Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

J. Mol. Biol. (1990) **214**, 437–453

Conformation of an RNA Pseudoknot

Joseph D. Puglisi†, Jacqueline R. Wyatt and Ignacio Tinoco Jr‡

Department of Chemistry and Laboratory of Chemical Biodynamics
University of California, Berkeley, CA 94720, U.S.A.

(Received 23 November 1989; accepted 9 March 1990)

The structure of the 5′ GCGAUUUCUGACCGCUUUUUUGUCAG 3′ RNA oligonucleotide was investigated using biochemical and chemical probes and nuclear magnetic resonance spectroscopy. Formation of a pseudoknot is indicated by the imino proton spectrum. Imino protons are observed consistent with formation of two helical stem regions; nuclear Overhauser enhancements between imino protons show that the two stem regions stack to form a continuous helix. In the stem regions, nucleotide conformations (3′-*endo*, *anti*) and internucleotide distances, derived from two-dimensional correlated, spectroscopy and two-dimensional nuclear Overhauser effect spectra, are characteristic of *A*-form geometry. The data suggest minor distortion in helical stacking at the junctions of stems and loops. The model of the pseudoknot is consistent with the structure originally proposed by Pleij *et al.*

1. Introduction

The RNA from various plant viruses have structures at their 3′ termini that are recognized by tRNA processing enzymes (Haenni *et al.*, 1982; Rietveld *et al.*, 1982; Joshi *et al.*, 1983; Guerrier-Takada *et al.*, 1988) even though the secondary structures of these viral RNAs do not resemble the typical tRNA cloverleaf (Florentz *et al.*, 1982; Rietveld *et al.*, 1983). Pleij and co-workers proposed that a novel tertiary interaction, a pseudoknot allows turnip yellows mosaic virus (TYMV§) RNA to adopt a three-dimensional structure similar to that of tRNA (Rietveld *et al.*, 1983, 1984; Pleij *et al.*, 1985, 1986). The exact function of the tRNA-like structures at the 3′ ends of the plant viral RNAs is unclear. The tRNA-like structure, and thus the pseudoknot, is required for viral replication in brome mosaic virus RNA (Dreher & Hall, 1988*a*) and may be a relic of a primitive telomere (Weiner & Maizels, 1987). Pseudoknots have been proposed to play a functional role in a number of other RNAs (Moazed & Noller, 1987; James *et al.*, 1988; Brierley *et al.*, 1989; Tang & Draper, 1989; for a review, see Schimmel, 1989). A pseudoknot occurs, as shown in Figure 1, when nucleotides within a loop pair with nucleotides outside the loop, forming a second stem (Studnicka *et al.*, 1978). An idealized model for pseudoknot structure was proposed on the basis of biochemical data and structural analogy to tRNA (Pleij *et al.*, 1985; Dumas *et al.*, 1987). The stem regions stack coaxially in an undistorted *A*-form geometry. The two loop regions in the pseudoknot are inequivalent; loop 1 crosses the deep major groove of stem 2 and loop 2 crosses the shallow and wide minor groove of stem 1. Chemical and enzymatic mapping experiments, mutational analysis and the functional similarities of the viral RNAs to normal tRNA are consistent with this model (Rietveld *et al.*, 1982; Joshi *et al.*, 1983; van Belkum *et al.*, 1988; Dreher & Hall, 1988*b*). An oligonucleotide of 19 nucleotides was shown to form a pseudoknot using nuclease mapping and absorbance melting studies (Puglisi *et al.*, 1988). However, no structural data on the three-dimensional conformation of a pseudoknot have been presented.

Nuclear magnetic resonance (n.m.r.) provides the most detailed structural information on nucleic acids in solution. Most n.m.r. studies of RNA have concentrated on the exchangeable, base-pairing protons of large biochemically isolated molecules. The structures of $tRNA^{Phe}$ (76 nucleotides) (Reid, 1981; Heerschap, 1985; Hall *et al.*, 1989) and *Escherichia coli* 5 S rRNA (122 nucleotides) Leontis

† Present address: Institut de Biologie Moleculaire et Cellulaire de CNRS, 15, rue René Descartes, 67084 Strasbourg, France.

‡ Author to whom all correspondence should be addressed.

§ Abbreviations used: TYMV, turnip yellow mosaic virus; n.m.r., nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; NOESY, 2-dimensional nuclear Overhauser effect spectroscopy; s.w., sweep width; COSY, 2-dimensional correlated spectroscopy; DQF, double-quantum filtered; p.p.m., parts per million; DEP, diethylpyrocarbonate.

0022-2836/90/140437-17 $03.00/0

5' GCGAUUUCUGACCGCUUUUUUGUCAG 3'

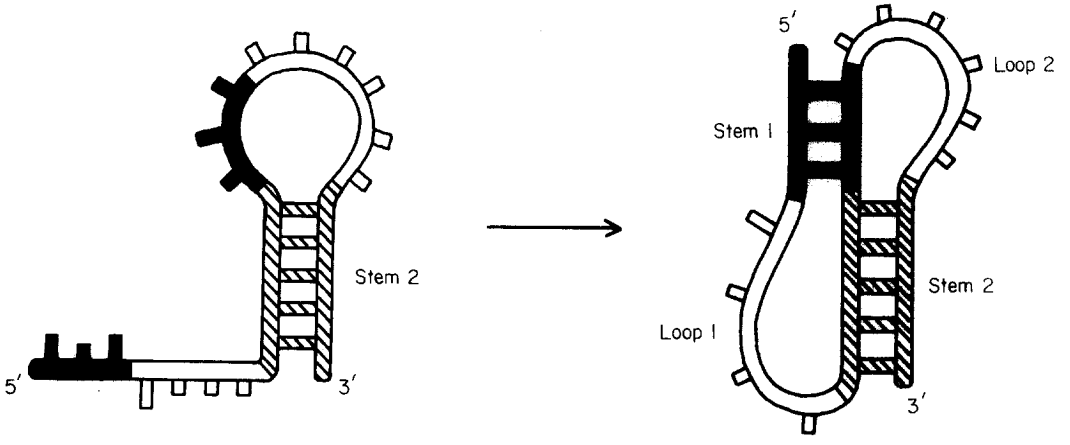


Figure 1. Sequence (top) and proposed folding (bottom) for the RNA oligonucleotide. Stem and loop regions in the pseudoknot structure are indicated; stem 1 is black and stem 2 is hatched.

& Moore, 1986; Gewirth & Moore, 1988) have been studied using imino proton n.m.r. data. The imino protons primarily give information about base-pairing. The plant viral RNAs containing pseudoknots are quite large (>5000 nucleotides) but Pleij and co-workers used RNase H cleavage to yield much smaller fragments (112, 88 and 44 nucleotides) of TYMV RNA (van Belkum *et al.*, 1989). The n.m.r. spectrum of the imino region of even the smallest fragments is quite complex. Much more detailed data on individual nucleotide and strand conformations can be obtained from the non-exchangeable protons but there have been few such studies of RNA (Clore *et al.*, 1984; Happ *et al.*, 1988; Varani *et al.*, 1989; Chou *et al.*, 1989; Davis *et al.*, 1990; Puglisi *et al.*, 1990).

This paper describes characterization of the pseudoknot structure formed by an oligonucleotide of 26 nucleotides with the sequence shown in Figure 1. Both biochemical and physical methods were used to determine the structure formed by the oligonucleotide. Nuclease mapping experiments provided initial evidence that a pseudoknot structure is formed by this sequence. One-dimensionl NOE experiments in water permitted the assignment of the imino resonances and demonstrated the coaxial stacking of the two stem regions in the pseudoknot. Two-dimensional n.m.r. experiments on the non-exchangeable resonances provided estimates of interproton distances and ribose torsion angles, thus yielding a detailed view of the structure of this pseudoknot in solution.

2. Materials and Methods

(a) *Synthesis of oligonucleotides*

RNA oligonucleotides were synthesized using bacteriophage T7 RNA polymerase and a synthetic DNA template with a double-stranded promoter region and a single-stranded template region (Milligan *et al.*, 1987). T7 RNA polymerase was purified as described (Davenloo *et al.*, 1984). DNA oligonucleotides were synthesized using standard solid-phase phosphoramidite chemistry. Transcription reactions were incubated at 37°C for approximately 2 h in 40 mM-Tris·HCl (pH 8·1 at 37°C), 1 mM-spermidine, 5 mM-dithiothreitol, 0·01% (v/v) Triton X100, 80 mg polyethylene glycol/ml (relative molecular mass = 8000), 36·8 mM-$MgCl_2$, 4 mM each nucleoside triphosphate and 5 mM-GMP. The DNA concentration was 0·3 μM per strand and T7 RNA polymerase was used at 30 units/μl. Typical yields were between 10 to 15 nmol of strand per ml of reaction (2 to 3 A_{260} units/ml). For n.m.r. quantities, reaction volumes were 50 to 75 ml. Reactions were quenched by the addition of EDTA to 50 mM. The reaction mixtures were extracted with phenol, then with chloroform and were precipitated with ethanol. Full-length transcripts were isolated from a denaturing (7 M-urea) 20% (w/v) polyacrylamide gel. Reactions yielded 2 major bands corresponding to the correct length and to transcript with an additional nucleotide at the 3' terminus. Gel electrophoresis allowed good resolution of these products. Resolution was worse when reactions were primed using GMP. Incorporation of GMP at position 1 was about 80% efficient as determined by 2-dimensional thin-layer chromatography analysis of total RNase digests of transcripts primed with GMP and labeled with [α-^{32}P]CTP (Nishimura *et al.*, 1967).

RNA transcripts were electroeluted from the gel using an Elutrap electro-elution apparatus (Schleicher & Schuell) and then precipitated with ethanol. Final purification involved extensive dialysis (approx. 24 h) using a microdialysis apparatus (Bethesda Research) first *versus* 5 mM-EDTA, 10 mM-sodium phosphate (pH 6·4), then *versus* water.

The sequence and length were verified by 2-dimensional thin-layer chromatography analysis of products of complete enzymatic hydrolysis of internally labeled or 3' end-labeled transcripts (Kuchino *et al.*, 1980) and by partial digestion of 5' end-labeled transcript with base-specific RNases. A Bio-Rad TSK-125 gel-filtration column

was used to ensure that the structure formed was monomeric even at high strand concentration. A buffer identical with the n.m.r. buffer was used and retention time was compared with the retention times of RNA molecules of known structure and size.

Enzymatic and chemical structure mapping experiments were performed as described in the following paper (Wyatt *et al.*, 1990).

(b) *Deuteriation of nucleotides*

GTP and GMP were deuteriated at the 8 position of guanine by incubating in ^2H_2O at 50°C for 24 h (Benevides *et al.*, 1984). UTP was deuteriated at the 5 position of uridine using a modification of the method of Hayatsu (Hayatsu *et al.*, 1970; Wataya & Hayatsu, 1972; Brush *et al.*, 1988). A stock solution of $(NH_4)_2SO_3$ (2·5M) was lyophilized twice from ^2H_2O; 120 mg UTP was dissolved in 1·0 ml of 2·5 M $(N^2H_4)_2SO_3$ and adjusted to pH 8·0 using NaO^2H. The UTP was incubated for 28 h at 50°C. The reaction was monitored by the disappearance of the H5 resonance and the collapse of the H6 doublet into a singlet in the n.m.r. spectrum. Incubation at higher temperatures or higher pH resulted in hydrolysis to UDP and UMP. The majority of the product was not the d^5UTP, but the adduct of bisulfite to UTP (Wataya & Hayatsu, 1972). The reaction mixture was cooled to 22°C and the pH raised to 9·5 to 10·0 by addition of NaO^2H. Incubation at 22°C for 2 h resulted in complete conversion of the bisulfite adduct to the desired d^5UTP as monitored by n.m.r. The d^5UTP was desalted on a G-10 size-exclusion column and lyophilized. Deuteriated nucleotides were used in transcription reactions without further purification.

(c) *n.m.r. spectroscopy*

n.m.r. spectra were acquired on a Nicolet GN-500 (500 MHz ^1H) n.m.r. spectrometer. n.m.r. samples were prepared by dissolving purified RNA in approx. 300 μl of 5 mM-$MgCl_2$, 50 mM-NaCl, 10 mM-sodium phosphate (pH 6·4) and dialyzing *versus* this buffer for 24 h. The final volume of the sample was adjusted to 350 μl by addition of buffer and lyophilized. For studies of exchangeable protons, the sample was dissolved in 95% H_2O/5% ^2H_2O up to 350 μl volume. For studies of non-exchangeable protons, samples were lyophilized twice from ^2H_2O and resuspended in 350 μl of 99·96% ^2H_2O (Aldrich). Concentrations of n.m.r. samples were 1·0 to 1·5 mM-strand. All n.m.r. spectra are referenced relative to trimethylsilylpropionate (TSP).

n.m.r. spectra of exchangeable protons were acquired using a 1-3-3-1 pulse sequence (Hore, 1983). The carrier was placed on the water resonance with a delay time of 125 μs and a spectral width of 15,500 Hz. One-dimensional NOE were measured using decoupler irradiation with a 1-3-3-1 pulse sequece. The irradiation time was 600 ms. All 1-dimensional NOE measurements were done at 4·0°C. NOE difference spectra were obtained by interleaving on and off-resonance irradiation. Typically, 4000 scans were taken to achieve good signal to noise. Double-precision acquisition allowed the use of high (16 bit) digitizer resolution. Decoupler spillover was determined by irradiating at variable offsets from a peak. Spillover was estimated at 100 Hz at the decoupler power used. Difference spectra were processed using 5 Hz line broadening.

NOESY spectra were obtained using the standard pulse sequence and phasing cycling for pure phase spectra (Bodenhausen *et al.*, 1984). Data sets were 2000 in t_2 with a sweep-width (s.w.) of 4464·28 Hz. A total of 450 to 500 t_1 values were collected with 128 scans per t_1. A relaxation delay of 1·8 to 2·0 s was used. Total acquisition time was approx. 30 h. NOESY spectra were obtained at 400 ms for assignment of the spectrum and 150 ms for distance determination. NOESY spectra at short (< 150 ms) mixing time had poor signal-to-noise. All NOESY spectra were measured at 23°C. Spectra were processed using a skewed sine bell function with a 30° phase shift and a skew factor of 0·7.

Interproton distances were approximated from NOE cross-peak volumes at short mixing times. The NOE intensities of the fixed CH6–CH5 distance (2·45 Å; 1 Å = 0·1 nm) was used as a reference. These intensities varied by only 10 to 20% among the 6 cytosine residues. Volume integrals of resolved cross peaks were measured using an ellipse based method in FTNMR (D. Hare, Infinity Systems). The mixing time of 150 ms was adequate for estimating intermediate range distances (> 2·5 Å).

Phase-sensitive COSY and double-quantum-filtered COSY (DQF-COSY) spectra were obtained using the standard pulse sequences and phase cycling (Piantini *et al.*, 1982; Bodenhausen *et al.*, 1984). Data sets were 2000 in t_2 with a s.w. of 4464·28 Hz; 128 scans were acquired per t_1; typical data sets consisted of 450 to 500 t_1 values. Data were processed using a sine bell with a 30° phase shift in both t_2 and t_1.

Double-quantum (DQ) spectra were obtained using the 90-τ/2-180-τ/2-90-t_1-90-t_2 pulse sequence with standard phase cycling (Braunschweiler *et al.*, 1983). The carrier was placed just upfield from the H2′/H3′/H4′/H5′/H5″ region of the spectrum. A s.w. of 400 Hz was used, which resulted in some foldover in the H5–H6 region. The delay time (τ) was set to 50 ms to maximize excitation of DQ coherences with coupling of approx. 10 Hz in an AX spin system. Data sets were 2000 in t_2; 400 t_1 values with 96 scans per t_1 were acquired for a total experiment time of 24 h. Data were apodized using a sine bell with a 30° phase shift.

(d) *Assignment strategy*

The assignment of the n.m.r. spectrum of an RNA molecule is a more difficult task than the assignment of the equivalent DNA spectrum, and requires modification of the DNA assignment strategy (Hare *et al.*, 1983). There are 2 major problems: one involves the intrinsic chemical shifts of protons in RNA and the second arises from the structural diversity of RNA molecules. The resonances in a 1-dimensional spectrum of DNA are more dispersed than those in a 1-dimensional spectrum of RNA. Aromatic and sugar H1′ protons resonate in approximately the same regions in both RNA and DNA. In DNA the H2′/H2″ protons resonate far upfield at 2 to 3 p.p.m., whereas H3′ protons resonate at 4·5 to 5·0 p.p.m.; only the H4′ and H5′/H5″ resonances overlap at 4·0 to 4·5 p.p.m. Unfortunately, the H2′, H3′, H4′, H5′/H5″ resonances of RNA all overlap in the region from 3·9 to 5·0 p.p.m.

To assign a non-exchangeable n.m.r. spectrum of RNA, the double-stranded stem regions must first be identified. Data from chemical and enzymatic mapping experiments give a crude view of secondary structure; assignments of exchangeable imino proton resonances help confirm base-paired regions. In an *A*-form duplex, a connectivity pathway exists from H8/H6 to the H1′ on its own sugar (3·5 Å) and to the H1′ on the 5′-neighboring sugar (4·6 Å) (Haasnoot *et al.*, 1984). Although the distance to the 5′-

neighboring H1′ is quite long, the distance from the aromatic H6/H8 to 5′-neighboring H2′ proton is very short (2·0 Å) (see Wüthrich (1986) for short distances in an idealized *A*-form helix). Therefore, spin diffusion from the H2′ to the H1′ will give rise to a relatively intense 3′ (H8/H6) → 5′ (H1′) cross peak in the NOESY spectrum at long mixing times.

RNA aromatic protons can only be identified as pyrimidine or purine; in DNA, pyrimidines can be distinguished in NOESY by the thymidine H6 to methyl cross peak. Incorporation of selectively deuteriated nucleotides during RNA synthesis permits the distinction of nucleotide type, by comparison to the spectrum of the non-deuteriated molecule. The AH2 protons are important, since they provide a link to the more readily assigned exchangeable imino protons. An AH2 proton in a double-stranded region is easily identified by an NOE from the A·U base-paired imino proton. The AH2 protons usually have no close distances to non-exchangeable protons in a *B*-form helix. However, in an *A*-form helix, the AH2 protons have NOE connectivities in the 3′ direction to the H1′ on the same strand, and in the 5′ direction to the H1′ on the opposite strand (both distances approx. 4 Å). An AH2 misassigned as a purine H8 will lead to an erroneous assignment pathway that shunts across strands. AH2 protons generally have only very weak NOEs in the H2′/H3′/H4′/H5′/H5″ region. The AH2 protons, once assigned from the imino spectrum, act as a check on the H8/H6 → H1′ assignment pathway. Assignment of non-helical regions follows the above protocol if the nucleotides are in a stacked, relatively *A*-form geometry (i.e. the anticodon loop of $tRNA^{Phe}$); however, assignment pathways are much more difficult to obtain in unstacked regions. This is a serious problem for RNA structure determination by n.m.r., since frequently the non-helical regions of the structure are of primary interest.

The assignment of the H8/H6 protons in an *A*-form stem can be checked by the strong internucleotide NOE from 3′ (H8/H6) ↔ 5′ (H2′) (approx. 2·Å). In general, each H8/H6 proton in RNA will have a number of NOEs to resonances in the H2′, H3′, H4′, H5′/H5″ region between 3·9 and 5·0 p.p.m. The NOE to the H2′ should be the strongest. This can be confirmed by checking the H1′/H5 (5·0 to 6·4 p.p.m.) to H2′/H3′/H4′/H5′/H5″ (3·9 to 5·0 p.p.m.) region of the NOESY. A strong H1′ to H2′ (distance 2·5 to 2·7 Å) cross peak identifies the H2′ on the same nucleotide. In DNA, this assignment can be confirmed unambiguously by looking for COSY connectivities between the H1′ and H2′/H2″. However, $J_{1'2'}$ depends strongly on sugar pucker: $J_{1'2'}$ is approximately 8 Hz for C-2′-*endo* sugar pucker and $J_{1'2'}$ is less than or equal to 1 Hz for C-3′-*endo* sugar pucker (Altona, 1982). The small *J* value for C-3′-*endo* puckers results in no observed H1′–H2′ COSY cross peaks in *A*-form stem regions. Identification of C-2′-*endo* sugars is obvious, since they do have strong H1′–H2′ COSY cross peaks. The assignments of H8/H2/H6/H5 base protons and H1′/H2′ sugar protons are the most reliable in a complex RNA spectrum.

Further assignment of the sugar region in DNA is relatively easy since H3′ protons resonate separately from H2′/H2″ and H4′/H5′/H5″ protons. In RNA the situation is more complex. H3′ resonances can be identified by intranucleotide H8/H6 ↔ H3′ cross peaks (2·8 Å in *A*-form nucleotides) as well as internucleotide 3′ (H8/H6) ↔ 5′ (H3′) connectivities (approx. 3 Å in *A*-form helices). For long mixing times, spin diffusion intrasugar H1′–H3′ cross peaks can be observed. Intrasugar H1′–H4′ cross peaks can sometimes be observed for C-3′-*endo* or C-2′-*endo* sugars (distance 3·4 Å). This cross peak should be very strong for intermediate sugar puckers (i.e. pseudorotation angle $P = 90°$, distance 2·2 Å).

The H2′/H3′/H4′/H5′/H5″ region of a DQF-COSY spectrum of an RNA molecule is crowded. This region in principle contains H2′ to H3′, H3′ to H4′, H4′ to H5′/H5″ and H5′ to H5″ cross peaks. H5′ and H5″ are usually degenerate, but can be identified by a strong COSY cross peak ($J = -15$ Hz) when they are non-degenerate. If the torsion angle γ is approximately g^+, as found in both *A*-form and *B*-form helices (Saenger, 1984) the H4′ to H5′/H5″ couplings are small ($J < 3$ Hz). In non-standard conformations ($\gamma = g^-$ or *trans*) the coupling to the H5′ or H5″ can be quite large ($J > 10$ Hz). The major cross peaks in this region for RNA are H3′–H4′ connectivities, since $J_{3'4'}$ is 8 Hz for C-3′-*endo* sugars (Altona, 1982). H2′–H3′ cross peaks are much weaker since $J_{2'3'}$ is only 5 to 6 Hz; the intensity is further attenuated by the complex antiphase fine structure of this cross peak arising from the strong passive coupling to H4′ and 3′ phosphorus. The antiphase nature of COSY cross peaks makes linewidth a critical factor in determining the amount of cancellation of cross peak components. This is a problem with larger RNA molecules; even in the DQF-COSY of the 26 nucleotide molecule, very few H2′–H3′ cross peaks are observed, and only a small fraction of H3′–H4′ cross peaks are resolved.

A DQ spectrum contains *J*-coupling information similar to that in a DQF-COSY. *J*-coupled cross peaks will occur at the 2 frequencies of the coupled spins in t_2, at the same DQ frequency $(\nu_1+\nu_2)$ in t_1. The advantage of DQ spectroscopy is that there is no diagonal, so that connectivities close to the diagonal in a DQF-COSY are easily resolved in a DQ spectrum. The excitation of DQ transitions depends on the DQ preparation period (τ), with maximal excitation for $\tau = (2J)^{-1}$ in an AX spin system. For DQ spectra with $\tau = 50$ ms, H3′–H4′ DQ transitions are maximally excited. In practice, DQ spectra can be useful in further assigning sugar spin systems when used in conjunction with the other methods described above.

3. Results and Discussion

(a) *Characterization of global structure*

(i) *Enzymatic and chemical mapping*

Single and double-strand-specific enzymes and chemical reagents were used to probe the structure of the oligonucleotide in buffer containing 5 mM-Mg^{2+} and 60 mM-Na^+. At and below 55 °C, the oligonucleotide is resistant to single-strand-specific nuclease S_1 cleavage except in the loop regions of the pseudoknot. At 4 °C, strong cleavage occurs after the second, third and fourth nucleotides in loop 2 (U_{17}, U_{18} and U_{19}) (Fig. 2). At 22 °C and 37 °C, weaker cleavage also occurs in loop 1 and after the other nucleotides in loop 2 (data not shown). The molecule is completely resistant to cleavage by RNase T_1 (specific for single-stranded G) at temperatures at and below 37 °C. Nucleotides in loop 2, which bridges the minor groove, are more accessible to single-strand-specific probes than those in loop 1; this is also observed for the pseudoknot formed by TYMV RNA (van Belkum *et al.*, 1988).

RNase V_1, which is specific for helical or stacked regions of RNA structure (Auron *et al.*, 1982),

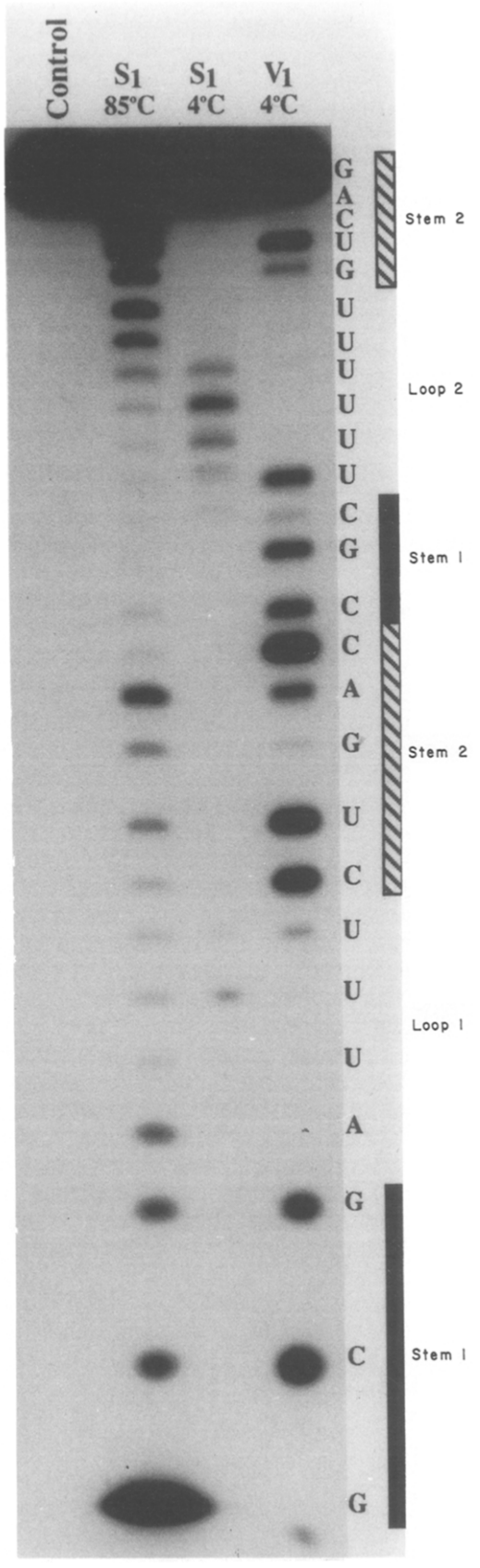


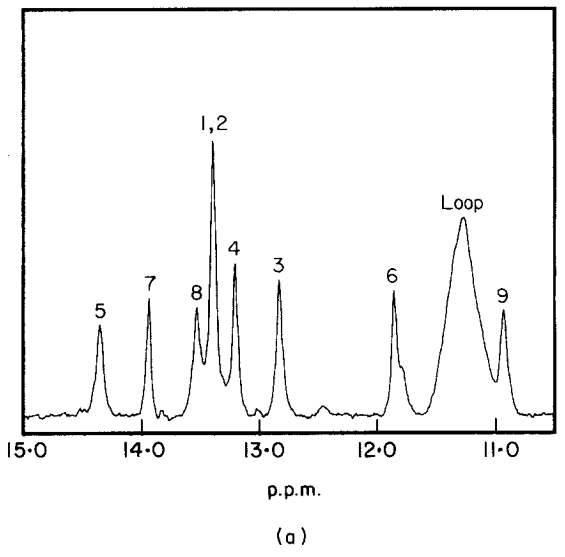
Figure 2. Autoradiograms of gel electrophoretic analysis of single-strand-specific nuclease S_1 and double-strand-specific RNase V_1 digestions of the oligonucleotide. RNA was ^{32}P-labeled at the 5′ end. Lane 1, control, no enzyme; lane 2, nuclease S_1 cleavage at 85°C in 5 mM-NaCl, 5 mM-Mes (pH 6·3) (denaturing conditions); lane 3, nuclease S_1 cleavage at 4°C in 5 mM-$MgCl_2$, 60 mM-NaCl, 5 mM-Mes (pH 6·3); lane 4, RNase V_1 cleavage at 4°C in 5 mM-$MgCl_2$, 60 mM-NaCl, 5 mM-Tris·HCl (pH 8·1). Each radioactive band corresponds to cleavage at the nucleotide indicated by the sequence at the right of lane 4. The stem and loop regions of the pseudoknot are also indicated.

cleaves the oligonucleotide in the proposed stem regions of the pseudoknot (Fig. 2). Cleavage is observed after each of the nucleotides in stem 1 although cleavage is weak after G_1 and C_{15}. In stem 2, cleavage occurs following each nucleotide on the 5′ side of the stem (C_8 through C_{12}). Weak cleavage after G_{10} cannot be correlated to the pseudoknot structure, since a hairpin containing only the stem 2 region shows the same sequence-specific pattern of RNase V_1 cleavage (data not shown). On the 3′ side of stem 2, cleavage is observed only after residues G_{22} and U_{23}. Lack of cleavage near the 3′ terminus of stem 2 is consistent with RNase V_1 cleavage of other hairpins and duplexes. RNase V_1 does not cleave single-stranded control molecules; cleavage after loop nucleotides U_7 and U_{16} is correlated with stacking.

The solvent accessibilities of the three adenosine residues in the oligonucleotide were compared over a range of temperatures using diethyl pyrocarbonate (DEP) (see the following paper; Fig. 2). The carboxyethylation reaction at the N-7 position is sensitive to stacking (Peattie & Gilbert, 1980). Both A_{11} and A_{25}, residues in stem 2 of the pseudoknot structure, show the same extent of modification throughout the temperature range from 4°C to 85°C. As expected for nucleotides in duplex regions, A_{11} and A_{25} are only slightly reactive at low temperatures. The mid-point of the percentage cleavage *versus* temperature curve is approximately 65°C, roughly the same as the optically observed melting temperature (see the following paper). Below 45°C, A_4, the first nucleotide of loop 1, is about one third as accessible to modification as it is in the denatured oligonucleotide at 85°C. The mid-point of the melting transition of A_4 is approximately 5 to 10 deg.C lower than that of A_{11} and A_{24}. The DEP modification data, like the cleavage patterns obtained using enzymatic probes, support the existence of stem 2. The increased DEP modification of A_4 as compared with the adenosine residues in stem 2 suggests that this residue is part of a loop region; however, A_4 also shows a melting transition so it is partially protected from solvent by the folding of the pseudoknot. Loop 1 is proposed to bridge stem 2 across the major groove of the helix.

(ii) *Exchangeable n.m.r. spectrum*

The n.m.r. spectrum of the exchangeable imino protons in 5 mM-$MgCl_2$, 50 mM-NaCl, 10 mM-sodium phosphate (pH 6·4) at 4°C is shown in Figure 3. The total of eight base-paired imino resonances is consistent with pseudoknot formation. The two resonances farthest downfield, at 13·92 and 14·35 p.p.m.,



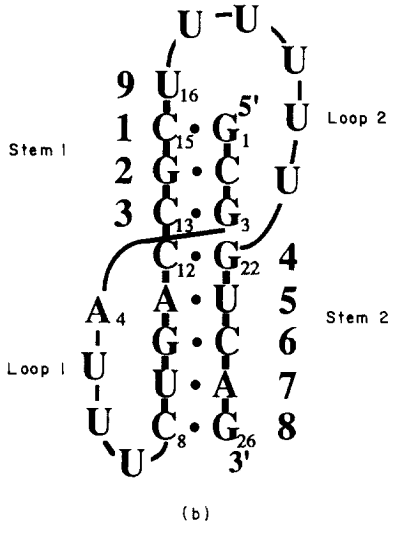


Figure 3. (a) n.m.r. spectrum of the exchangeable imino protons of the oligonucleotide in 5 mM-$MgCl_2$, 50 mM-NaCl, 10 mM-sodium phosphate (pH 6·4) at 23 °C. (b) Schematic folding of the oligonucleotide; numbered base-pairs corresponding to resonances in the spectrum.

resonate in a chemical shift region characteristic of imino protons from A·U base-pairs. The resonances below 11·5 p.p.m. are most likely due to non-hydrogen-bonded imino protons on uridine residues. The imino resonances were assigned using one-dimensional NOE measurements. An NOE is observed between protons that are less than 4·5 to 5·0 Å apart. Since the inter-base-pair distance in both *A*-form and *B*-form helices is approximately 3·5 to 3·8 Å, an NOE is expected from an imino resonance of one base-pair and the imino protons corresponding to the two flanking base-pairs. In addition, uridine imino protons involved in a Watson–Crick A·U base-pair will have a strong NOE to the adenine H2, which is separated by only 1·9 Å from the imino proton. The assignments of the imino protons are summarized in Figure 3(b) and Table 1.

Table 1

Assignment of the exchangeable imino protons in 5 mM-$MgCl_2$, 50 mM-NaCl, 10 mM-sodium phosphate (pH 6·3) at 5·0 °C

	Imino resonance†	Base-pair	p.p.m.
Stem 1	1	$G_1 \cdot C_{15}$	13·38
	2	$C_2 \cdot G_{14}$	13·38
	3	$G_3 \cdot C_{13}$	12·83
Loop 1		A_4 to U_7	Approx. 11·3
Stem 2	8	$C_8 \cdot G_{26}$	13·52
	7	$U_9 \cdot A_{25}$	13·93
	6	$G_{10} \cdot C_{24}$	11·85
	5	$A_{11} \cdot U_{23}$	14·35
	4	$C_{12} \cdot G_{22}$	13·20
Loop 2	9	U_{16}	10·93
		U_{17} to U_{21}	Approx. 11·3

† Numbers correspond to resonances from either G(H1) or U(H3) in the exchangeable spectrum shown in Fig. 3.

The observed pattern of NOE connectivities can only arise if all the imino resonances originate from a single structure rather than an equilibrium between the two possible hairpin structures. The 5′ hairpin, corresponding to formation of only stem 1, has a completely different imino spectrum (Puglisi *et al.*, 1990). The 3′ hairpin structure, corresponding to stem 2, gives rise to imino resonances corresponding to peaks 4, 5, 6, 7 and 8. The NOE between imino protons 3 and 4 indicates that both stems are present within the same structure. Thus, the imino spectrum does not originate from either of the two stable hairpin structures or an equilibrium mixture of the two. No other relatively stable secondary structures can be predicted. Therefore, the oligonucleotide forms a structure with eight continuous stacked base-pairs, consistent with the pseudoknot model of Pleij *et al.* (1985).

The imino protons in nucleic acids exchange with solvent due to base-pair opening and base catalysis. As the temperature is raised, imino resonances broaden due to an increased exchange rate. Normally, imino resonances corresponding to terminal base-pairs in a duplex broaden first because of fraying at the ends of helices. Figure 4 shows the imino spectra as a function of temperature; the optical melting temperature under these conditions is 63 °C (see the following paper). At 5 °C, the resonances arising from base-pairs are all relatively sharp. The broad peak between 11·0 and 11·5 p.p.m. is due to non-hydrogen-bonded loop imino protons, which are not well protected from solvent. The imino proton assigned to U_{16}, which gives the sharper resonance at 10·93 p.p.m., is better protected from solvent. At 25 °C, all the non-hydrogen-bonded imino resonances have broadened into the baseline. Resonance 8 (terminal $C_8 \cdot G_{26}$ pair) has also broadened significantly. As the temperature is raised to 35 °C, resonances 1, 2 and 3 broaden. These resonances correspond to stem 1 in the pseudoknot structure. In stem 2, resonance 7 ($U_9 \cdot A_{25}$ base-pair) also broadens. By 45 °C, only the

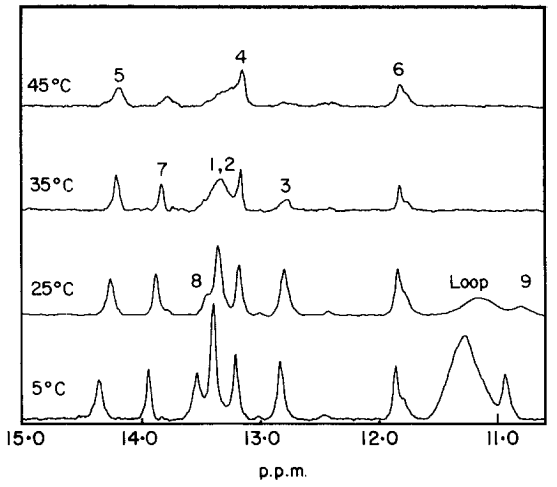


Figure 4. n.m.r. spectra of the exchangeable imino protons as a function of temperature of the oligonucleotide in 5 mM-$MgCl_2$, 50 mM-NaCl, 10 mM-sodium phosphate (pH 6·4). Individual resonances are numbered as for Fig. 3.

imino resonances 4, 5 and 6 from the core of stem 2 remain. Resonance 4, which corresponds to the terminal $C_{12}\cdot G_{22}$ base-pair of stem 2, remains relatively sharp up to 50°C, whereas the imino resonance from $G_3\cdot C_{13}$ pair 3 has completely broadened. The sequential order of broadening of the imino resonances could be due to differences in either the thermodynamic stability or the kinetic behavior of the two stems. Data discussed in the following paper support the latter.

(b) *Detailed structure of the pseudoknot*

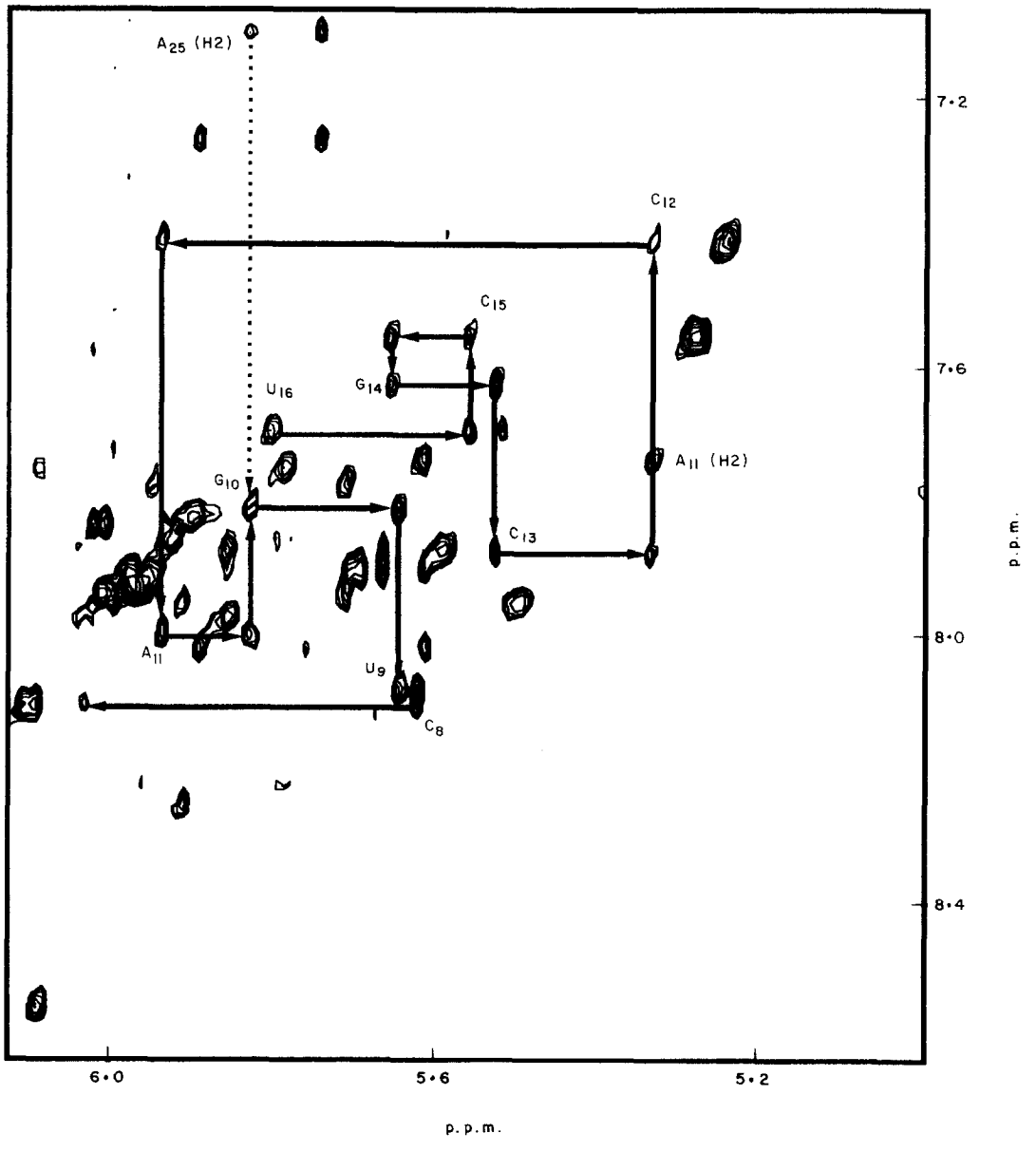
(i) *Assignment of the non-exchangeable n.m.r. spectrum*

The region between 7·0 p.p.m. and 8·6 p.p.m. contains resonances from AH8, GH8, CH6 and UH6 as well as AH2 protons. For the oligonucleotide described here, this region should contain 29 resonances. UH5, CH5 and sugar H1′ protons resonate in the region between 5·0 p.p.m. and 6·2 p.p.m. The spectrum can be significantly simplified in these two regions through selective deuteriation. Deuteriation of pyrimidine residues at the 5 position is experimentally straightforward for either the nucleotides or the DNA or RNA strands (Wataya & Hayatsu, 1972; Brush *et al.*, 1988). Since 11 of the 26 nucleotides of the oligonucleotide are uridine residues, UPT deuteriated at the 5 position was incorporated using T7 RNA polymerase. The region between 5·0 p.p.m. and 6·2 p.p.m. is significantly simplified by the elimination of 11 UH5 resonances. Deuteriation of UH5s also improves the appearance of the 7·0 to 8·6 p.p.m. region of the spectrum because the UH6 resonances are singlets; the presence of deuterium in the H5 position eliminates the splitting of these resonances.

The first step in the assignment of the non-exchangeable spectrum was the identification of uridine and cytidine H6 resonances. The COSY spectrum of the deuteriated oligonucleotide exhibits six cross peaks due to H5–H6 *J*-coupling (not shown), and allowed the unambiguous identification of the six cytidine H6 and H5 resonances. Figure 5 shows the H8/H6 to H1′/H5 region of the 400 millisecond mixing time NOESY spectrum. Strong cross peaks arise at the chemical shifts of the CH5–H6 protons in the NOESY spectrum due to the short fixed 2·45 Å. This NOESY spectrum was compared with the NOESY spectrum of the non-deuteriated molecule under identical conditions. Strong cross peaks missing in the NOESY spectrum of the deuteriated molecule, due to UH5–H6 interactions, permitted the identification of 6 of 11 UH6 resonances; the other uridine resonances were not sufficiently resolved to allow assignment.

The assignment of the pyrimidine residues aids in the confirmation of the aromatic to H1′ assignment paths shown in Figure 5; the paths are consistent with both sequence and secondary structure. Connectivity path 1 (Fig. 5(a)) encompasses nine nucleotides: 3′ U-C-R-C-C-R-R/U-U-C 5′ (R is a purine). Only one region of sequence is consistent with this assignment path: 3′ U_{16}-C_{15}-G_{14}-C_{13}-C_{12}-A_{11}-G_{10}-U_9-C_8 5′. This path spans the two-stem regions of the pseudoknot. Connectivity path 2 (Fig. 5(b)) encompasses five nucleotides: 3′ R-R-C-U/R-R 5′. This path corresponds to the 3′ terminus: 3′ G_{26}-A_{25}-C_{24}-U_{23}-G_{22} 5′. Both regions are double-stranded stems on the basis of assignment of the imino spectrum and chemical and enzymatic mapping. One-dimensional NOEs from the two A·U imino protons allowed assignment of the resonance at 7·02 p.p.m. as A_{25}(H2) and that at 7·71 p.p.m. as A_{11}(H2). Both AH2 positions have NOE connectivities to two H1′ protons. A_{25}(H2) has NOEs to G_{26}(H1′) and G_{10}(H1′); A_{11}(H2) has NOEs to C_{12}(H1′) and C_{24}(H1′). These NOEs from AH2 to the 3′-neighboring H1′ on the same strand and the H1′ proton across strand in the 5′ direction are exactly those expected for an *A*-form helical geometry. Both assignment paths can be confirmed using the 3′ H8/H6 ↔ 5′ H2′ and H1′–H2′ NOE cross peaks (see below).

The final connectivity path (Fig. 5(b)) involves the stem region at the 5′ terminus. The purine at 7·77 p.p.m. is assigned at G_3(H8), since it has a NOE to its own H1′(5·94 p.p.m.), and to the C_2(H1′) (5·71 p.p.m.). The C_2(H6) (7·93 p.p.m.) was the only unassigned CH6–CH5 cross peak in the COSY spectrum. The C_2(H5) (5·96 p.p.m.) has a second weak cross peak at 5·50 p.p.m. due to a small amount of 5′ hairpin structure present (Puglisi *et al.*, 1990). The 5′-terminal G_1(H8) was identified by incorporation of d^8GMP during transcription. Since T7 RNA polymerase preferentially primes synthesis with GMP rather than GTP, the d^8G is selectively incorporated at position 1 (Sampson & Uhlenbeck, 1988). G_1(H8) (8·24 p.p.m.) has a NOE to its own H1′ (5·90 p.p.m.) and to the C_2(H1′).



(a)

Fig. 5.

NOE connectivities are also observed between loop 1 nucleotides (Fig. 5(b)). The farthest downfield aromatic resonance at 8·53 p.p.m. is assigned as A_4(H8) and the resonance at 8·21 p.p.m. is A_4(H2). These assignments were confirmed by the presence of both these resonances after incorporation of d^8GTP (data not shown). The A_4(H8) has a strong cross peak to the A_4(H1′) at 6·09 p.p.m. The UH6 at 7·74 p.p.m. has a NOE to the A_4(H1′). This proton, assigned as the U_5(H6), also has a NOE to the U_5(H1′) at 5·78 p.p.m. Further assignments are summarized in Table 2.

(ii) *Nucleotide conformation: sugar pucker*

The individual nucleotide conformations can be determined using both *J*-coupling and NOE information. Scalar (*J*) coupling information is used to characterize the sugar conformation as north (approx. C-3′-*endo*, $P = 18°$) or south (approx. C-2′-*endo*, $P = 162°$). For smaller RNA molecules, the sugar conformation can be more precisely determined using additional coupling information (i.e. H2′–H3′, H3′–H4′) (Davis *et al.*, 1990); this information is limited in the case of the 26 nucleotide molecule studied here. The two families of puckers represent points of minimum energy for the pseudorotation phase angle, *P* (Olson & Sussman, 1982; Olson, 1982). Since the rate of interconversion between these two conformers is fast on the n.m.r. time-scale, the value of $J_{1'2'}$ for a nucleotide is an average of $J_{1'2'}$ for north ($J_{1'2} < 1$ Hz) and south ($J_{1'2'} = 8{\cdot}0$ Hz) conformers weighted by the relative population of each. An empirical formula for determining the percentage of each conformer has been described (van den Hoogen, 1988):

$$\%N = 114{\cdot}9 - 14{\cdot}5 \cdot (J_{1'2'}). \qquad (1)$$

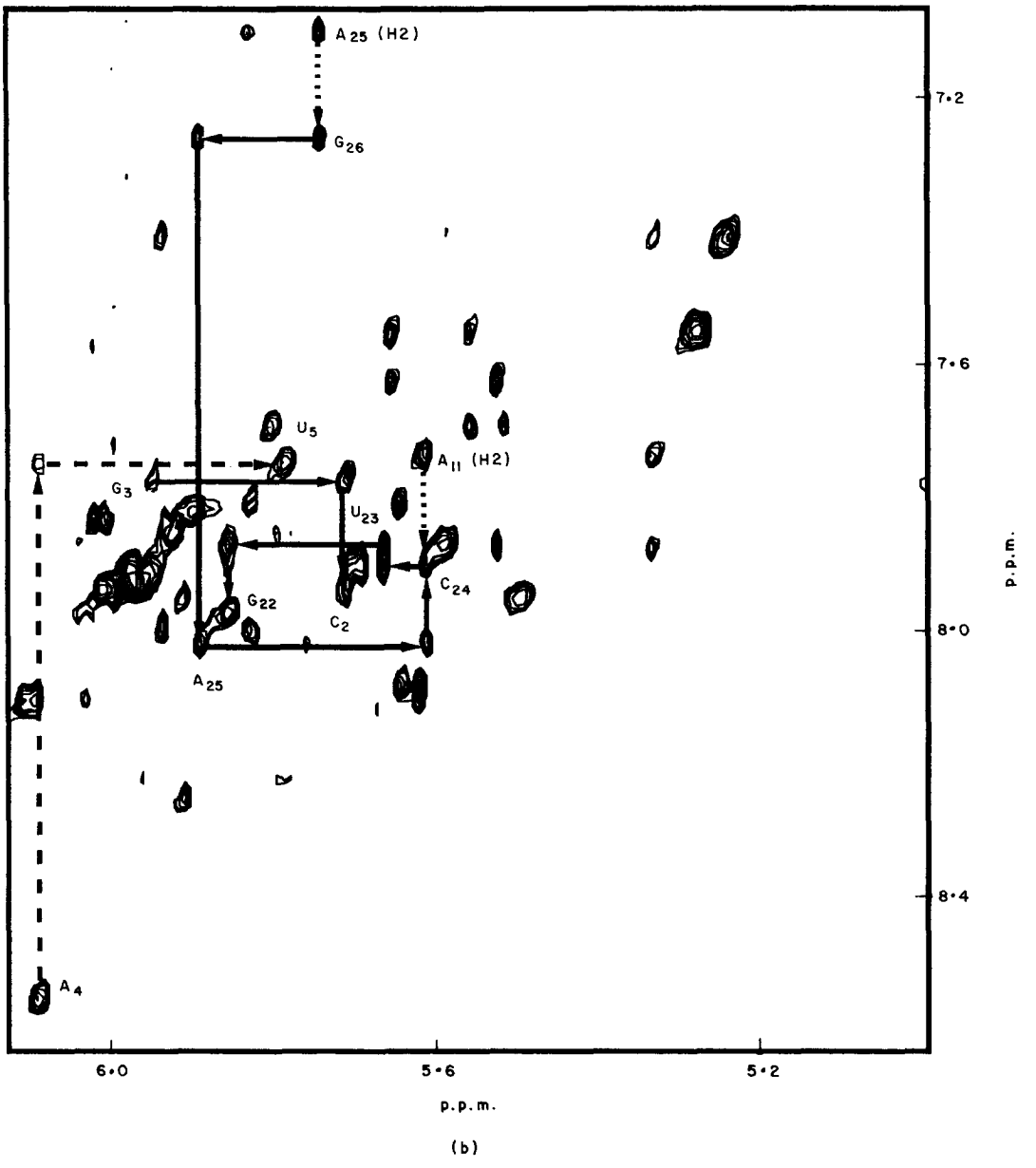


Figure 5. A portion of the NOESY spectrum of d^{5}U-incorporated oligonucleotide showing NOEs between H8/H6/H2 and H5/H1′ protons. NOEs between H8/H6 and H1′ protons are indicated by continuous lines and those between AH2 and H1′ are indicated by dotted lines. Labeled NOE cross peaks indicated a NOE from that nucleotide's aromatic proton to its own H1′ sugar proton. Conditions were: 5 mM-$MgCl_2$, 50 mNaCl, 10 mM-sodium phosphate (pH 6·4) at 23°C. The mixing time was 400 ms. (a) H8/H6 to H1′ NOE connectivity pathway 1, which spans nucleotides U_{16} to C_8. (b) H8/H6 to H1′ NOE connectivity pathway 2 spans G_{26} to G_{22} and from G_3 to C_2. Note that there is no NOE between G_{22}(H8) and G_3(H1′). The broken line indicates the NOEs between A_4 and U_5 in loop 1.

The H1′–H2′ region of the DQF-COSY is shown in Figure 6. Certain nucleotides have observable H1′–H2′ cross peaks. The unresolved cross peaks correspond to H1′–H2′ cross peaks for loop uridine residues. The strong resolved cross peaks labeled in Figure 6 also belong to loop uridine residues; one has been assigned to U_{16}; the other corresponds to an unassigned loop residue. Weaker H1′–H2′ cross peaks are also observed for U_5, G_{22}, A_{25} and G_{26}. These cross peaks have $J_{1'2'}$ less than 7 Hz; the smaller coupling results in cancellation of antiphase components in the cross peak, and attenuated intensity (Widmer & Wüthrich, 1987; Bax & Lerner, 1988). The A_4(H1′–H2′) cross peak is very weak and can only be observed in contour plots at very low levels in the COSY. A larger linewidth for these resonances could explain the reduced intensity of this cross peak in relation to the G_{26} cross peak, which also has a small coupling. The sugar conformations (%N) calculated from these cross peaks using equation (1) are listed in Table 3. The sugar pucker was estimated as greater than 85% north in the absence of a detectable H1′–H2′ cross peak (i.e. $J_{1'2'}$ values less than 2 Hz).

(iii) *Nucleotide conformation: glycosidic torsion angle*

The glycosidic torsion angle (χ) can be approximately determined for many nucleotides in the

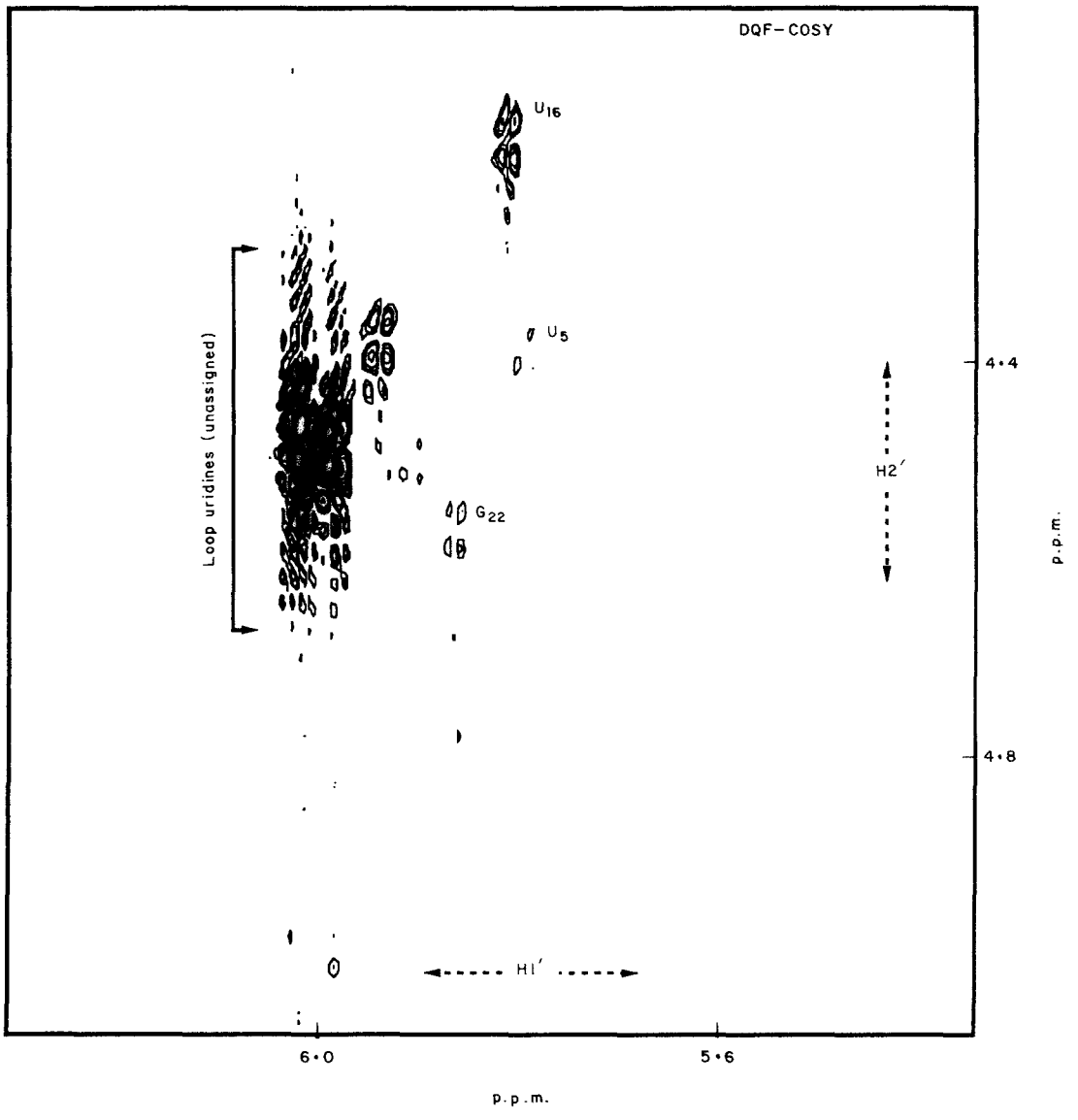


Figure 6. Portion of the double-quantum filtered COSY spectrum of d^5U-incorporated oligonucleotide showing cross peaks between sugar H1′ and H2′ protons. Cross peaks are only observed for nucleotides with a significant population of S-type sugars; these cross peaks are labeled. The unresolved peaks between 6 and 4·5 p.p.m. are due to unassigned loop nucleotides.

oligonucleotide. The method employed here makes use of the known sugar pucker (determined as described above) and the intranucleotide distances (determined from the NOESY). Contour plots that relate both pseudorotation phase angle and glycosidic torsion angle to values of intranucleotide interproton distances were presented by Wüthrich (1986). If the sugar adopts a north conformation, as do most of the sugars in the stem region, and the value of χ is *anti*, then a short H8 or H6 to H3′ and a long (approx. 3·8 Å) H8 or H6 to H2′ distance results. Assuming a C-3′-*endo* sugar pucker ($P = 18°$), an approximation, since the pseudorotation phase angle is not known precisely, the value of χ can be estimated from the H8/H6–H3′ distance. The nucleotide conformations determined in this way are listed in Table 3.

Interproton distances (listed in Table 4) were estimated from a single NOESY experiment at a relatively short mixing time (150 ms). This method will yield accurate distances provided that the two-spin approximation is valid and that all interproton vectors are tumbling with the same correlation time. Cases where the two-spin approximation is not valid are discussed below. The approximation of a single correlation time can be checked by comparing the integrated intensities of the H5–H6 cross peaks, which arise from a fixed interproton distance of 2·45 Å. Any differences (at short mixing times) in cross peak intensity are due to differences in the motional behavior of the different nucleotides. The six CH5–CH6 cross peaks from the 150 millisecond NOESY used as references all had integrated intensities within 10 to 20% of each other. A single correlation time is thus a reasonable approximation. At the low concentrations (approx. 1 mM) of RNA used in these two-dimensional experiments, 150 milliseconds was the shortest mixing time to

Table 2

Assignments of non-exchangeable protons in 5 mM-$MgCl_2$, 50 mM-NaCl, 10 mM-sodium phosphate (pH 6·4) at 23°C

Nucleotide	H8/H6	H2/H5	H1′	H2′	H3′	H4′
G_1	8·24		5·90	4·93	4·83	4·61
C_2	7·93	5·96	5·71	4·43		
G_3	7·77		5·94	4·97	4·51	
A_4	8·53	8·21	6·09	4·93	4·83	4·65
U_5	7·74	5·70	5·78	4·38		
U_6 to U_7	—	—	—	—	—	—
C_8	8·09	6·10	5·62	4·45	4·66	
U_9	8·07	5·63	5·64	4·68		
G_{10}	7·80		5·82	4·53	4·74	
A_{11}	7·99	7·74	5·93	4·55	4·66	
C_{12}	7·42	5·24	5·33	4·37		
C_{13}	7·87	5·58	5·52	4·49	4·39	
G_{14}	7·62		5·65	4·45	4·11	
C_{15}	7·55	5·27	5·55	4·27		
U_{16}	7·69	5·51	5·80	4·16	4·56	4·04
U_{17} to U_{21}	—	—	—	—	—	—
G_{22}	7·96		5·85	4·55	4·63	4·23
U_{23}	7·87		5·66	4·60	4·75	
C_{24}	7·90	5·70	5·61	4·58	4·29	
A_{25}	8·01	7·11	5·89	4·51	4·70	
G_{26}	7·27		5·74	3·96	4·22	4·42

give an adequate signal-to-noise ratio. This mixing time will definitely lead to spin diffusion (Keepers & James, 1984; Chazin *et al.*, 1986). The magnetization transferred by spin diffusion will result in an underestimation of long interproton distances and an overestimation of short distances (<2·4 Å). Most distances between 2·5 Å and 3·4 Å seemed approximately correct, however, as shown by the reasonable values obtained for the relatively short (2·5 to 2·8 Å) H1′–H2′ distance (see below).

Table 3

Nucleotide conformations determined by n.m.r.

Nucleotide	Glycosidic torsion angle χ (°)†	Sugar conformation (% north)‡
G_1	—	—
C_2	*Anti*	>85
G_3	−155±25	>85
A_4	Between −60 and 0	60
U_5	*Anti*	<20
U_6 to U_7	—	—
C_8	−160±25	>85
U_9	−170±25	>85
G_{10}	−175±25	>85
A_{11}	−170±25	>85
C_{12}	−170±25	>85
C_{13}	−165±25	>85
G_{14}	−170±25	>85
C_{15}	−165±25	>85
U_{16}	*Anti*	<5
U_{17} to U_{21}	—	—
G_{22}	*Anti*	<20
U_{23}	−165±25	>85
C_{24}	−180±25	>85
A_{25}	−155±25	—
G_{26}	*Anti*	60

† The glycosidic torsion angle (χ) was estimated graphically using both H8/H6–H1′ and H8/H6–H3′ NOEs. Errors in χ are estimated assuming errors of ±0·5 Å in distances.

‡ Sugar conformation was estimated using H1′–H2′ coupling constants determined from the DQF-COSY using eqn (1) (see the text). Sugar molecules were assumed to be in fast equilibrium between N ($P = 18°$) and S ($P = 160°$) conformers.

Table 4

Intranucleotide proton–proton distances (Å) calculated from 150 ms NOESY

Nucleotide	H8/H6↔H1′	H8/H6↔H2′	H8/H6↔H3′	H1′↔H2′
G_1				
C_2	3·5	—	—	2·5
G_3	3·3	3·6	2·5	—
A_4	(2·9)	(3·0)	(3·1)	2·7
U_5	3·2	(2·5)	(3·0)	2·9
U_6 to U_7	—	—	—	—
C_8	3·3	—	2·6	2·6
U_9	3·3	—	2·9	2·7
G_{10}	3·4	—	3·1	2·6
A_{11}	3·6	—	3·0	2·5
C_{12}	3·5	—	2·8	2·6
C_{13}	3·6	—	2·7	2·6
G_{14}	3·4	—	3·0	2·5
C_{15}	3·7	3·7	2·8	2·6
U_{16}	(3·1)	<2·4	(2·5)	2·8
U_{17} to U_{24}	—	—	—	—
G_{22}	(3·0)	(2·4)	(3·0)	2·8
U_{23}	3·4	—	2·7	2·7
C_{24}	3·5	—	3·2	2·6
A_{25}	3·4	—	2·7	2·5
G_{26}	(3·2)	(3·1)	(2·6)	2·5

Distances in parentheses are the least accurate, since these nucleotides have significant populations of both N and S conformers; this results in $1/r^6$ averaging of the distance determined by NOE. Approximate error limits on other distances are ±0·4 Å.

Nucleotides C_8 through C_{15} all adopt north sugar conformations as shown by the lack of an observable H1′–H2′ COSY cross peak for any of these residues. These nucleotides all have intranucleotide proton–proton distances (2·7 to 3·1 Å) consistent with an *anti* value of χ. The approximate values of χ determined for these nucleotides range between −160° and −175°; the value of χ in an *A*-RNA duplex determined from fiber diffraction is −170° (Saenger, 1984). U_{16}, which marks the end of this connectivity path, has an S-type (C-2′-*endo*) sugar conformation. The values calculated for both the H6–H2′ distance (<2·5 Å) and H6–H3′ distance (<2·7 Å) are short. No glycosidic torsion angle is consistent with an S-type sugar and both short H6–H2′ and H6–H3′ distances. The apparent short distance between H6 and H3′ is probably due to spin diffusion from the strong H6–H2′ connectivity (the distance could be approximately 2 Å). The short H6 to H2′ distance suggests that this nucleotide has an *anti* value of χ.

The sugar conformations at the 3′ end of the sequence are affected by terminal end fraying. Both G_{26} and A_{25} have small (<3 Hz) H1′–H2′ coupling. These nucleotides have sugar puckers in fast exchange between N and S-type sugar conformations. G_{26} has an especially large (approx. 40%) population of S-type sugar. The distance deter-

mined by NOE is not a simple average of the distances in the two conformations, but is weighted by the $1/r^6$ dependence towards the shorter distance. For example, G_{26}(H8) has short apparent distances to both the H2′ (3·1 Å) and H3′ (2·6 Å). Nonetheless, these distances are consistent with an *anti* value of χ. A_{25} has H8–H1′ (3·4 Å) and H8–H3′ (2·7 Å) distances consistent with an *anti* glycosidic torsion angle. The population of S-type sugar conformation must be low, since there is not a strong H8–H2′ NOE. C_{24} and U_{23} both have pure N-type sugar conformations; the H6–H1′ and H6–H3′ distances give values of χ consistent with *A*-form geometry. G_{22} mostly adopts a south sugar conformation (approx. 80% S), despite its location in a duplex region (see discussion below). The calculated H8–H1′ distance is short (3·0 Å), most likely due to overlap of cross peaks and/or spin diffusion from the strong H8-H2′ cross peak. The calculated H8–H3′ distance (3·0 Å) is short due to the population of N-type sugar present. G22 has an *anti* value of χ.

G_1 has an N-type pucker. C_2 has an N-type sugar conformation; the intranucleotide H6–H1′ distance (3·5 Å) is consistent with an *anti* value of χ. G_3 also has a pure N-type sugar conformation and both H8–H1′ (3·4 Å) and H8–H3′ (2·5 Å) are consistent with an *A*-form value of χ ($-155°$). Residue A_4 may have a non-standard nucleotide conformation. The sugar conformation is mostly N-type. The H8–H1′ distance is quite short (2·8 Å), as are both the H8–H2′ (3·0 Å) and H8–H3′ (3·1 Å). Spin diffusion does not seem to be a significant problem with this residue since both the H1′–H2′ (2·7 Å) and H1′–H4′ (2·9 Å) are consistent with the N conformation. The intranucleotide distances are roughly consistent with a χ value in the high *anti* range ($-60°$ to 0°); however, this nucleotide may be in equilibrium among various conformations. The sugar conformation of U5 is a mixture of N and S-type puckers. The H6–H1′ distance (3·2 Å) is consistent with an *anti* conformation. The conformation of the remaining loop residues cannot be determined due to spectral overlap and the lack of a clear connectivity path. Many of the loop nucleotides must have some population of S-type sugars, as shown by the number of strong overlapping H1′–H2′ cross peaks in the COSY.

(iv) *Strand conformation*

The analysis of nucleotide conformations demonstrates that the residues composing the stem regions adopt C-3′-*endo* sugar puckers with glycosidic torsion angles consistent with *A*-form helical geometry. The internucleotide NOEs are also consistent with an *A*-form helical structure in the stem regions. The strong 3′ H8/H6 ↔ 5′ H2′ NOEs that are expected for *A*-form are observed. These internucleotide NOEs are difficult to interpret in terms of distances. The expected 3′ H8/H6 ↔ 5′ H2′ distance for *A*-form is extremely short (2·0 Å); spin diffusion will cause an overestimation of these

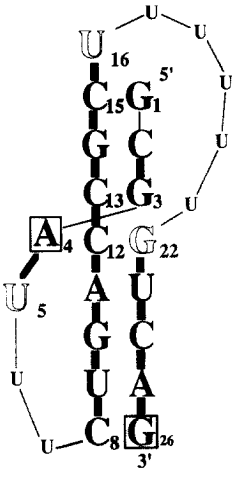


Figure 7. Summary of the n.m.r. data on the oligonucleotide. Each of the nucleotides indicated in large type adopts an *anti* glycosidic torsion angle. Sugar conformation is indicated by letter type: C-3′-*endo*, bold-face type; C-2′-*endo*, outline; equilibrium between conformations, box. Internucleotide NOEs are represented by the thickness of the line connecting 2 nucleotides. Thick lines represent strong 3′ H8/H6 ↔ 5′ H2′ NOEs, consistent with *A*-form stacking. Thin lines indicate no observed NOE interaction. No NOE was observed between the non-exchangeable protons of G_{22} and G_3. Loop regions are largely undefined (smaller letters).

distances unless very short mixing times (<50 ms) are used. The intranucleotide H8/H6–H2′ distances for this molecule derived from the 150 millisecond mixing time NOESY range from 2·3 to 2·6 Å; these distances are longer than those usually observed in *A*-form helices, consistent with the presence of spin diffusion. For certain nucleotides, the 3′ H8/H6–5′ H3′ NOE could also be observed, also in agreement with the overall *A*-form geometry. The presence of relatively intense AH2 ↔ H1′ NOEs further confirm the *A* conformation of the stem regions; these NOEs would not be as intense in a *B*-form helix. The AH2–H1′ NOEs actually appear more intense than predicted by classical *A*-form geometry. Chou *et al.* (1989) observed similar effects in an RNA oligonucleotide; they proposed that sequence-specific variations are responsible for a reduction in cross-strand distances, since the calculated distances are derived from fiber diffraction studies. Internucleotide NOEs have only been interpreted in a qualitative manner; no attempt has been made to gain specific geometries of each dinucleotide step. The NOE data are illustrated schematically in Figure 7.

The n.m.r. data described above lead to the following conclusions concerning the structure of the pseudoknot. Residues C_8 to C_{15} form a continuous *A*-form helix. There seems to be little perturbation of the structure along the strand at the junction of stem 1 and stem 2. The NOE connectivity is continuous from C_{12} to C_{13}; the H5–H5 NOE observed between these two nucleotides at 400 milliseconds mixing time confirms that these

two bases are stacked in an *A*-like fashion. The eight base-pairs form a continuous helix with no break in NOE connectivity between imino resonance 3 and resonance 4 at the junction between the two stem regions. U_{16} is stacked upon C_{15}, continuing the *A*-form connectivity into loop 2. The non-hydrogen-bonded imino proton of U_{16} at 10·93 p.p.m. gave a NOE to the (degenerate) imino resonances belonging to base-pairs 1 and 2. It appears that U_{16} marks the turning point of loop 2 from a continuous stacked conformation; there are no NOEs between U_{16} and U_{17}, and U_{16} adopts a C-2′-*endo* conformation.

The 3′ side of stem 2 (G_{22} through G_{26}) also forms a continuous *A*-like helix. There is more heterogeneity in the sugar conformation at the 3′ end of stem 2 than in the nucleotides to which they are base-paired. Both A_{25} and G_{26} have significant populations of S-type sugars, while U_9 and C_8 have pure N-type sugar conformation. On the 5′ side of stem 1 there is some ambiguity in connectivities from C_2 to G_1; however, C_2 and G_3 are stacked in an *A*-like conformation. The non-exchangeable and imino NOE data discussed above suggested that there was no significant perturbation in stacking between C_{12} and C_{13} at the junction of stem 1 and stem 2, where the two stem regions are covalently linked. However, the two helical regions do not appear to form a completely continuous *A*-like helix. There is no NOE connectivity between G_3 and G_{22}, G_{22} adopts primarily an S-type sugar pucker. Even though they are not covalently linked, these nucleotides should exhibit the NOEs characteristic of *A*-form stacking if the two stems were perfectly continuous. The perturbation in helical structure must be minor. There is a NOE between imino protons on G_3 and G_{22} (resonances 3 and 4, respectively) and the stacking between C_{12} and C_{13} appears *A*-like. As discussed below, this perturbation is not surprising since both C_3 and G_{22} are involved in loop–helix junctions.

Information concerning the loop conformations in the pseudoknot is scant. Loop 1 is better defined than loop 2. Loop 1 crosses the deep major groove of stem 2. The 5′ nucleotide in loop 1 (A_4) shows no NOEs to G_3. The A_4 nucleotide may have a non-standard conformation and is stacked upon U_5. The far downfield chemical shift of A_4(H8) may be indicative of its environment within the loop. There are no NOEs from U_5 to the unassigned U_6. C_8 has a weak NOE to H1′, which probably belongs to U_7. There may be stacking of U_7 on C_8 at the 5′ terminus of stem 2, as suggested both by n.m.r. data and the double-strand-specific RNase V_1 cleavage at this nucleotide. The residues of loop 1 are cleaved much more weakly by nuclease S_1 than loop 2 nucleotides, indicating some protection relative to single strands. Model building suggests that both A_4 and U_5, even when stacked, can fit within the major groove. There is a possibility for tertiary interactions between the A_4 and the bases in stem 2 (i.e. U_9 or C_{24}). However, the N-7 atom of A_4 is more accessible to modification by DEP than either of the base-paired adenine residues. Thus, N-7 must be relatively accessible to solvent and not involved in hydrogen bonding. This does not, however, preclude interactions with other portions of the base ring. Alternatively, the loop region may be dynamically more flexible than the stem; the A_4(H8) resonance is slightly broader than other resolved aromatic resonances.

Connecting loop 2 is poorly defined by the n.m.r. data. It consists of six uridine residues (U_{16}–U_{21}), of which only U_{16} was definitely assigned. U_{16} is stacked at the 3′ end of stem 1 (on C_{15}); there is no NOE connectivity to U_{17}. U_{16}, as well as many of the uridine residues in the loop, adopts a south-type sugar conformation. This loop crosses the minor groove of stem 1. Nuclease S_1 cleaves strongly after three loop residues, U_{17}, U_{18} and U_{19}. Loop 2 joins stem 2 at G_{22}. The conformation of G_{22} is perturbed by the loop; it adopts a south-type sugar pucker. The S-type sugar pucker of U_{16} is adopted by nucleotides in tRNA at similar positions where the phosphodiester backbone changes direction. The π-turn observed in the anticodon loop of $tRNA^{Phe}$ results from changes in the values of the phosphodiester angles α and ζ (Saenger, 1984). This type of conformation may occur in the pseudoknot, but the values of the phosphodiester angles are probably *gauche–gauche*, rather than *gauche–trans*, since no peaks are shifted downfield in the ^{31}P n.m.r. spectrum (data not shown) (Gorenstein, 1981).

(c) *Structural model*

A low-resolution model of the structure of the pseudoknot has been constructed based on the results described above using the INSIGHT program on an Evans and Sutherland computer graphics terminal. The co-ordinates used for generation of the stems are based on an *A*-form helical structure from the fiber diffraction data of Arnott *et al.* (1973). Loop regions were constructed to be consistent with the n.m.r. data. A detailed representation is shown in Figure 8. The phosphate backbone from C_8 to U_{16} is a continuous unperturbed *A*-form helix. On the opposite backbone, the two stem regions (G_1 to G_3 and G_{22} to G_{26}) are stacked in *A*-form geometry without distortion. The region around G_3 and G_{22} contains close steric and electrostatic contacts between nucleotides from loop 2 and loop 1. It appears impossible to maintain a continuous, unperturbed helix without generating unfavorable interactions between phosphate molecules (highlighted in Fig. 9). The perturbation between G_{22} and G_3 is therefore unsurprising. G_{22} adopts an S-type sugar conformation that, possibly coupled with other structural changes, may move the phosphates from loop 1 and 2 nucleotides further apart. A slight increase in the twist of the helix at the junction of G_{22} and G_3 would relieve further some of the close interactions between the two loops, while maintaining the coaxial stacking of the two helices. The electrostatic interactions in this region may also explain the requirement for divalent

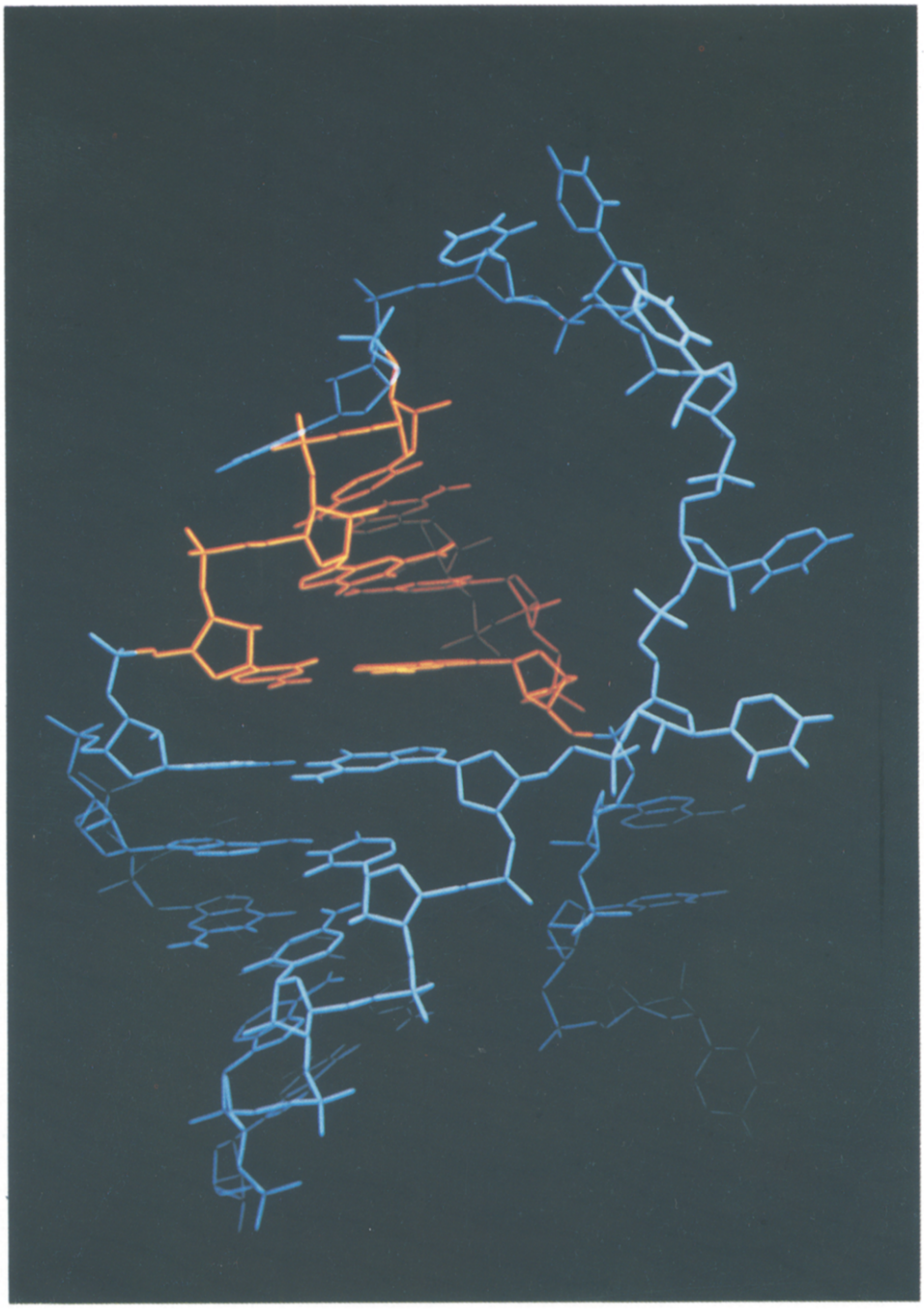


Figure 8. Detailed model of the pseudoknot showing the stacking of stem 1 (in red) on stem 2 (blue). The model was built with little distortion in *A*-form stacking between G_{22} and G_3 (see the text).

cations or high concentrations of monovalent cations for pseudoknot formation (see the following paper).

Loop 1 crosses the major groove of stem 2. This distance is quite short, even though the loop spans five base-pairs. The interphosphate distance (from G_3 to C_8) is only 12 Å, assuming *A*-form stacking. Loop residues A_4 and U_5 are shown stacked in the groove. The loop 1 residues could potentially interact with nucleotides in the major groove. In order to form base–triple interactions of the type observed in tRNA, some changes from regular *A*-form geometry must occur, since the major groove in *A*-form RNA is only 4·1 Å wide. In the structure of $tRNA^{Phe}$, the D-stem is slightly underwound (10·3 base-pairs/turn); this adjustment in twist causes a widening of the major groove and allows triple interactions to occur in that groove. Similar adjustments in geometry are observed in the fiber diffraction structure for triple stranded poly(U)·poly(A)·poly(U); the major groove is widened to allow the third poly(U) strand to form the triple (Saenger, 1984). Global alterations in structure may allow triple–base interactions between loops and stems in the pseudoknot, but there is no evidence for particular interactions.

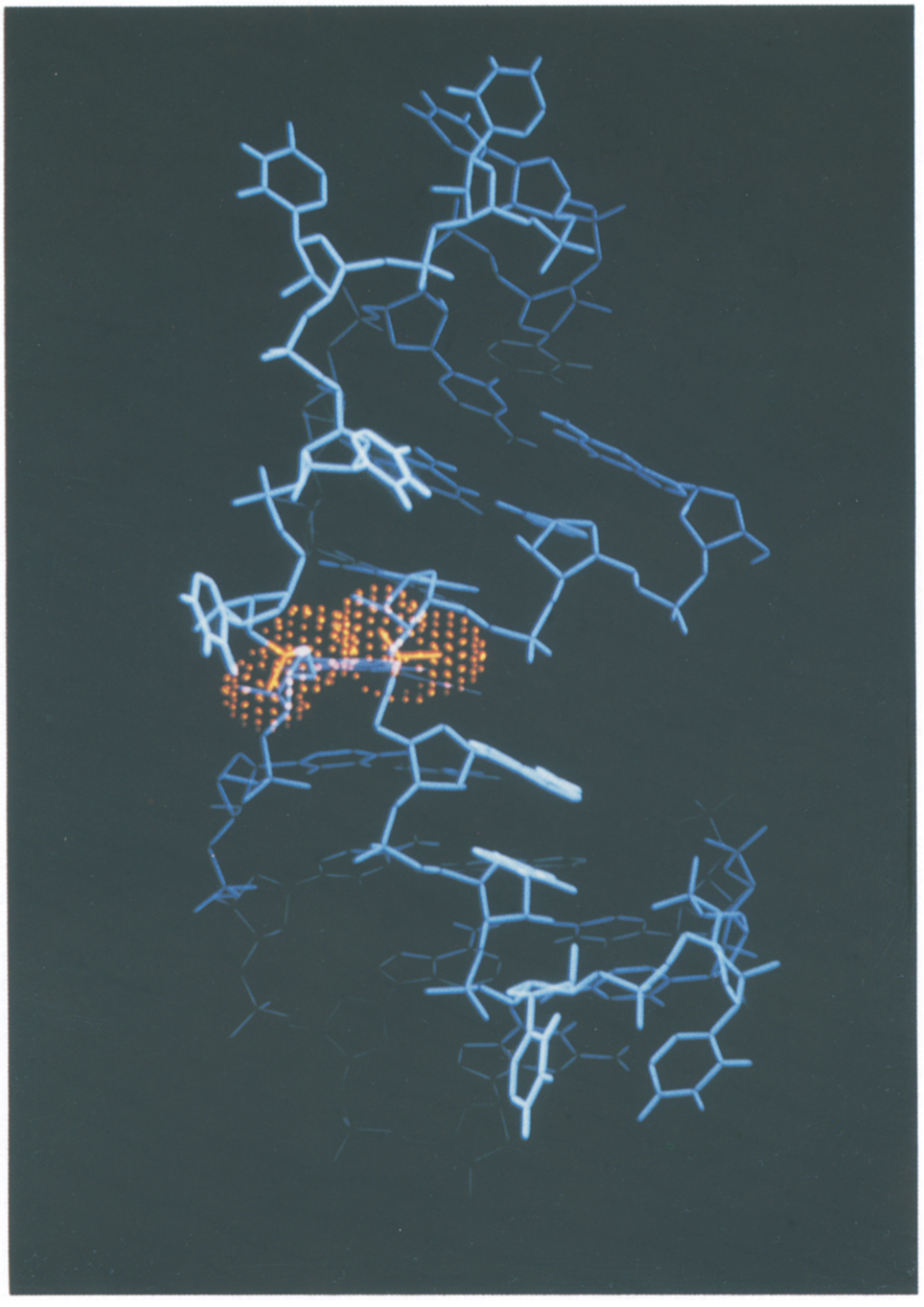


Figure 9. View of the loop side of the pseudoknot model. Loop 1 crosses the major groove of stem 2 (lower) and loop 2 crosses the minor groove of stem 1 (upper). A_4 and U_5 in loop 1 are stacked within the major groove. Loop 2 is poorly defined by the n.m.r. data. In this undistorted model, the 2 loop regions can come into close contact. Two phosphate molecules with van der Waals' radii that come into particularly close contact are highlighted in red.

Loop 2 crosses the wide, shallow minor groove of stem 1. The distance (18 Å) across this groove is longer than the distance across the major groove, even though only three base-pairs are spanned. Pseudoknot formation described here and variants with shorter loops (see the following paper) demonstrate that a loop size of six nucleotides is more than sufficient to bridge three base-pairs. Most of the loop residues probably have an S-type sugar pucker. A south-type sugar enables a nucleotide to span a greater distance (7·0 Å) than a nucleotide in a north conformation (5·9 Å). The phosphodiester backbone probably changes direction between nucleotides U_{16} and U_{17}. There does not appear to be significant stacking in loop 2.

4. Conclusions

These results show unequivocally that RNA can form a pseudoknot structure with coaxial stacking of stems. The sequence studied here forms a single unimolecular structure in 5 mM-Mg^{2+}; the n.m.r. data are not consistent with an equilibrium between

structures. This possibility could not be excluded in our previous studies on pseudoknot sequences (Puglisi *et al.*, 1988). Since the pseudoknot structure forms in the absence of further stabilization from interactions with RNA outside of the pseudoknot region, pseudoknots are viable forms of local RNA structure.

The structural model consistent with the n.m.r. data is in good agreement with the original proposals of Pleij *et al.* (1985) and later model building studies on the TYMV RNA (Dumas *et al.*, 1987). The two stem regions stack in a coaxial geometry. Although data indicate some distortion from *A*-form stacking between stem 1 and stem 2 at the junction between stem and loop regions, the model is still consistent with the observed biological activity of the pseudoknots in the viral tRNA-like structures. The stacked stem regions of the pseudoknot structure in TYMV RNA mimics the seven base-pairs in the acceptor stem of $tRNA^{Val}$, allowing aminoacylation by valyl-tRNA synthetase (Pleij *et al.*, 1985). Florentz & Giegé (1986) have mapped the interactions of valyl-tRNA synthetase with the TYMV RNA using chemical and enzymatic probes. The interactions with the "acceptor stem" formed by the pseudoknot are along the region equivalent to C_8 to C_{15}. It is this covalently connected phosphodiester backbone that will appear to the synthetase as a continuous *A*-form helix (Fig. 8). Mutations in the base-pairs of a similar pseudoknot in brome mosaic virus RNA drastically reduced the aminoacylation by synthetase (Dreher & Hall, 1988*b*). Thus, a pseudoknot has two sides: one is a fairly undistorted *A*-form helix (the left side of the model shown in Fig. 8) and the other contains both loop regions as well as the junctions of the loops with the stems (Fig. 9). Even if undistorted, the loop-junction side of the structure would not appear to a protein as an *A*-form helix.

The loop regions of the pseudoknot described in this study were purposely made long to allow them to span the stem regions easily; the effects of loop size on pseudoknot stability are described in the following paper. Structurally, the loop regions are less well characterized than the stems. Loop 1 has a better-defined structure than loop 2. The proposed distortion in the pseudoknot *A*-form geometry occurs in the region of the junction of the two loops with the stems (Fig. 9). Distortion of the structure, mainly localized to the stack between G_3 and G_{22} probably reduces unfavorable electrostatic interactions between phosphate groups at the loop–stem junctions.

Features of the pseudoknot other than the duplex region may be useful in biological functions. Proteins may recognize unique loop conformations. McPheeters *et al.* (1988) propose that bacteriophage T4 gene *32* protein regulates translation of its own mRNA by co-operative binding to the message. The nucleation site for the binding is proposed to be a pseudoknot. The authors suggest that gene *32* protein recognizes a conformation adopted by the larger loop in their pseudoknot (equivalent to loop 2). Since loop 2 is sterically less hindered than loop 1, it may be more accessible for protein binding.

The structure studied here represents a specific type of pseudoknot interaction; the stem regions are contiguous and stacked, and the loop regions are relatively short. In general, pseudoknot interactions can form where the stems are not adjacent. An examination of the present structure shows that if one or two nucleotides are inserted between the two stem regions (i.e. between C_{12} and C_{13}) close electrostatic interactions between the loops could be minimized although the coaxial stacking of the two stems would be eliminated. In the proposed pseudoknot required for efficient frameshifting in coronavirus RNA the two stem regions are separated by a nucleotide (Brierley *et al.*, 1989).

The loop regions of the pseudoknot can be made shorter as detailed in the following paper, or can be extremely long with secondary structure of their own. Pseudoknots with large loop sizes have been proposed in the structure of RNase P (James *et al.*, 1988). In 16 S rRNA, a pseudoknot is formed by pairing nucleotides 17 through 19 and 916 through 918 (Moazed & Noller, 1987). Loop 2 of this pseudoknot contains most of the 16 S rRNA secondary structure. These long-range pseudoknots are an effective mechanism for bringing regions distant in sequence into close spatial proximity. Such a pseudoknot interaction has been proposed in the structural model for mitochondrial group I introns (Davies *et al.*, 1982) and may be recognized by specific proteins (Akins & Lambowitz, 1987; Delahodde *et al.*, 1989). In its versatile structures and functions RNA seems more akin to proteins than DNA. The results presented here demonstrate the usefulness of n.m.r. studies on model oligonucleotides to determine biologically relevant RNA structures.

The authors thank Mr David Koh for synthesizing DNA oligonucleotides, Ms Barbara Dengler for general assistance, and Mr Jeff Pelton and Mr John Hubbard for their assistance with the computer graphics facility. We also thank Professor David Wemmer, Dr Peter Davis and Dr Gabriele Varani for advice and useful discussions. This work was supported in part by National Institutes of Health grant GM10840 and the Department of Energy, Office of Energy Research, Office of Health and Environmental Research under grant DE-FG03-86ER60406.

References

Akins, R. A. & Lambowitz, A. M. (1987). *Cell*, **50**, 331–345.

Altona, C. (1982). *Recl. Trav. Chim. Pays-Bas*, **101**, 413–434.

Arnott, S., Hukins, D. W. L., Dover, S. D., Fuller, W. & Hodgson, A. R. (1973). *J. Mol. Biol.* **81**, 107–122.

Auron, P. E., Weber, L. D. & Rich, A. (1982). *Biochemistry*, **21**, 4700–4706.

Bax, A. & Lerner, L. (1988). *J. Magn. Reson.* **79**, 429–438.

Benevides, J. M., Lemeur, D. & Thomas, G. J., Jr (1984). *Biopolymers*, **23**, 1011–1024.

Bodenhausen, G., Kogler, H. & Ernst, R. R. (1984). *J. Magn. Reson.* 370–388.

Braunschweiler, L., Bodenhausen, G. & Ernst, R. R. (1983). *Mol. Phys.* **48**, 535–560.
Brierley, I., Digard, P. & Inglis, S. C. (1989). *Cell*, **57**, 537–547.
Brush, C. K., Stone, M. P. & Harris, T. M. (1988). *Biochemistry*, **27**, 115–122.
Chazin, W. J., Wüthrich, K., Hyberts, S., Rance, M., Denny, W. A. & Leupin, W. (1986). *J. Mol. Biol.* **190**, 439–453.
Chou, S.-H., Flynn, P. & Reid, B. (1989). *Biochemistry*, **28**, 2422–2435.
Clore, G. M., Gronenborn, A. M., Piper, E. A., McLaughlin, L. W., Graeser, E. & van Boom, J. H. (1984). *Biochem. J.* **221**, 737–751.
Davenloo, P., Rosenberg, A. H., Dunn, J. J. & Sudier, F. W. (1984). *Proc. Nat. Acad. Sci., U.S.A.* **81**, 2035–2039.
Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A. & Scazzocchio, C. (1982). *Nature (London)*, **300**, 719–724.
Davis, P. W., Adamiak, R. W. & Tinoco, I., Jr (1990). *Biopolymers*, **29**, 109–122.
Delahodde, A., Goguel, V., Becam, A. M., Creusot, F., Perea, J., Banroques, J. & Jacq, C. (1989). *Cell*, **56**, 431–441.
Dreher, T. W. & Hall, T. C. (1988*a*). *J. Mol. Biol.* **201**, 31–40.
Dreher, T. W. & Hall, T. C. (1988*b*). *J. Mol. Biol.* **201**, 41–55.
Dumas, P., Moras, D., Florentz, C., Giegé, R., Verlaan, P., van Belkum, A. & Pleij, C. W. A. (1987). *J. Biomol. Struct. Dynam.* **4**, 707–728.
Florentz, C. & Giegé, R. (1986). *J. Mol. Biol.* **191**, 117–130.
Florentz, C., Briand, J. P., Romby, P., Hirth, L., Ebel, J. P. & Giegé, R. (1982). *EMBO J.* **1**, 269–276.
Gewirth, D. T. & Moore. P. B. (1988). *Nucl. Acids Res.* **16**, 10717–10732.
Gorenstein, D. G. (1981). *Annu. Rev. Biophys. Bioeng.* **10**, 355–386.
Guerrier-Takada, C., van Belkum, A., Pleij, C. W. A. & Altman, S. (1988). *Cell*, **53**, 267–272.
Haasnoot, C. A. G., Westerink, H. P., van der Marel, G. A. & van Boom, J. H. (1984). *J. Biomol. Struct. Dynam.* **2**, 345–360.
Haenni, A.-L., Joshi, S. & Chapeville, F. (1982). *Prog. Nucl. Acids Res. Mol. Biol.* **27**, 85–104.
Hall, K. B., Sampson, J. R., Uhlenbeck, O. C. & Redfield, A. G. (1989). *Biochemistry*, **28**, 5794–5801.
Happ, C. S., Happ, E., Nilges, M., Gronenborn, A. M. & Clore, G. M. (1988). *Biochemistry*, **27**, 1735–1743.
Hare, D. R., Wemmer, D. E., Chou, S.-H., Drobny, G. & Reid, B. R. (1983). *J. Mol. Biol.* **171**, 319–336.
Hayatsu, H., Wataya, Y., Kai, K. & Iida, S. (1970). *Biochemistry*, **9**, 2858–2865.
Heerschap, A. (1985). Ph.D. thesis, Catholic University of Nijmegen.
Hore, P. J. (1983). *J. Magn. Reson.* **55**, 283–300.
James, B. D., Olsen, G. J., Liu, J. & Pace, N. R. (1988). *Cell*, **52**, 19–26.
Joshi, R. L., Joshi, S., Chapeville, F. & Haenni, A. L. (1983). *EMBO J.* **2**, 1123–1127.
Keepers, J. W. & James, T. L. (1984). *J. Magn. Reson.* **57**, 404–426.
Kuchino, Y., Watanabe, S., Harada, F. & Nishimura, S. (1980). *Biochemistry*, **19**, 2085–2089.
Leontis, N. B. & Moore, P. B. (1986). *Biochemistry*, **25**, 3916–3925.
McPheeters, D. S., Stormo, G. D. & Gold, L. (1988). *J. Mol. Biol.* **201**, 517–535.
Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. (1987). *Nucl. Acids Res.* **15**, 8783–8798.
Moazed, D. & Noller, H. F. (1987). *Nature (London)*, **327**, 389–394.
Nishimura, S., Harada, F., Narushima, U. & Seno, T. (1967). *Biochim. Biophys. Acta*, **142**, 133–148.
Olson, W. K. (1982). *J. Amer. Chem. Soc.* **104**, 278–286.
Olson, W. K. & Sussman, J. L. (1982). *J. Amer. Chem. Soc.* **104**, 270–278.
Peattie, D. A. & Gilbert, W. (1980). *Proc. Nat. Acad. Sci., U.S.A.* **77**, 4679–4682.
Piantini, U., Sørensen, O. W. & Ernst, R. R. (1982). *J. Amer. Chem. Soc.* **104**, 6800–6801.
Pleij, C. W. A., Rietveld, K. & Bosch, L. (1985). *Nucl. Acids Res.* **13**, 1717–1731.
Pleij, C. W. A., van Belkum, A., Rietveld, K. & Bosch, L. (1986). In *Structure and Dynamics of RNA* (van Knippenberg, P. H. & Hilbers, C. W., eds), pp. 87–98, Plenum Press, New York.
Puglisi, J. D., Wyatt, J. R. & Tinoco, I., Jr (1988). *Nature (London)*, **331**, 283–286.
Puglisi, J. D., Wyatt, J. R. & Tinoco, I., Jr (1990). *Biochemistry*, **29**, 4215–4226.
Reid, B. R. (1981). *Annu. Rev. Biochem.* **50**, 969–996.
Rietveld, K., van Poelgeest, R., Pleij, C. W. A., van Boom, J. H. & Bosch, L. (1982). *Nucl. Acids Res.* **10**, 1929–1946.
Rietveld, K., Pleij, C. W. A. & Bosch, L. (1983). *EMBO J.* **2**, 1079–1085.
Rietveld, K., Linschooten, K., Pleij, C. W. A. & Bosch, L. (1984). *EMBO J.* **3**, 2613–2619.
Saenger, W. (1984). *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
Sampson, J. R. & Uhlenbeck, O. C. (1988). *Proc. Nat. Acad. Sci., U.S.A.* **85**, 1033–1037.
Schimmel, P. (1989). *Cell*, **58**, 9–12.
Studnicka, G. M., Rahn, G. M., Cummings, I. W. & Salser, W. A. (1978). *Nucl. Acids Res.* **5**, 3365–3387.
Tang, C. K. & Draper, D. E. (1989). *Cell*, **57**, 531–536.
van Belkum, A., Verlaan, P., Kun, J. B., Pleij, C. & Bosch, L. (1988). *Nucl. Acids Res.* **16**, 1931–1950.
van Belkum, A., Wiersema, P. J., Joordens, J., Pleij, C., Hilbers, C. W. & Bosch, L. (1989). *Eur. J. Biochem.* **183**, 591–601.
van den Hoogen, F. (1988). Ph.D Thesis, University of Leiden.
Varani, G., Wimberly, B. & Tinoco, I., Jr (1989). *Biochemistry*, **28**, 7760–7772.
Wataya, Y. & Hayatsu, H. (1972). *Biochemistry*, **11**, 3583–3588.
Weiner, A. M. & Maizels, N. (1987). *Proc. Nat. Acad. Sci., U.S.A.* **84**, 7383–7387.
Widmer, H. & Wüthrich, K. (1987). *J. Magn. Reson.* **74**, 316–336.
Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, New York.
Wyatt, J. R., Puglisi, J. D. & Tinoco, I., Jr (1990). *J. Mol. Biol.* **214**, 455–470.

Edited by M. F. Moody