C-G-C-C-T-T-A-T-G-G-C-G-5′

(strand 1, G_1–C_{13}) and

5′-T-A-A-T-A-C-G-A-C-T-C-A-C-T-A-T-A-G-3′

3′-A-T-T-A-T-G-C-T-G-A-G-T-G-A-T-A-T-C-G-C-C-A-T-C-A-T-G-C-G-G-5′

(strand 2, G_{14}–C_{26}).

Oligoribonucleotides were phenol extracted and purified by 20% polyacrylamide gel electrophoresis. RNA was eluted from

[1] Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser effect spectroscopy; 2QF-COSY, two-quantum-filtered correlated spectroscopy; TFIIIA, transcription factor IIIA.

the gel by using a Schleicher & Schuell electroelution apparatus and subsequently ethanol precipitated twice. Multivalent ions were removed by dialyzing extensively against a high EDTA buffer (10 mM phosphate/5 mM EDTA, pH 7) and then passing the solution through the chelating agent Chelex-100 (Bio-Rad, Richmond, CA). Finally, oligonucleotides were dialyzed several hours against 10 mM phosphate/1 mM EDTA (pH 7). Typical yields were 3–5 ODU of purified RNA/mL of reaction; several 10-mL reactions yielded more than 40 mg of purified RNA.

The fidelity of transcription was checked by partial digestion of ^{32}P 5′-labeled oligonucleotides with ribonucleases A (specific for C and U), T_1 (G), and U_2 (A) (Donis-Keller et al., 1977). 5′-Terminal phosphates were removed with calf intestinal phosphatase (Boehringer-Mannheim), and re-added by incubation with T4 polynucleotide kinase (USB) and radioactive ATP. Transcription reactions without GMP as a primer yield homogeneous molecules; ^{32}P-labeled oligonucleotides migrate as a single band on gels after 5′-phosphates are removed. When GMP is added to the reaction mixture to improve the yield of large-scale syntheses, oligonucleotides migrate in preparative gels as partially resolved doublets of approximately equimolar species differing in their 5′-terminal phosphates. Terminal monophosphates can be removed by incubation with calf intestinal phosphatase. However, this procedure would lead to loss of about 50% of the oligonucleotides as a consequence of Mg^{2+}-catalyzed hydrolysis.

Thermodynamics. Absorbance versus temperature profiles (melts) were recorded at 260 nm on a thermoelectrically controlled Gilford 250 spectrophotometer interfaced with an Apple IIE computer. The temperature was increased at a rate of 1 °C/min; control experiments at lower heating rates yielded identical melting curves. The samples were dialyzed 24 h against 10 mM phosphate, 0.1 mM EDTA, and 1 M NaCl (pH 7) and preheated at 90 °C before each measurement. Cells of different path length (0.07–10 mm) were used to allow measurements over a 200-fold range of nucleotide concentrations (≈0.01 to ≈1.7 mM). Oligonucleotide concentrations were determined from high-temperature absorbances using the nearest-neighbor approximation to estimate extinction coefficients. Thermodynamic parameters for helix formation were calculated by assuming a two-state model for the single strand to duplex transition. The equilibrium constant is

$$K = 2f/[(1-f)^2 C_t]$$

where f is the fraction of duplex and C_t the total oligonucleotide concentration. Standard techniques of extrapolation of upper and lower base lines were employed. Standard enthalpies and entropies were calculated from the concentration dependence of the melting temperature.

Size Exclusion Chromatography. Size exclusion chromatography was done at flow rate of 0.5 mL/min on a Bio-Rad TSK-125 column preequilibrated with the appropriate buffer. Samples were heat-denatured and allowed to equilibrate at room temperature for several hours before injection.

NMR. Proton NMR experiments were recorded at 500 MHz on a GE GN-500 spectrometer; phosphorus spectra (at 202 MHz) were also done on the GN-500 spectrometer. Most NMR experiments were done on one sample prepared by dialyzing overnight against 10 mM phosphate, 100 mM NaCl, and 1 mM EDTA (pH 7). Some two-dimensional experiments were repeated at slightly different salt concentration (150 mM NaCl) or temperature (40 °C) to help resolve overlapping cross-peaks. Samples used in exchangeable proton experiments were prepared by adding a 10:1 mixture of H_2O and D_2O to dry RNA. For nonexchangeable proton experiments, RNA was lyophilized several times with 99.8% D_2O (Cambridge Isotopes); then, approximately 15 mg of RNA was dissolved in 0.4 mL of 99.96% D_2O (Aldrich) to a final strand concentration of ≈3.5 mM. The sample was not removed from the tube (Aldrich) over a period of a few months; no appreciable degradation was detected.

Exchangeable proton spectra were recorded by using the 1331 solvent suppression scheme (Hore, 1983) with 10000-Hz sweep width. One-dimensional NOE experiments were performed with a preirradiation time of 600 ms. Three experiments with different decoupler powers were used to help distinguish true NOE's from spillover artifacts.

2D NMR spectra were recorded in the phase-sensitive mode using the TPPI method (Marion & Wüthrich, 1983) and preirradiation at low decoupler power of the residual HDO peak. A total of 400–450 FID's were collected, and 2K complex data points were recorded in t_2. The sweep width was 4000 Hz. The repetition delay was set to ≈2 s, and 80–96 scans were averaged for each FID; the total acquisition time ≤24 h. Data analysis was done on a μVAX computer using Dr. D. Hare's software. Data were zero-filled to 1K real-points in t_1 and apodized prior to Fourier transformation using phase-shifted skewed sine bells in both dimensions; the skewedness was 0.7. The same data were apodized with different phase shifts to yield spectra with improved resolution (15–30° phase shift) or signal/noise (40–80° phase shift).

NOESY spectra were recorded by using the experimental scheme of Macura and Ernst (1979) at different mixing times (60, 100, 150, 300, and 500 ms). Multiple-quantum contributions were eliminated by incrementing the mixing time systematically together with the evolution time t_1 (Macura et al., 1982). 2QF-COSY spectra were collected by using the pulse sequence of Müller et al. (1986); composite mixing pulses were employed to reduce pulse inhomogeneities (Müller et al., 1986). Double-quantum spectra were recorded by using the sequence (Braunschweiler et al., 1983)

$$(\pi/2)_x - \tau/2 - (\pi)_x - \tau/2 - (\pi/2)_x - t_1 - (\pi/2)_\phi - t_2$$

The phase (ϕ) of the mixing pulse was cycled according to $x,y,-x,-y$, with alternate addition and subtraction of the corresponding FID's (Braunschweiler et al., 1983); CYCLOPS was also incorporated for a basic cycle of 16 scans. The excitation delay (τ) was set to 50 ms. Composite pulses were used for the mixing pulse and the π pulse in the middle of the excitation sequence.

Interproton Distances. Interproton distances were evaluated from the cross-peak intensities in NOESY experiments at short mixing times (60–150 ms). Cross-peak volume integration was done by using Dr. D. Hare's software. H5–H6 distances in pyrimidines (2.45 Å) were used as references; only H5–H6 cross-peak intensities within the same base pair or immediately adjacent were used.

The effect of apodization was estimated by comparing distances from spectra apodized with sine bells of different phase shift. Apodization effects are smaller than the experimental uncertainty. Obviously, higher resolution spectra are superior for crowded spectral regions.

It is well-known that spin diffusion can lead to incorrect interproton distances. Whenever the signal/noise ratio was sufficient for reliable integration, only data at the shortest mixing times (60 and 100 ms) were used. The importance of spin diffusion was estimated by comparing distances evaluated from experiments at different mixing times. Mixing times of 60–150 ms appear to be short enough to estimate within ±0.5 Å distances between 3 and 4.5 Å. This is consistent with the results for deoxyoligonucleotides. One exception is represented

by the internucleotide H1′(i)–H8/H6(i + 1) distance, due to the very favorable spin diffusion pathway H1′(i)–H2′(i)–H8/H6(i + 1) characteristic of A-form helices (see below). Distances shorter than 3 Å can be obtained from the shortest mixing time NOESY (60 ms), with the exception of H2′(i)–H8/H6(i + 1). Shorter mixing times (≤40 ms) are probably necessary to yield reliable distances when protons are closer than 2–2.3 Å.

A second problem is internal motion. It has been suggested that the correlation time of sugar protons may be faster than that of base protons (Gronenborn & Clore, 1985). Since the H1′–H3′ and H1′–H2′ distances are not very sensitive to conformation (Wüthrich, 1986), they can be used to investigate whether the same correlation time (and consequently the same reference distance) can be used for both base and sugar protons. H1′–H2′ and H1′–H3′ distances were evaluated by using H5–H6 distances as reference. All H1′–H2′ and H1′–H3′ distances were found within the range defined by covalent constraints, 2.4–2.9 and 3.6–4 Å, respectively. This suggests that base and sugar protons experience similar correlation times and the same reference distance (H5–H6) can be used for all protons in this molecule. It is more difficult to estimate the relevance of other internal motions, since little is known about nucleic acids' internal dynamics in solution. The conformational purity of the ribose moiety (see below) is encouraging, but other motions might be present. The $1/r^6$ dependence of the intensity of NOESY cross-peaks will skew all distance estimates toward shorter values. We decided to interpret distances only qualitatively, similarly to what is usually done in protein studies (Wüthrich, 1986).

Imino Proton Exchange. The kinetics of exchange of imino protons with the solvent were studied following the procedure of Leroy et al. (1985a). Approximately 10 mg of RNA was dialyzed against 100 mM NaCl, 1 mM EDTA, and 10 mM sodium cacodylate. All buffers were passed through Chelex-100 to remove multivalent paramagnetic ions. The pH was adjusted to 7.45 in the NMR tube and checked throughout the experiment with a microelectrode; when Tris was added as an exchange catalyst, the Tris proton resonance was followed to monitor pH variations. The intrinsic exchange rate of imino protons in nucleic acids is independent of pH or catalyst, between pH 7 and 9 (Leroy et al., 1985a). The choice of pH 7.45 and Tris catalyst was a compromise between two different necessities. One would like to keep the pH as low as possible to minimize RNA hydrolysis, but higher pH favors exchange by increasing the catalyst concentration. Among the catalysts suggested by Leroy et al. (1985a), Tris (pK = 8.18) better matches these requirements. However, the buffer pH (7.45) is well below the catalyst pK. Consequently, high quantities of Tris (≥0.2–0.3 M) are required to induce appreciable broadening of some resonances.

The Tris concentration, [B], was increased in small increments up to 0.8 M by adding small amounts of a concentrated Tris solution (2 M), whose pH was also adjusted to 7.45. Dilution effects were taken into account in evaluating the catalyst concentration, [C]:

$$[C] = [B]/(1 + 10^{pK-pH})$$

Following Leroy et al. (1985a), the catalyst-induced exchange rate is $k_{ex} = 1/\tau_{ex} = 1/\pi\Delta$, where Δ is the excess line width at half-height induced by the catalyst. Line widths were measured from spectra recorded by using the 1331 solvent suppression scheme; base-line distortion effects were negligible in the region of interest. A 3-Hz line broadening contribution was introduced to improve the signal/noise ratio. The range of exchange times accessible with this technique is approximately 3–150 ms. Plots of exchange times, τ_{ex}, vs inverse catalyst concentration, 1/[C], yield straight lines of intercept τ_o, the intrinsic base pair opening lifetime (Leroy et al., 1985a). Activation enthalpies for base pair opening were determined from Arrhenius plots of the intrinsic lifetime τ_o vs $1/T$.

Results

Thermodynamics. Absorbance vs temperature melting data were analyzed with a two-state model to obtain thermodynamic parameters for duplex formation in 1 M NaCl. A standard free energy of formation at 37 °C of $\Delta G°(37\ °C) = -17.7 \pm 1$ kcal mol^{-1}, a standard enthalpy of $\Delta H° = -127 \pm 14$ kcal mol^{-1}, and a standard entropy of -352 ± 40 cal mol^{-1} K^{-1} were obtained; the melting temperature at 400 μM total strand concentration was 70 °C.

Measured thermodynamic quantities can be compared with estimates for an internal loop or a fully paired helix. An internal loop of six nucleotides is calculated (Freier et al., 1986) to have $\Delta G°(37\ °C) = -17.3$ kcal mol^{-1}. However, the imino proton spectra indicate eight bases have rapidly exchanging imino protons (see below); therefore, we also calculated an internal loop of eight nucleotides; $\Delta G°(37\ °C) = -15.9$ kcal mol^{-1}. The comparison indicates that the duplex is about as stable at 37 °C as expected for an internal loop of six nucleotides, but significantly more stable than expected for a loop of eight nucleotides. A reasonable upper estimate for the duplex stability can be obtained by replacing the two G·A mismatches by U·A base pairs. This gives calculated (Freier et al., 1986) values of $\Delta G°(37\ °C) = -21.4$ kcal mol^{-1} and $\Delta H° = -117.7$ kcal mol^{-1}. Thus, the duplex molecule is not as stable at 37 °C as a fully Watson–Crick-paired helix, but its enthalpy of formation is slightly more favorable. Even if the two G·A mismatches are replaced by G·C pairs, the calculated enthalpy of the duplex is only $\Delta H° = -130.4$ kcal mol^{-1}. The large negative $\Delta H°$ measured for the duplex implies very favorable stacking (relative to the single strand) in the loop.

Exchangeable Protons. The low-field section of the exchangeable NMR proton spectrum is shown in Figure 2A; spectra in the presence of $MgCl_2$ (5 mM $MgCl_2$/100 mM NaCl) and high NaCl concentration (1.15 M) are also shown. Spectra at different temperatures are reported in Figure 2B. Below 40 °C, seven major resonances (labeled in Figure 2A) and a few less intense peaks (a–c) are present. Peaks a–c are due to a small population of single-stranded species in equilibrium with the duplex. Size exclusion chromatography reveals the presence of 5–10% lower molecular weight species at NMR concentrations. The intensity of the smaller peaks in the spectra of Figure 2A,B is consistent with an estimated 90–95% duplex. The observation of imino proton resonances for the single strands indicates the formation of hairpin loops.

The assignment of the exchangeable proton spectrum was accomplished by one-dimensional NOE experiments. The most downfield resonance was assigned to an A·U base pair from its chemical shift and the strong NOE to a sharp nonexchangeable proton at 7.11 ppm (AH2). All other imino resonances have one strong NOE to an exchangeable proton between 8 and 8.5 ppm, where hydrogen-bonded C amino protons are usually found (Fazakerley et al., 1984). Most of them have a second, weaker NOE to another exchangeable proton between 6.5 and 7 ppm, where non-hydrogen-bonded C amino protons resonate (Fazakerley et al., 1984). All resonances except the first should be assigned to G·C pairs.

Of the possible A·U pairs, A_9·U_{18} is the most likely to form, and a single NOE is observed between the leftmost resonance and a G imino proton. The leftmost resonance can then be

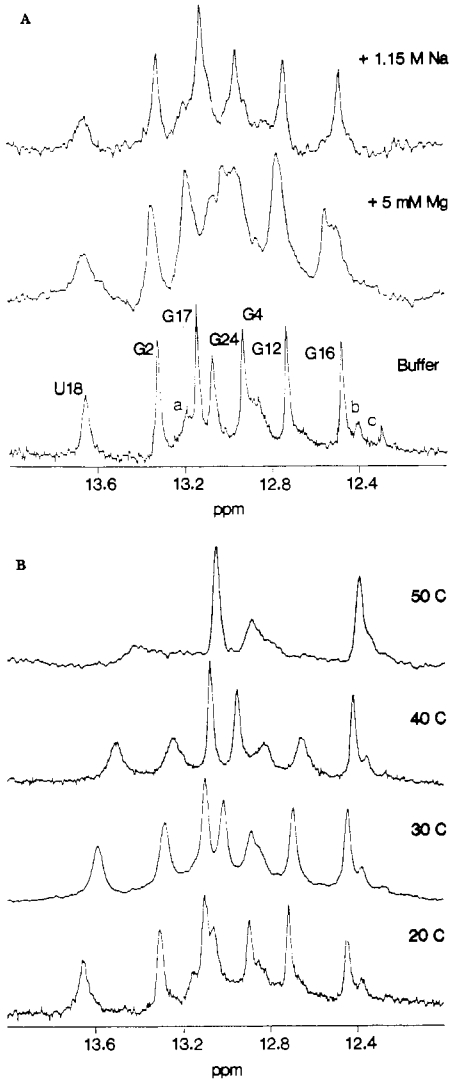


FIGURE 2: Low-field section of the exchangeable proton spectrum at 500 MHz in different solvents (A) and at different temperatures (B). No filter was applied to the spectra of (A) to show the broadening induced by high NaCl concentrations and by Mg^{2+}. 1-Hz line broadening was applied to the spectra in (B). Spectra in (A) are at 30 °C; from the bottom, the solvent is 10 mM phosphate buffer (pH 7) plus 150 mM NaCl, buffer plus 5 mM $MgCl_2$ plus 100 mM NaCl, and buffer plus 1.15 M NaCl, respectively. Major peaks are labeled; peaks a–c are from minor hairpin species. Spectra in (B) are in 100 mM NaCl/10 mM phosphate at pH 7; from the bottom, temperatures are 20, 30, 40, and 50 °C, respectively. Chemical shifts are relative to TSP.

assigned to U_{18}. From there, it is possible to assign sequentially the imino protons of G_{17}, G_{16}, and G_{12} using imino to imino NOE's. Of the three remaining peaks, one resonance (G_{24}) has NOE's to two other resonances (G_2 and G_4), whereas no NOE was observed between the other two. The two remaining resonances can be assigned either to G_2 or to G_4 and cannot be distinguished on the basis of NOE's alone. C amino protons were assigned from the NOE's from the corresponding G imino. The downfield resonance (8–8.5 ppm) was assigned to the hydrogen-bonded amino proton and the upfield resonance (6.5–7 ppm) to the non-hydrogen-bonded one.

The disappearance of imino resonances at increasing temperature (Figure 2B) is consistent with the proposed assignments. As observed in deoxyoligonucleotides, imino resonances shift upfield at increasing temperature before disappearing when exchange with the solvent becomes fast. Terminal base pairs do not contribute to the imino spectrum above ≈20 °C due to end-fraying. One resonance, G_4, broadens significantly

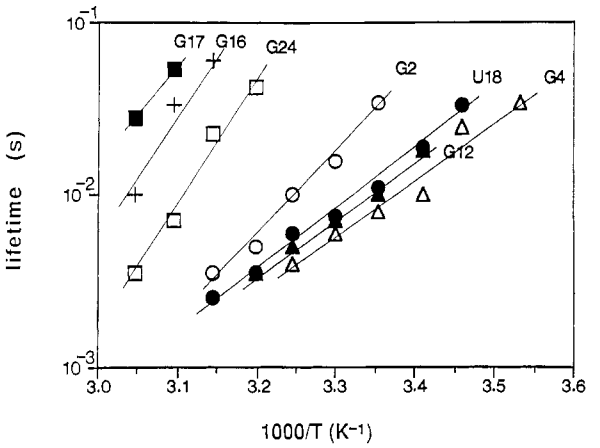


FIGURE 3: Arrhenius plot for base pair opening. Intrinsic base opening lifetimes are reported vs inverse absolute temperature. Each line is labeled as in Figure 2A. Experimental uncertainty is estimated to range between 1 ms for lifetimes shorter than 10 ms to approximately 10 ms for the longest intrinsic opening lifetimes.

at 35 °C and disappears above 45 °C. Peaks corresponding to U_{18}, G_{12}, and G_2 broaden at somewhat higher temperature (40 °C). The G_2 resonance was assigned on the basis of its solvent exchange behavior, analogous to G_{12} which occupies the same position close to the end of the duplex. Resonances from imino protons in the core of the longer stem are still relatively sharp at 55 °C, close to the absorbance melting temperature (≈70 °C).

The spectrum at lower pH and in the absence of phosphates (1 mM acetate, pH 5.8) is essentially identical with that at neutral pH. The main difference is a very broad shoulder at 11 ppm observable only below 10 °C. Spectra in Mg^{2+} and in 1.15 M NaCl (Figure 2A) are also similar to the low-salt spectra; in particular, no new resonance is observed above 13 ppm, where A·U base pairs resonate. In the presence of Mg^{2+}, the peaks are somewhat broader and minor-form peaks more intense. Broadening is likely due to nonspecific aggregation, which is indeed observed at lower temperatures in the presence or absence of Mg^{2+}. However, the U_{18} imino resonance is specifically broadened by Mg^{2+} and by high NaCl concentrations.

Imino Proton Exchange. Imino proton exchange lifetimes were measured as described under Materials and Methods. As predicted by Leroy et al. (1985a), plots of exchange lifetime τ_{ex} vs inverse catalyst concentration 1/[C] yield straight lines. The exception is represented by some data at high Tris concentration. The intercepts yield intrinsic base opening lifetimes (τ_0) and are very reproducible over different experiments. Slopes are related to the dissociation constant for base opening, the pK of individual nucleotides, and the accessibility of the catalyst to the bases (Leroy et al., 1985a). Slopes are less reproducible and were not analyzed quantitatively. Intrinsic lifetimes τ_0 are plotted in Figure 3 vs inverse absolute temperature $1/T$ to obtain activation enthalpies for base pair opening.

Two groups of protons can be distinguished in Figure 3. Three protons in the interior of the stems have long opening lifetimes ($\tau_0 > 50$ ms at 40 °C) and high activation enthalpies (≈35 kcal/mol). The two protons in the longer stem, G_{16} and G_{17}, have slightly longer lifetimes than G_{24} in the shorter stem. The protons in the more rapidly exchanging group are located next to the loop (G_4, U_{18}) or close to the end of the molecule (G_2, G_{12}). Their activation enthalpies are between 15 and 20 kcal/mol, considerably lower than those of the first group of protons. Their open-state lifetimes at 40 °C range between

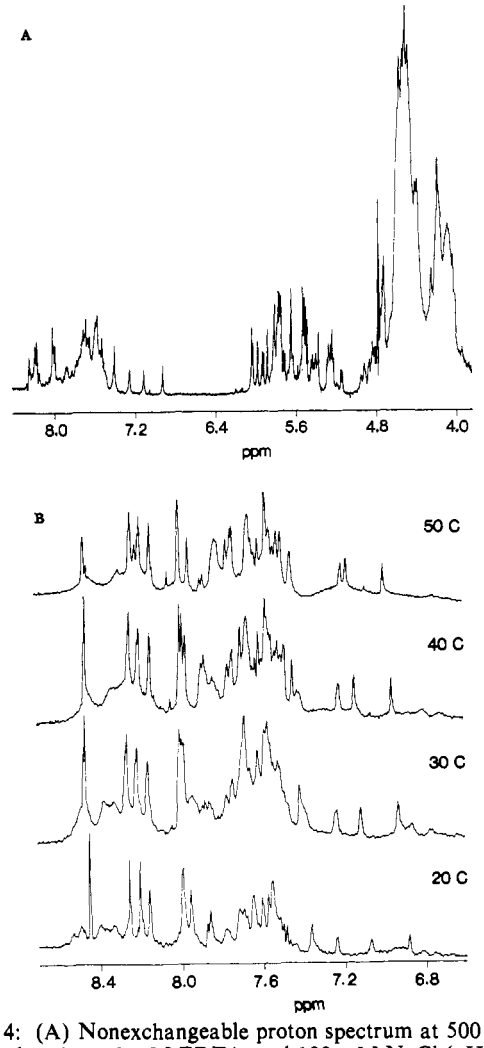


FIGURE 4: (A) Nonexchangeable proton spectrum at 500 MHz in 10 mM phosphate, 1 mM EDTA, and 100 mM NaCl (pH 7) at 30 °C. 0.25-Hz line broadening was applied before Fourier transformation. (B) Aromatic and amino section of the proton spectrum at different temperatures in the same buffer of (A). Spectra were recorded by using the 1331 solvent suppression scheme and a 10000-Hz sweep width; 0.25-Hz line broadening was applied before Fourier transformation. From the bottom, temperatures are 20, 30, 40, and 50 °C, respectively. Chemical shifts are relative to TSP.

3 and 6 ms, G_2 and G_{12} having the larger values. Activation enthalpies are very similar within each group, with the exception of U_{18} which has the lowest activation enthalpy for base opening.

Phosphorus. All ^{31}P peaks in the spectrum at 202 MHz resonate between −3.5 and −5 ppm, as expected in regular A- or B-form helices. The small chemical shift dispersion of ^{31}P resonances rules out nonstandard conformations (such as observed in Z-DNA and in tRNA crystals) for the backbone angles immediately adjacent to the phosphorus (Gorenstein, 1984).

Nonexchangeable Protons. The nonexchangeable proton spectrum is shown in Figure 4A. A portion of the proton spectrum containing both aromatic and amino protons is shown in Figure 4B for temperatures between 20 and 50 °C; all broad resonances below 8 ppm and above 7 ppm are from amino protons. Some peaks shift upfield on raising the temperature below melting, but the two rightmost peaks (A_9H2 and $A_{19}H2$) display a significant downfield shift. A few nonexchangeable protons have broad lines; this can be appreciated for the G_5H8 resonance, well isolated at 7.9 ppm in the spectrum at 30 °C. At temperatures above 55 °C, all resonances broaden as melting to single strands occurs. Unfortunately, the overlap

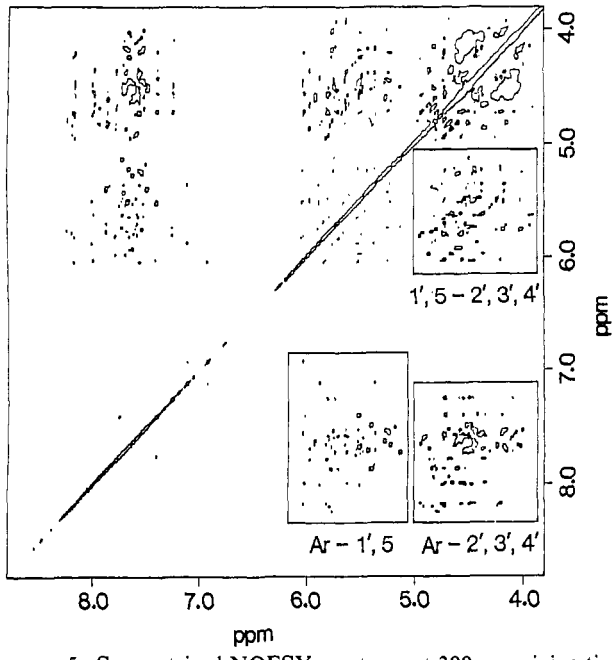


FIGURE 5: Symmetrized NOESY spectrum at 300-ms mixing time in 10 mM phosphate, 1 mM EDTA, and 100 mM NaCl (pH 7) at 30 °C; RNA concentration is 3.5 mM (in strands). The aromatic to H1′–H5 and aromatic to H2′–H3′–H4′ sections are expanded in Figure 6 and Figure 8, respectively.

of most resonances makes it impossible to follow the premelting and the melting of individual protons. Below 15 °C, all resonances broaden at high RNA concentration. Since this effect is less evident at lower RNA concentration, it was attributed to aggregation.

Assignment of the Nonexchangeable Proton Spectrum. The assignment of the nonexchangeable proton spectrum was accomplished following largely the strategy developed for B-form deoxyoligonucleotides, i.e., using the connectivities H8/H6(i) → H1′(i) → H8/H6(i + 1) (Wüthrich, 1986). Unlike in B-DNA, the H1′(i)–H8/H6(i + 1) distance is relatively long in A-RNA (≈4.6 Å). However, the spin diffusion pathway H1′(i) → H2′(i) → H8/H6(i + 1) provides the A-RNA aromatic to H1′ cross-peaks with intensity comparable to the corresponding B-DNA cross-peak (≈3.6 Å) at all mixing times ≥100 ms. The assignment procedure is discussed in detail in the following section. Chemical shifts of all assigned protons are reported in Table I.

A symmetrized NOESY spectrum at 300-ms mixing time is shown in Figure 5. An expansion of the H8/H6–H1′/H5 region of the same nonsymmetrized spectrum is shown in Figure 6; assignment pathways for one of the strands (G_1–C_{13}) are shown. As a first step toward assignment, pyrimidines and purines were distinguished from the cross-peaks observed in 2QF-COSY experiments (Figure 7), and the corresponding, very strong cross-peaks in NOESY spectra. Several minor H5–H6 cross-peaks can be observed in Figure 7. They can be distinguished from major peaks because the intensities of corresponding NOESY cross-peaks are very different. Minor-form NOESY cross-peaks have much weaker intensities due to lower concentration (1:10) and faster correlation time (1:2).

A convenient start for assignments is the A_9H2 resonance, which was identified at 7.11 ppm using the strong NOE from the U_{18} imino (see above). In A-RNA, AH2's lie in the minor groove, close to two H1′ protons on both strands, here $C_{10}H1'$ and $A_{19}H1'$. Cross-peak a in Figure 6 was tentatively assigned to A_9H2–$C_{10}H1'$. Then, C_{10} and C_{11} were identified as one of the only two Py-Py steps from the H5–H5 cross-peak

Table I: Chemical Shifts (ppm) Relative to TSP of Assigned Proton Resonances for the RNA Duplex in 10 mM Phosphate (pH 7)/100 mM NaCl at 30 °C[a]

	H8/6	H2/5	H1′	H2′	H3′	H4′	H5′/5″	imino/amino
5′G_1	8.19	na[b]	5.79	4.92	4.72	4.52	4.24/4.08	
	8.16	na	5.82	4.95	4.75	4.58	4.25/4.07	
G_2	7.61	na	5.92	4.61	4.56	4.26		13.23
	7.65	na	5.93	4.62	4.57	4.27		13.23
C_3	7.62	5.28	5.53	4.47	4.13	4.57		8.25/6.96
G_4	7.51	na	5.65	4.38	4.68			12.93
G_5	7.94	na	5.81	4.52	4.82			
A_6	7.78		5.98	4.87	4.35	4.56	4.18/4.08	
A_7	7.68							
U_8	7.48	5.40						
A_9	8.17	7.11	6.03	4.45	4.74		4.15/4.06	
C_{10}	7.55	5.23	5.37	4.17	4.38			8.41/7.08
C_{11}	7.70	5.43	5.48	4.50	4.52	4.40		8.48
G_{12}	7.59	na	5.75	4.44	4.59			12.76
3′C_{13}	7.53	5.27	5.76	4.01	4.18	4.21		
5′G_{14}	8.19	na	5.78	4.85	4.75	4.55		
	8.23	na	5.77	4.80	4.75	4.50		
C_{15}	7.86	5.39	5.73	4.73	4.60	4.53		8.20
	7.86	5.39	5.70	4.71	4.61	4.52		8.20
G_{16}	7.57	na	5.80	4.73	4.55	4.51		12.49
G_{17}	7.24	na	5.74	4.49	4.47			13.15
U_{18}	7.70	5.13	5.51	4.46				13.71
A_{19}	8.00	6.92	5.88	4.43	4.73	4.60		
G_{20}		na						
U_{21}	7.58	4.95	5.62	4.66	4.35	4.57		
A_{22}	7.99		6.02	4.72	4.56		4.20/4.11	
C_{23}	7.39	5.52						8.53/6.83
G_{24}	7.56	na	5.65	4.45	4.63		4.21/4.07	13.11
C_{25}	7.64	5.23	5.50	4.25		4.39		8.50/6.86
3′C_{26}	7.66	5.52	5.77	4.04	4.19			

[a] Two sets of assignments are reported for some nucleotides; they correspond to molecules with 5′-mono- and 5′-triphosphates. [b] Not applicable.

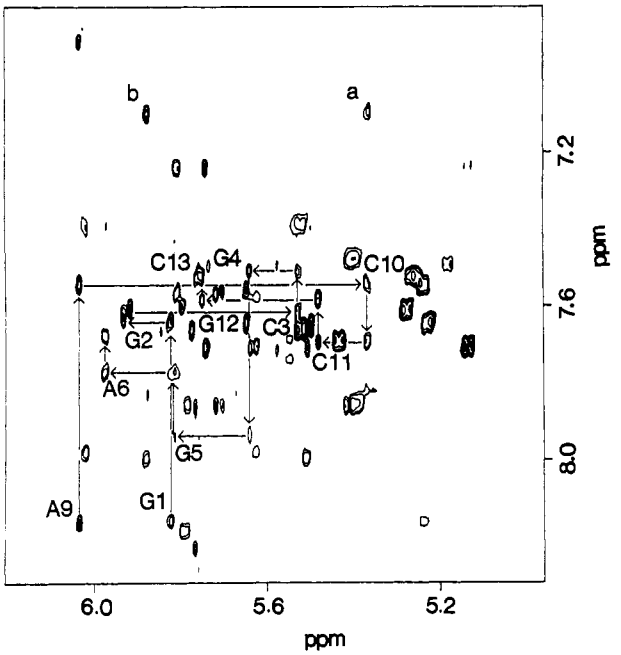


FIGURE 6: Aromatic to H1′–H5 section of the NOESY spectrum at 300-ms mixing time. In order to enhance the resolution, the spectrum was not symmetrized, and the apodization was accomplished by 15° phase-shifted skewed sine bells. Assignment pathways for one of the strands (G_1–C_{13}) are shown; intranucleotide H1′–H8/H6 cross-peaks are labeled. Peaks a and b (intra- and interstrand AH2 to H1′ cross-peaks) represent useful starts for assignments, as discussed in the text.

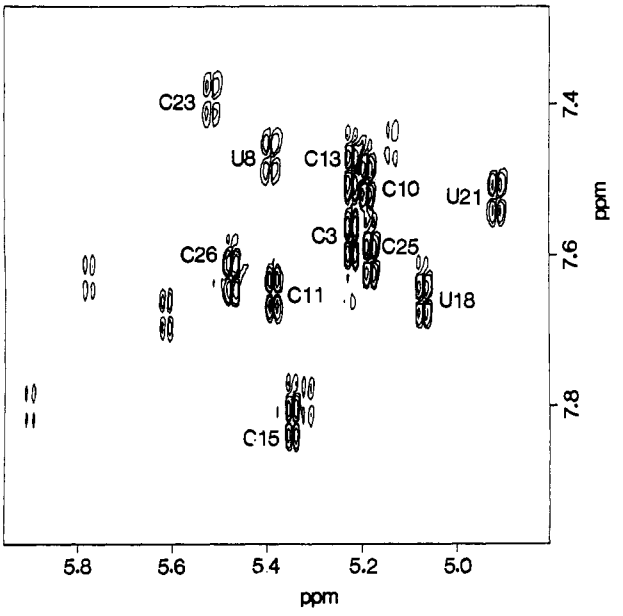


FIGURE 7: Double-quantum-filtered COSY spectrum in 150 mM NaCl, 10 mM phosphate, and 1 mM EDTA (pH 7) at 40 °C. Data were apodized with a 15° phase-shifted skewed sine bell prior to Fourier transformation. Major cross-peaks are labeled; all other peaks are from minor species.

characteristic of A-form helices. This identification leads to the assignment of the five nucleotides A_9–C_{13}. Similarly, cross-peak b was assigned to A_9H2–A_{19}H1′. From there, the connectivity pathway Pu-Py-Pu-Pu-Py-Pu (G_{14}–A_{19}) was established. The assignment of G_{14}H8 was further confirmed by using oligonucleotides selectively deuterated at G_{14}H8. One strand of the shorter stem (G_{24}–C_{26}) was assigned by identifying the second H5–H5 cross-peak (C_{25}–C_{26}). The other side was assigned by locating G_1H8 using selectively deuterated oligonucleotides; the following connectivity is present: Pu-Pu-Py-Pu-Pu-Pu-Pu (G_1–A_7).

Connectivities involving H1′ and aromatic protons for U_8, G_{20}, and C_{23} are missing. However, U_8H6 and C_{23}H6 were identified by using cross-peak to aromatic protons (A_9H8 and G_{24}H8, respectively) observed at very long mixing time (500 ms, data not shown). These connectivities show the partial

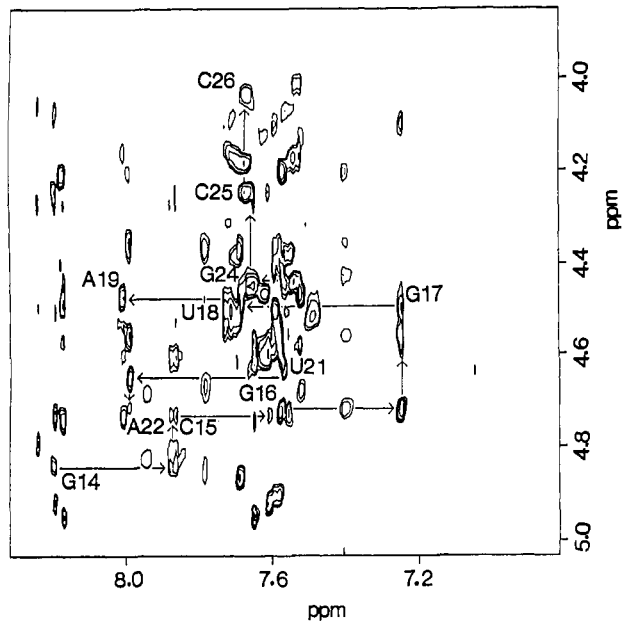


FIGURE 8: Aromatic to H2′–H3′–H4′–H5′–H5″ section of the 300-ms NOESY spectrum. Connectivity pathways for the second strand (G_{14}–C_{26}) are shown. Intranucleotide H2′–H8/H6 cross-peaks are labeled.

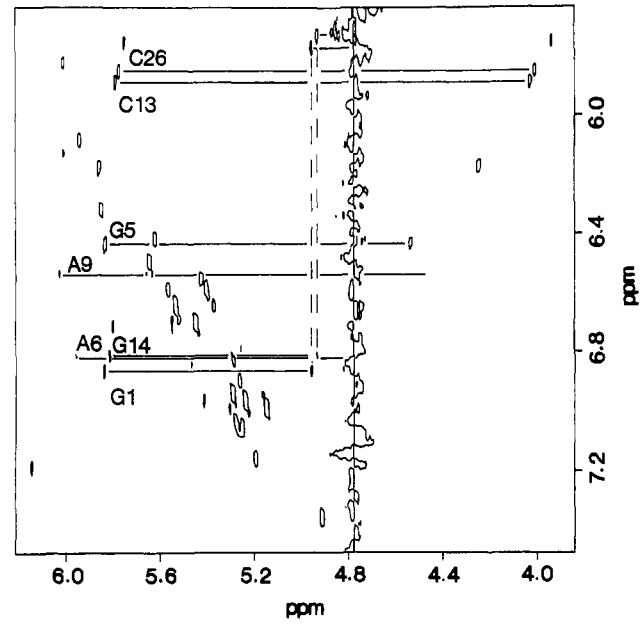


FIGURE 9: H1′ to H2′ section of a two-quantum spectrum in 100 mM NaCl, 10 mM phosphate, and 1 mM EDTA (pH 7) at 30 °C; only positive levels are shown. Data were apodized with a 15° phase-shifted skewed sine bell before Fourier transformation. The horizontal dimension (ω_2) has the same meaning as in conventional one- or two-dimensional spectroscopy. Two-quantum states are detected in the vertical dimension (ω_1); chemical shifts in ω_1 correspond to sums of the offsets from the carrier of the frequencies of the resonances detected in ω_2. The carrier was at 3.9 ppm. H1′–H2′ cross-peaks at ($\omega_{1'}$, $\omega_{1'} + \omega_{2'}$) and ($\omega_{2'}$, $\omega_{1'} + \omega_{2'}$) are joined by horizontal lines; ($\omega_{1'}$, $\omega_{1'} + \omega_{2'}$) cross-peaks are labeled. For G_1 and G_{14}, vertical lines join ($\omega_{2'}$, $\omega_{1'} + \omega_{2'}$) and ($\omega_{2'}$, $\omega_{2'} + \omega_{3'}$) cross-peaks; horizontal lines to ($\omega_{3'}$, $\omega_{2'} + \omega_{3'}$) peaks are also drawn. These two-quantum connectivities have been used to assign sequentially the sugar protons of G_1, A_6, C_{13}, G_{14}, and C_{26}. Nonlabeled peaks on the left of the HDO ridge are folded H5–H6 cross-peaks at (ω_5, $\omega_5 + \omega_6$).

conservation of stacking between the loop and the stems. From C_{23}H6, it is possible to establish the connectivity Py-Pu-Py (U_{21}-A_{22}-C_{23}); the upfield shift of U_{21}H5 (4.95 ppm) is remarkable. It was not possible to identify reliably any of the protons belonging to G_{20}.

H5–H6 COSY cross-peaks for U_8 and C_{23} are weaker than other H5–H6 cross-peaks (Figure 7), and even weaker at lower temperatures. This can be attributed to broadening and cancellation of antiphase components, probably a consequence of exchange between different conformations. H5–H6 NOESY cross-peaks for U_8 and C_{23}, as well as aromatic to sugar proton cross-peaks for G_5, are also broad. U_8, C_{23}, and G_5 are all located at loop–stem junctions.

Resonances belonging to G_1 and G_2 are doublets. Two close but distinct resonances can be observed, each about 50% as intense as other well-resolved resonances. Likewise, resonances belonging to G_{14} and C_{15} are also doublets. This heterogeneity can be attributed to the presence of 5′-terminal mono- and triphosphates (see Materials and Methods).

It is important to have alternative assignment pathways to ensure consistent assignments and resolve ambiguities in H1′–H6/H8 connectivities. Independent connectivity pathways to confirm the assignments are provided by very short H2′(i)–H8/H6(i + 1) distances (≈2 Å in A-form RNA) and longer but still observable H2′(i)–H1′(i + 1) distances (≈4 Å). Unlike DNA, all RNA sugar protons except H1′ resonate in the same spectral range (4–5 ppm). Thus, H2′ protons must be distinguished from other sugar protons. In principle, H2′ resonances could be identified by using the scalar coupling to H1′. Unfortunately, H1′–H2′ cross-peaks are not expected in COSY or two-quantum experiments on RNA because of very small $J_{1',2'}$ (<2 Hz for N sugar pucker (Altona, 1982)). In fact, only a few, weak H1′–H2′ cross-peaks are observed in either 2QF-COSY or two-quantum spectra (see below); this indicates that the sugar pucker is in the N range of conformers (3′-endo) for most residues (Altona, 1982). For N pucker, the intranucleotide H1′–H2′ distance (2.4–2.8 Å) is shorter than the corresponding H1′–H3′ (3.6 Å) and H1′–H4′ (3.0–3.6 Å) (Wüthrich, 1986). Thus, the H2′ protons were identified as the strongest H1′–sugar cross-peaks in low mixing time NOESYs (60–100 ms). It was then possible to establish independent assignment pathways using H2′–H8/H6 connectivities (Figure 8). For some nucleotides, useful internucleotide H2′(i)–H1′(i + 1) and H2′(i)–H5(i + 1) connectivities were also found.

Assignment of the remaining sugar protons (H3′ and H4′) provides useful structural information. A few H3′ and H4′ protons were assigned by using the relatively strong $J_{2',3'}$ (5 Hz) and $J_{3',4'}$ (8 Hz) couplings. Two-quantum spectroscopy was used because there is not a diagonal as in COSY or NOESY experiments; consequently, the sugar–sugar region is less crowded than in conventional correlated spectra. Connectivities for the sequential assignment of sugar protons using two-quantum spectroscopy are shown in Figure 9 for G_1 and G_{14}. Since H1′ protons were assigned as described above, some H2′ protons could be identified by observing two-quantum cross-peaks at ($\omega_{1'}$, $\omega_{1'} + \omega_{2'}$) and ($\omega_{2'}$, $\omega_{1'} + \omega_{2'}$) (Figure 9); horizontal lines connect ($\omega_{1'}$, $\omega_{1'} + \omega_{2'}$) and ($\omega_{2'}$, $\omega_{1'} + \omega_{2'}$) cross-peaks in Figure 9. ($\omega_{2'}$, $\omega_{1'} + \omega_{2'}$) cross-peaks are connected by vertical lines to ($\omega_{2'}$, $\omega_{2'} + \omega_{3'}$) cross-peaks for G_1 and G_{14}; in turn, these are horizontally connected to ($\omega_{3'}$, $\omega_{2'} + \omega_{3'}$). From ($\omega_{3'}$, $\omega_{2'} + \omega_{3'}$) cross-peaks, it is possible to extend the sequential walk to H4′ (not shown). One further advantage of two-quantum spectroscopy is that two-quantum cross-peaks, for instance, ($\omega_{2'}$, $\omega_{2'} + \omega_{3'}$) and ($\omega_{3'}$, $\omega_{2'} + \omega_{3'}$), are in symmetrical positions with respect to the two-quantum diagonal; this symmetry can be used to identify unambiguously cross-peaks very close to the HDO ridge (as for G_1 and G_{14}), or in overlapped regions. Furthermore, each cross-peak contains information on the resonance frequency of both protons;

this is very useful when overlap or noise prevents accurate measurements for one of the cross-peaks.

Since connectivities are generated by scalar couplings, the above procedure is very reliable. However, even two-quantum spectra are too crowded for a reliable analysis in oligonucleotides of this size. Only a few nucleotides were assigned by using scalar connectivities. Most H3′ and H4′ were assigned instead by using the intensities of NOESY cross-peaks to aromatic protons at short mixing times. All H1′–H8/H6 cross-peaks are weak, showing that the conformation around the glycosidic angle is anti for all nucleotides (Wüthrich, 1986). For anti glycosidic angle and N sugar pucker, H3′ protons are close (2.5–3 Å) to base protons (both on the same and on the neighboring nucleotide), whereas H4′ protons are always distant from aromatic protons (Wüthrich, 1986). Assignments using interproton distances are consistent with the few assignments from scalar connectivities.

The assignment of approximately half of the H4′ protons, and a few H3′, was not possible due to spectral overlap. Some H5′ and H5″ were tentatively identified from weak cross-peaks to H1′ protons at long mixing time (>150 ms), probably due to spin diffusion. For the few cases reported in Table I, the assignment of H5′ and H5″ was supported by the observation of intranucleotide NOESY cross-peaks with H2′, H3′, or H4′.

The chemical shift dispersion of H4′ protons is very poor. On the other hand, H2′ and H3′ protons have good dispersion; in particular, H2′ protons can be anywhere between 4 and 5 ppm and are more dispersed than in DNA. This is likely due to N sugar pucker and A-like stacking, which both enhance ring current effects for H2′ and H3′ protons. Assignments of the proton spectra of a few RNA oligonucleotides in this laboratory suggest that the upfield shift of 3′-terminal H2′ (4 ppm) and the downfield shift of 5′-terminal H2′ (4.8–4.9 ppm) are general features of oligoribonucleotides. In general, 5′-terminal sugar resonances are downfield shifted with respect to all other sugar protons. Downfield shifts are also observed whenever stacking is partially disrupted on the 3′ side (such as for A_9). On the contrary, 3′-terminal sugar protons display anomalous upfield shifts.

Structure. Of the seven degrees of freedom per nucleotide which define RNA structure (Saenger, 1983), sugar pucker and glycosidic torsion angle are the most easily accessible to NMR studies. In oligonucleotides, the sugar undergoes fast exchange between N (close to $C_{3'}$-endo) and S (close to $C_{2'}$-endo) conformers (Altona, 1982; Rinkel & Altona, 1986). The percentage of time the sugar is in the major conformer and the pseudorotation phase angle (P) of the major conformer can be determined from scalar couplings of the sugar protons (Rinkel & Altona, 1986). Glycosidic angles (χ) can be evaluated from intranucleotide aromatic–sugar distances. The detailed procedure is outlined in the following paragraphs. Percentage of the predominant (N) conformer, pseudorotation phase angles, and glycosidic angles are reported in Table II.

A sensitive measurement of deviations from pure N sugar pucker is the observation of H1′–H2′ cross-peaks in 2QF-COSY or two-quantum experiments (Figure 9). We found that two-quantum experiments have considerable advantages in observation of low-intensity peaks, particularly those that lie close to the diagonal in 2QF-COSY. This is likely due to the reduction of t_1 noise caused by the absence of a diagonal. The region of a phase-sensitive two-quantum spectrum where H1′–H2′ cross-peaks are expected is shown in Figure 9. All unlabeled cross-peaks on the left of the HDO ridge correspond to H5–H6 cross-peaks, and several minor form peaks are present. The choice of carrier frequency and sweep width led

Table II: Structural Parameters for the RNA Duplex in 10 mM Phosphate (pH 7)/100 mM NaCl at 30 °C[a]

sequence	% north	pseudorotation phase angle, P (deg)	glycosidic torsion angle, χ (deg)
$5'G_1$	80	−90/−10	−140 ± 10
G_2	>90	−70/+30	−140 ± 10
C_3	>90	−70/+30	−180 ± 30
G_4	>90	−70/+30	−170 ± 30
G_5	75	−90/+40	−135 ± 25
A_6	90	−90/0	−150 ± 30
A_7			
U_8			
A_9	90	−60/+40	−140 ± 30
C_{10}	>90	−50/+20	−155 ± 15
C_{11}	>90	−50/+20	anti
G_{12}	>90	−70/+30	anti
$3'C_{13}$	50	−90/+50	−140 ± 20
$5'G_{14}$	85	−90/+20	−150 ± 40
C_{15}	>90	−70/+30	−155 ± 15
G_{16}	>90	−70/+30	anti
G_{17}	>90	−70/+30	−170 ± 30
U_{18}	>90	−70/+30	−140 ± 40
A_{19}	>90	−70/+30	−165 ± 25
G_{20}			
U_{21}	>90	−70/+30	−180 ± 20
A_{22}	>90	−70/+30	−150 ± 20
C_{23}			
G_{24}	>90	−70/+30	−160 ± 20
C_{25}	>90	−70/+30	−130 ± 30
$3'C_{26}$	50		anti

[a] The % north refers to the percentage of time the sugar pucker is in the north family of conformers ($C_{3'}$-endo family).

to foldover in ω_1 for these cross-peaks; this was tolerated to reduce the sweep width to 4000 Hz, thereby improving the signal/noise ratio and digital resolution. Only a few H1′–H2′ cross-peaks corresponding to terminal or loop nucleotides can be observed and are labeled. Significantly, they correspond to either terminal (C_{13} and C_{26}, more intense than G_1 and G_{14}) or loop (G_5, A_6, A_9) residues. For all other assigned nucleotides, lack of H1′–H2′ cross-peaks in 2QF-COSY or two-quantum experiments indicates that the H1′–H2′ coupling is smaller than the line width, $J_{1',2'} \leq 2$ Hz. This implies that the sugar pucker is N more than 90% of the time (Altona, 1982; Rinkel & Altona, 1986); thus, very little conformational flexibility is observed for most residues. The close proximity of H8/H6 to H3′ protons also points toward N as the major conformer.

The small values of $J_{1',2'}$ restrict the possible values of P to the range −70° to +30° ($C_{3'}$-endo is 18°, $C_{2'}$-exo is −18°). A more precise definition of the pseudorotation phase angle could be obtained from $J_{3',4'}$, very sensitive to small variations of P in the N range of conformers, if the sugar–sugar region of the 2QF-COSY spectrum were less crowded. An alternative is represented by the measurement of interproton distances. In the N range of conformers, the H1′–H4′ distance varies between 2.8 Å (P = 40°) and 3.6 Å (P = −90°). With the exception of a few favorable cases, however, the precision with which interproton distances can be defined (±0.5 to ±1 Å) is not sufficient to define the sugar conformation more precisely.

Two-quantum spectroscopy also provides a means of estimating the value of $J_{1',2'}$, hardly obtainable from correlated experiments since the coupling is of similar magnitude as the line width. The excitation of two-quantum states involving H1′ and H2′ is proportional to sin ($\pi J_{1',2'}\tau$) (Braunschweiler et al., 1983), where τ is the excitation time (50 ms in the present experiment); correspondingly, the intensity of H5–H6

cross-peaks is proportional to sin ($\pi J_{56}\tau$). Excited two-quantum states involving H1′ and H2′ are labeled during t_1 with frequency $\omega_{1'} + \omega_{2'}$, and magnetization is transferred back to H1′ and H2′ by the mixing pulse only when the H1′–H2′ coupling is resolved (i.e., when $J_{1',2'} > 2$ Hz) (Braunschweiler et al., 1983). Two-quantum spectroscopy can then be used to roughly estimate $J_{1',2'}$ by comparing the intensities of H1′–H2′ and H5–H6 cross-peaks ($J_{5,6} = 7$ Hz).

Intermediate $J_{1',2'}$ values (2–5 Hz) can be due either to north/south conformational equilibrium or to nonstandard conformations close to $O_{4'}$-endo, $P = 90°$. The second possibility is ruled out by relatively long H1′–H4′ distances; these protons are very near (2.2–2.4 Å) when the sugar pucker is close to $O_{4'}$-endo (Wüthrich, 1986). The approximate % N was then estimated from $J_{1',2'}$ knowing that $J_{1',2'} \cong 1$ Hz for N sugar pucker and $J_{1',2'} \cong 8$ Hz for S pucker (Altona, 1982).

Glycosidic angles χ were determined by using distances involving sugar and aromatic protons. Relatively long H1′–H8/H6 distances (≥3–3.5 Å for all nucleotides) restrict the available conformational space to the anti range of conformers. More precise values were obtained by comparing measured distances with published plots of aromatic to sugar proton distances vs sugar pucker and glycosidic angle (Wüthrich, 1986). The H3′–H8/H6 distances are very sensitive to variation of χ when the sugar pucker is N. Glycosidic angles are close to the value expected for A-RNA (−155°) for most residues (Table II).

Discussion

Our data provide a detailed picture of the solution structure and dynamics of an oligoribonucleotide with a purine-rich internal loop similar to loop E of eukaryotic 5S ribosomal RNAs. The present oligonucleotide closely matches the sequence of *X. laevis* 5S RNA (Figure 1). U_{73} was omitted in the hope of favoring the formation of G·A pairs, which would stabilize the loop structure.

Base Pairs. Exchangeable proton NMR data strongly favor an open loop, as opposed to closed conformations with mismatched base pairs. This conclusion is based mainly on the observation of a single A·U base pair, even at low temperature, at high NaCl concentration, and in the presence of Mg^{2+} (Figure 2A). Three A·U pairs should be observed for the closed conformation containing two G·A pairs, similar to the structure proposed by Andersen et al. (1984a,b) for loop E in *X. laevis* 5S RNA (Figure 1). The absence of syn nucleotides rules out base pairing schemes not involving imino protons; therefore, mismatched base pairs do not form in this oligonucleotide. Further support for this conclusion comes from the solvent exchange behavior of imino protons from the two base pairs flanking the internal loop. Both G_4 and U_{18} have short open-state lifetimes and small activation enthalpies for base pair opening. They behave similarly to both G_2 and G_{12}, close to the end of the helix, and very differently from base pairs in the core of the stems. In the structure proposed by Andersen et al. (1984a,b) for loop E in 5S RNA, U_{73} is bulged out, making G·A mismatch formation even less favorable than in the present molecule. Deoxyoligonucleotides containing one G·A mismatch were studied by Fazakerley et al. (1986). Mismatched pairs formed only when flanked by G·C base pairs, but not when flanked by A·T pairs. This is consistent with our finding of an open RNA loop when two G·A mismatches are separated by an A·U base pair and one G·A is flanked by A·U pairs.

Magnesium ions frequently play an important role in stabilizing RNA structure. In the present fragment, however, no new imino resonances are observed, and the nonexchangeable proton spectrum is essentially unmodified when Mg^{2+} is added. Therefore, Mg^{2+} does not affect the conformation of this oligonucleotide significantly. However, the U_{18} imino is specifically broadened at high salt concentration and in Mg^{2+}. This might be due to a destabilization of the $A_9 \cdot U_{18}$ pair, or to the appearance of a new structure in intermediate exchange. It is not possible to investigate the structure in Mg^{2+} by 2D NMR. The time required is too long, since Mg^{2+} catalyzes the hydrolysis of one of the strands (G_1–C_{13}) very efficiently.

Stems. In DNA, long double-helical tracts are the rule. In RNA, instead, relatively short double-helical stems connect single-stranded regions. Thus, the structure of short RNA helices and the differences and similarities with long double-helical tracts are of obvious interest.

The structures of the stems

$$5'\text{-}G_1\text{-}G\text{-}C\text{-}G_4\text{-}3'$$

$$3'\text{-}C_{26}\text{-}C\text{-}G\text{-}C_{23}\text{-}5'$$

and

$$5'\text{-}C_{10}\text{-}C\text{-}G\text{-}C_{13}\text{-}3'$$

$$3'\text{-}G_{17}\text{-}G\text{-}C\text{-}G_{14}\text{-}5'$$

are very close to the classical A-RNA geometries observed by fiber diffraction. With the exception of the terminal residues, the sugar conformation is purely N (Table II); the pseudorotation phase angle P is between −70° and +30°. The pucker of classical A-form RNA is $C_{3'}$-endo, $P = +18°$, and the pucker of most nucleotides in the crystal structure of tRNA is $C_{2'}$-exo, $P = -18°$ (Saenger, 1983). Also, the glycosidic angles are close to the value expected for A-RNA (−155°).

Information on the relative position of neighboring nucleotides can be obtained from internucleotide interproton distances. Sugar to base proton distances [H2′/H3′(i)–H8/H6(i + 1)] are generally close to those expected for classical A-RNA. Also, the distances involving H5 protons, H5(i)–H5(i + 1) and H2′/H3′(i)–H5(i + 1), generally agree with classical A-RNA values (Wüthrich, 1986). The observation of AH2 to H1′ cross-peaks across the minor groove is another typical feature of A-like helices. In general, the pattern of interproton distances closely matches that expected for A-form RNA. Thus, structural perturbations induced by the internal loop do not extend beyond the nucleotides at the loop–stem junctions.

Considerable flexibility in the sugar pucker is a common feature of 3′-terminal and, to a lesser extent, 5′-terminal nucleotides in deoxyoligonucleotides (Rinkel & Altona, 1986). A similar behavior is found here, as revealed by nonvanishing populations of S sugar conformers. C_{26} appears to be more flexible than C_{13} and G_1 more flexible than G_{14}, showing the influence of sequence on terminal nucleotide conformational freedom. Perhaps surprisingly, the different phosphate charges at the 5′-termini of the molecule do not have significant structural consequences. In fact, interproton distances involving molecules with either a mono- or a triphosphate on the 5′-nucleotide are very similar. The effect of terminal phosphates on the chemical shifts of nucleotides close to the 5′ end of both strands (Table I) may be due to electrostatic rather than structural effects.

The dynamics of short RNA helices are also of interest. Intrinsic base opening lifetimes for base pairs flanking the loop (U_{18} and G_4) and close to helix termini (G_2 and G_{12}) are between 3 and 6 ms at 40 °C. These values are similar to those found for the RNA polynucleotides poly(rA)·poly(rU) and poly(rI)·poly(rC) (≈2 ms at 37 °C) (Leroy et al., 1985a).

Activation enthalpies for the same base pairs are between 15 and 20 kcal/mol, again in agreement with the results of Leroy et al. (1985a), who report activation enthalpies between 18 and 20 kcal/mol. Lifetimes and enthalpies are much higher for internal G·C base pairs in the present oligonucleotide. To our knowledge, the lifetime of internal G·C pairs in RNA has been investigated only for tRNA (Leroy et al., 1985b). Intrinsic opening lifetimes of about 30 ms are found, comparable with our results (50–150 ms). In deoxyoligonucleotides, opening lifetimes for both internal G·C and A·T pairs are between 1 and 10 ms (Leroy et al., 1988). Thus, while the intrinsic opening rate of A·U pairs in RNA is similar to that of A·T pairs in DNA, this does not appear to hold true for G·C pairs. In fact, the lifetimes of the open state of internal G·C pairs in our RNA molecule (>50 ms at 35 °C) are much longer than in DNA oligonucleotides.

Loop Conformation. The analysis of interproton distances shows that extensive stacking is conserved in the loop and, to a lesser extent, at loop–stem junctions. In the core of the internal loop (U_{21}-A_{22} and A_6-A_7), the conformation is close to that of the stems, even if interstrand base pairing is broken. For both U_{21} and A_{22}, the sugar pucker is mainly N (>90%), and most internucleotide interproton distances are close to those expected for standard A form. However, the upfield shift of U_{21}H5, 4.95 ppm, is unusual. Since the U_{21}–A_{22} step conserves A-form stacking, this anomalous chemical shift should be attributed to the ring current of G_{20}. The glycosidic angle of U_{21} (−180° ± 20°) might play a role in positioning U_{21}H5 in the field of G_{20}; unfortunately, it was not possible to assign the protons of G_{20}. The sugar pucker is predominantly N also for A_6 and G_5, suggesting a strong tendency for ribonucleotides to conserve N sugar pucker. A_6 sugar protons are close to A_7H8, showing the conservation of stacking between A_6 and A_7. However, distortions from typical A-form distances are observed, and the A_7 sugar protons are involved in the dynamic processes discussed in the next section.

Both sides of the loop continue the stacking from the stems at the 3′ ends. Both U_8H6–A_9H8 and A_{22}H8–C_{23}H6 connectivities are observed at very long mixing times (500 ms), perhaps via spin diffusion. Unlike the 3′ sides, however, the two sides of the loop differ at the 5′ ends. On one side, G_5 stacks above G_4 for at least part of the time. On the opposite side, it is impossible to find any connectivity to or from G_{20} in NOESY spectra. This is likely a consequence of the conformation of (A_9)·(U_{18}-A_{19}). A_{19} stacks above U_{18}, conserving A-form connectivities, but, quite surprisingly, hydrogen bonding with U_8 is broken. While the A_9 sugar experiences some conformational freedom, probably because the U_8·A_{19} base pair is not formed, the sugar pucker for both U_{18} and A_{19} is more than 90% N (Table II). In general, interproton distances between U_{18} and A_{19} suggest the conservation of A-like stacking. However, the relatively short A_9H2–A_{19}H2 and A_9H1′–A_{19}H2 distances (≈3 Å) suggest a positive roll ("B-like") for A_{19}. The downfield shifts of A_9H2 and A_{19}H2 with increasing temperature (Figure 4B) suggest that some of these conformational features are temperature sensitive.

No sign of sharp turns or other unusual conformations is revealed by the analysis of the ^{31}P spectrum. All phosphorus resonances group together between −3.5 and −5 ppm. This shows that the backbone angles adjacent to the phosphorus, α and ζ, are in the canonical (g^-,g^-) conformation found in both A- and B-form double helices (Gorenstein, 1984). Consequently, the stems are bridged smoothly by the loop, without any sharp turn such as those observed in Z-RNA or in the crystal structure of tRNA. Since the conformation of the backbone and the base stacking in the loop are close to A form, one might ask whether A-form stacking and backbone can be conserved while interstrand hydrogen bonds are broken. Model building reveals that the disruption of hydrogen bonds can be accommodated smoothly in the loop with small distortions from A-form geometry for each strand. Opening the loop does not require major distortions from A-form stacking.

The conservation of stacking in the loop provides a structural basis for the large favorable enthalpy of duplex formation. Comparison with estimates for perfect Watson–Crick duplexes of similar sequence shows that stacking in the loop is as favorable as in base-paired helices. The thermodynamic results are very different from those on pyrimidine-rich internal loops (Gralla & Crothers, 1973). Phylogenetic studies have shown that most conserved internal loops in ribosomal RNA are purine rich (Noller, 1984; Leffers et al., 1988). It is tempting to suggest that the tendency of purines to conserve base stacking is responsible for the thermodynamic stability of these conserved internal loops.

The NMR data presented here suggest that the structure of loop E in *X. laevis* 5S RNA, whose sequence is very close to that of the present oligonucleotide, might be more similar to A form than generally believed (Andersen et al., 1984a,b; Romaniuk et al., 1988). This can explain why highly conserved nucleotides in loop E can be substituted and perfect Watson–Crick pairing restored with relatively small effects on TFIIIA binding (Romaniuk, 1989). It is interesting that structural deviations from A form are localized at the loop–stem junctions, where the strongest TFIIIA–RNA contacts are found (Christiansen et al., 1987).

Loop Dynamics. The presence of slow dynamics in the loop is revealed by selective broadening of some aromatic and sugar resonances. All broadened resonances are at loop–stem junctions (G_5, U_8, C_{23}); resonances in the stems and in the core of the loop (A_6, U_{21}, A_{22}) are sharp. Since most resonances are sharp, broadening must be attributed to intermediate exchange between different conformers. In order to lead to broadening, the rate of exchange has to be comparable to the frequency difference between the proton resonances in the two conformers; the populations of the conformers must also be comparable. Estimates of the frequency shift of aromatic protons in different magnetic environments range between 50 and 500 Hz (at 500 MHz); these are typical differences between mononucleotides and oligonucleotides. Consequently, the dynamic processes responsible for broadening are slow, in the range of 10^{-2}–10^{-4} s.

Estimates of magnetic shieldings for RNA have been made by Arter and Schmidt (1976). The quantitative agreement between predicted and observed chemical shifts is generally poor, but the qualitative agreement is good, particularly for aromatic protons. Arter and Schmidt (1976) predict that the chemical shift of aromatic protons is determined mainly by the ring current of the 5′ base; contributions from the 3′ base, from next-nearest neighbors, and from interstrand hydrogen bonding are generally smaller. For example, the chemical shift of C_{23}H6 and C_{23}H5 is mostly determined by the position of A_{22}. A motion in the millisecond range of A_{22} between conformations where the A_{22}–C_{23} stacking is conserved or disrupted would lead to broadening of the C_{23} resonances. The other alternative, motion of C_{23} relative to G_{24}, is unlikely since base pairing to G_4 is conserved. Furthermore, the contribution of G_{24} and C_{25} to the shielding of C_{23} base protons is minimal, and broadening would be observed only for exceptionally slow motion (slower than 1 s). Similarly, the chemical shift of G_5H8 is determined by the position of G_4, whereas G_4 chemical

shifts are determined by C_3. Thus, dynamics in the millisecond scale between conformers which conserve or do not conserve the G_4–G_5 stacking would broaden G_5 and maintain G_4 resonances sharp. Similar conclusions can be reached also for U_8; however, broadening of A_7 sugar resonances suggests that also this moiety and, perhaps, the phosphodiester backbone are involved in the same dynamics as U_8.

The previous analysis suggests that nucleotides at loop–stem junctions move between conformations stacked and unstacked above the stems. At the same time, resonances from the core of the loop are sharp. Therefore, nucleotides in the center of the loop are always experiencing the same magnetic environment when nucleotides at loop–stem junctions move. If A_6 were moving concertedly with G_5, the magnetic shielding of A_6 protons would not change, and A_6 resonances would be sharp while G_5 resonances would be broad. Similarly, if U_{21} were moving with A_{22}, the resonances of A_{22} would remain sharp while C_{23} resonances would be broad. Thus, a concerted motion of the loop nucleotides of each of the two strands would explain why nonexchangeable proton resonances from the center of the loop remain sharp while resonances at loop–stem junctions are broad. The collective character of this motion would also explain the slow time scale for the dynamics.

The nonobservation of resonances from loop imino protons is somewhat unexpected. In fact, the conservation of stacking in the loop suggests some protection from exchange; as in hairpin loops, one would expect to observe imino protons from non-base-paired nucleotides at low temperature and pH. Indeed, imino resonances from nonpaired nucleotides are observed at neutral pH in a second internal loop currently under investigation. We propose that the nonobservation of imino resonances from the loop has a dynamic explanation. If the concerted motion of the loop involved the transient formation of some interstrand hydrogen bonds ("breathing"), imino resonances from loop nucleotides would be very broad. In fact, the frequency difference between hydrogen-bonded and non-hydrogen-bonded imino resonances is approximately 1000 Hz at 500 MHz, or just about the rate of conformational exchange deduced from nonexchangeable proton data. At the same time, nonexchangeable proton resonances from the core of the loop would remain sharp because hydrogen bonding contributions to the chemical shift of aromatic and sugar protons are small (Arter & Schmidt, 1976). The transient formation of interstrand hydrogen bonds, coupled with increased solvent accessibility with respect to a regular A-form double helix, would explain why imino resonances from the loop are not observed. Thus, although base pairing is not directly observable, interstrand hydrogen bonds might be formed and disrupted on a millisecond time scale as the loop is "breathing".

Conclusions. Although common RNA secondary structure motifs, internal loops have been studied only slightly. Here we have used one- and two-dimensional NMR to investigate their conformation and dynamics. The quality of the NMR spectra shows that large quantities of essentially pure RNA oligonucleotides can be synthesized efficiently by T7 RNA polymerase (Milligan et al., 1987). The number of assigned protons and the amount of structural information are comparable to those obtainable with DNA oligonucleotides of similar size. In particular, most sugar protons (H2′, H3′, and H4′) can be assigned despite the spectral overlap.

Conformation and dynamics of the two helical stems are very similar to those expected for classical A-form RNA. This result is of general relevance, since short helical stems connecting single-stranded or looped regions are the main building block of RNA structure. The conformation of the loop is open, and no evidence is found for mismatched base pairs or syn nucleotides. However, extensive stacking and essentially A-like connectivities are conserved in the loop. The conservation of base stacking in the loop provides a structural basis for the large favorable enthalpy of duplex formation. The selective broadening of some resonances suggests a slow ("breathing") mode of the whole loop, in the millisecond time scale, perhaps involving the transient formation of some interstrand hydrogen bonds. Indeed, the main differences between this internal loop and classical A-form RNA are dynamic and not structural.

Major structural deviations from the A-form pattern are localized in nucleotides connecting the loop to the stems. This may be relevant for the specific recognition of loop E in *X. laevis* 5S RNA by TFIIIA and, perhaps, by ribosomal proteins. In fact, TFIIIA strongly protects nucleotides at loop–helix junctions in *X. laevis* 5S RNA (Christiansen et al., 1987).

Acknowledgments

We thank J. D. Puglisi and J. R. Wyatt for stimulating discussion and suggestions on the synthesis and characterization of RNA oligonucleotides, D. Koh for the synthesis of DNA oligonucleotides, and Barbara Dengler for managing the laboratory.

References

Altona, C. (1982) *Recl. Trav. Chim. Pays-Bas 101*, 413–433.
Andersen, J., Delihas, N., Hanas, J. S., & Wu, C.-W. (1984a) *Biochemistry 23*, 5752–5759.
Andersen, J., Delihas, N., Hanas, J. S., & Wu, C.-W. (1984b) *Biochemistry 23*, 5759–5766.
Arter, D. M., & Schmidt, P. G. (1976) *Nucleic Acids Res. 3*, 1437–1447.
Branch, A. D., Bauenfeld, D. J., & Robertson, H. D. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 6590–6594.
Braunschweiler, L., Bodenhausen, G., & Ernst, R. R. (1983) *Mol. Phys. 48*, 535–560.
Christiansen, J., Brown, R. S., Sprout, B. S., & Garret, R. A. (1987) *EMBO J. 6*, 453–460.
Donis-Keller, H., Maxam, A. M., & Gilbert, W. (1977) *Nucleic Acids Res. 4*, 2527–2538.
Fazakerley, G. V., van der Marel, G. A., van Boom, J. H., & Guschlbauer, W. (1984) *Nucleic Acids Res. 12*, 8269–8279.
Fazakerley, G. V., Quignard, E., Woisard, A., Guschlbauer, V., van der Marel, G. A., & van Boom, J. H. (1986) *EMBO J. 5*, 3697–3703.
Freier, S. M., Kiersek, J. A., Jaeger, J. A., Sujimoto, N., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373–9377.
Gewirth, D. T., Abo, S. R., Leontis, N. B., & Moore, P. B. (1987) *Biochemistry 26*, 5213–5220.
Gorenstein, D. G. (1984) *Phosphorus-31 NMR: Principles and Applications* (Gorenstein, D. G., Ed.) Academic Press, New York.
Gralla, J., & Crothers, D. M. (1973) *J. Mol. Biol. 78*, 301–319.
Gregory, R. J., Cahill, P. B. F., Thurlow, D. L., & Zimmerman, R. A. (1988) *J. Mol. Biol. 204*, 295–307.
Gronenborn, A. M., & Clore, G. M. (1985) *Prog. Nucl. Magn. Reson. Spectrosc. 17*, 1–33.
Hore, P. J. (1983) *J. Magn. Reson. 55*, 283–300.
Leffers, H., Egebjerg, J., Andersen, A., Christensen, T., & Garret, R. A. (1988) *J. Mol. Biol. 204*, 507–522.
Leroy, J.-L., Broseta, D., & Gueron, M. (1985a) *J. Mol. Biol. 184*, 165–178.
Leroy, J.-L., Bolo, N., Figueroa, N., Plateau, P., & Gueron, M. (1985b) *J. Biomol. Struct. Dyn. 2*, 915–939.

Leroy, J.-L., Kochoyan, M., Huyn-Dinh, T., & Gueron, M. (1988) *J. Mol. Biol. 200*, 223–238.
Li. S.-J., Wu, J., & Marshall, A. G. (1987) *Biochemistry 26*, 1578–1585.
Macura, S., & Ernst, R. R. (1979) *Mol. Phys. 41*, 95–117.
Macura, S., Wüthrich, K., & Ernst, R. R. (1982) *J. Magn. Reson. 46*, 269–282.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Milligan, J. F., Groebe, D. R., Witherell, G. W., & Uhlenbeck, O. C. (1987) *Nucleic Acids Res. 15*, 8783–8798.
Müller, N., Ernst, R. R., & Wüthrich, K. (1986) *J. Am. Chem. Soc. 108*, 6482–6492.
Noller, H. F. (1984) *Annu. Rev. Biochem. 53*, 119–162.
Rinkel, L. J., & Altona, C. (1986) *J. Biomol. Struct. Dyn. 4*, 621–649.
Romaniuk, P. J. (1989) *Biochemistry 28*, 1388–1395.
Romaniuk, P. J., de Stevenson, I. L., & Wong, H. H. A. (1987) *Nucleic Acids Res. 15*, 2737–2755.
Romaniuk, P. J., de Stevenson, I. L., Ehresmann, C., Romby, P., & Ehresmann, B. (1988) *Nucleic Acids Res. 16*, 2295–2312.
Romby, P., Westhof, E., Taukifimpfa, R., Mache, R., Ebel, J.-P., Ehresmann, C., & Ehresmann, B. (1988) *Biochemistry 27*, 4721–4731.
Saenger, W. (1983) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
Stern, S., Weiser, B., & Noller, H. F. (1988) *J. Mol. Biol. 204*, 447–481.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Chapter 11–13, Wiley, New York.

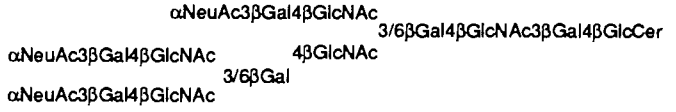
Novel Tri- and Tetrasialosylpoly-*N*-acetyllactosaminyl Gangliosides of Human Placenta: Structure Determination of Pentadeca- and Eicosaglycosylceramides by Methylation Analysis, Fast Atom Bombardment Mass Spectrometry, and ^{1}H NMR Spectroscopy†

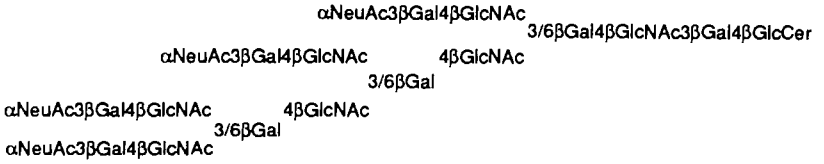
Steven B. Levery, Edward D. Nudelman, Mary Ellen K. Salyan, and Sen-itiroh Hakomori*

The Biomembrane Institute and Departments of Chemistry and Pathobiology, University of Washington, 201 Elliott Avenue West, Seattle, Washington 98104

Received February 8, 1989; Revised Manuscript Received May 24, 1989

ABSTRACT: A series of highly polar neolacto series (poly-*N*-acetyllactosaminyl) gangliosides were isolated from human placenta tissue and purified by HPLC and preparative HPTLC. Two of these ganglioside fractions (G-12 and G-13) were analyzed by 500-MHz ^{1}H NMR spectroscopy, GC–EIMS, $^+$FAB-MS, and sequential exoglycosidase treatments. Their structures have been identified as being of the repeating *N*-acetyllactosamine type, multiply branched through GlcNAcβ1→6/3 linkages, with every nonreducing Gal terminal α2→3-sialosylated, as shown below. These are among the highest molecular weight glycosphingolipids whose detailed oligosaccharide structures are presently known.





The presence of glycoconjugates having repeating *N*-acetyllactosamine units in certain types of cells and tissues has been known since they were initially found in highly polar glycolipids (Hakomori et al., 1972), as well as in long-chain glycosaminoglycans bound to transmembrane proteins of erythrocytes (Järnefelt et al., 1978; Fukuda, M., et al., 1979). Such glycoconjugates were originally identified as carriers of blood group ABH determinants and were found to be present in either branched or unbranched form (Hakomori et al., 1972, 1977). The unbranched species are present predominantly in fetal erythrocytes, while the branched forms are found mainly in adult erythrocytes (Watanabe & Hakomori, 1976). Subsequently, the unbranched form was identified as the classically known i antigen (Niemann et al., 1978) and the branched form as I antigen (Watanabe et al., 1979b). Thus, the ontogenic development of erythrocytes and erythroid cells is associated

†This study was supported by funds from National Institutes of Health Outstanding Investigator Grant CA42505 and from The Biomembrane Institute.

0006-2960/89/0428-7772$01.50/0