Identification of Three G·U Base Pairs in *Bacillus subtilis* Ribosomal 5S RNA via 500-MHz Proton Homonuclear Overhauser Enhancements†

Lee-Hong Chang‡ and Alan G. Marshall*,§

Department of Chemistry, The Ohio State University, Columbus, Ohio 43210

Received June 28, 1985; Revised Manuscript Received November 20, 1985

ABSTRACT: Three distinct G·U base pairs in *Bacillus subtilis* 5S RNA have been identified via homonuclear Overhauser enhancements (NOE) of their low-field (9–15 ppm) proton Fourier transform nuclear magnetic resonances at 11.75 T. With these G·U resonances as starting points, short segments of NOE connectivity can be established. One G·U-G·C-G·C segment (most probably G_4·C_{112}-G_5·C_{111}-U_6·G_{110}) can definitely be assigned to the terminal helix. The existence of at least part of the terminal helical stem of the secondary structure of a Gram-positive bacterial 5S RNA has thus been established for the first time by direct experimental observation. Addition of Mg^{2+} produces almost no conformational changes in the terminal stem but results in major conformational changes elsewhere in the structure, as reflected by changes in the ^{1}H 500-MHz low-field NMR spectrum. Assignment of the two remaining G·U base pairs will require further experiments (e.g., enzymatic-cleavage fragments). Finally, the implications of these results for analysis of RNA secondary structure are discussed.

An essential structural and functional constituent of the large ribosomal subunits of virtually all prokaryotic or eukaryotic ribosomes is 55 RNA (Erdmann, 1976). A rapidly expanding library of approximately 240 5S RNA primary nucleotide sequences has been compiled by Erdmann et al. (1985). It is generally believed that a universal secondary structure exists, and a large number of universal secondary structures has been proposed on the basis of comparative sequence analysis (Fox & Woese, 1975; Hori & Osawa, 1979; Luehrsen & Fox, 1981; Studnicka et al., 1981; De Wachter et al., 1982; Delihas & Andersen, 1982), enzymatic cleavage and chemical modification (Nishikawa & Takemura, 1974; Garrett et al., 1981; Pieler & Erdmann, 1982), and physical and spectroscopic measurements (Kearns & Wong, 1974; Osterberg et al., 1976; Luoma & Marshall, 1978a,b).

Although 5S RNA from the thermophilic bacterium *Thermus thermophilus* has recently been crystallized by Morikawa et al. (1982), their X-ray diffraction resolution was not high enough to enable a determination of the molecular structure. It therefore appears likely that the main outlines of the secondary and tertiary structure will first be found via other physical and spectroscopic techniques. In any case, it will be important to determine whether or not the structure in solution is the same as that in a crystal. Physical/spectroscopic techniques already applied to 5S RNA include Raman (Luoma & Marshall, 1978a,b) and Fourier transform infrared (FT-IR) spectroscopy (Burkey et al., 1983; Chang et al., 1984; Li et al., 1984b), UV and circular dichroism (CD) (Luoma et al., 1980), electron spin resonance (Luoma et al., 1982), differential scanning calorimetry (Li & Marshall, 1985; Chang & Marshall, 1986), ^{19}F NMR (Marshall & Smith, 1980), and ^{1}H Fourier-transform nuclear magnetic resonance (FT NMR) (Wong et al., 1972; Kearns & Wong, 1974; Luoma et al., 1980; Salemink et al., 1981; Kime & Moore, 1983a,b; Kime, 1984a,b; Li & Marshall, 1986).

Of the available physical/spectroscopic techniques, high-resolution water-suppressed ^{1}H high-field FT NMR homonuclear Overhauser enhancements (NOE) offer the most definitive method for determining the solution structure of RNA and DNA (Schimmel & Redfield, 1983; Reid, 1981). For example, virtually all of the hydrogen-bond protons found in the downfield (9–15 ppm) ^{1}H NMR spectrum have been identified (i.e., A·U, G·C, G·U, etc.) and assigned to particular primary sequence positions for several tRNA's [e.g., Johnston & Redfield (1981), Hare & Reid (1982a,b), and Heerschap et al. (1983a,b)] and small DNA's (Hare et al., 1983; Feigon et al., 1983; Patel et al., 1982; Scheek et al., 1984).

Extension of the ^{1}H NOE method from tRNA's to 5S RNA's presents several problems, all associated with the much larger molecular weight of 5S RNA (ca. 40000 compared to ca. 25000 for a typical tRNA). First, the secondary structure of tRNA was evident as soon as the first tRNA primary nucleotide sequence had been determined (Holley et al., 1965) and was a great help in identifying and assigning base pair ^{1}H NMR signals in tRNA's. In contrast, the secondary structure of 5S RNA (assuming that a universal secondary structure exists) has yet to be established by any direct measurements. Second, the crystal and molecular structure of $tRNA^{Phe}$ (Ladner et al., 1975; Sussman & Kim, 1976) was known *in advance* of the major NMR work. Third, the larger number of bases (ca. 120 in 5S RNA vs. ca. 80 in a typical tRNA) leads to several more base pairs in 5S RNA than in tRNA (Li et al, 1984b; Chang et al., 1984; Li & Marshall, 1985), so that there will be more frequent overlap between ^{1}H NMR signals from different base pairs. Fourth, the larger molecular weight leads to broader ^{1}H NMR signals, further exacerbating the peak-overlap problem. Fifth, the 50% lower molar concentration for the same weight/volume concentration reduces the signal-to-noise ratio, which is further reduced (even for the same molar concentration of base pairs and same observation period) by the broader line width and longer longitudinal magnetic relaxation time, T_1. Finally, G·U pairs (with two ^{1}H NMR visible imino protons per base pair), which are rare in tRNA's, are common in 5S RNA's, adding even

† This work was supported by grants (to A.G.M.) from the U.S. Public Health Service (NIH 1 R01 GM-29274; NIH 1 S10 RR-01458) and The Ohio State University.

‡ Present address: Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, CA 94143.

§ A.G.M. is also a member of the Department of Biochemistry.

0006-2960/86/0425-3056$01.50/0

more peaks to the downfield region of the ^{1}H NMR spectrum.

The eubacteria (prokaryotes) are divided into two major groups: the Gram-negative bacteria (e.g., *Escherichia coli*) and the Gram-positive bacteria (e.g., *Bacillus subtilis*). The primary nucleotide sequences of the 5S RNA's from the two phylogenetic groups differ in length: *E. coli* 5S RNA has 120 nucleotides, and *B. subtilis* 5S RNA has 116 (Marotta et al., 1976). When the primary nucleotide sequences are aligned, it is found that in *B. subtilis* 5S RNA two nucleotides are deleted from the 5′ end and another two from the 3′ end (Hori & Osawa, 1979). Reconstitution experiments have shown that *Bacillus stearothermophilus* 5S RNA is replaceable by *E. coli*, *B. subtilis*, and other prokaryotic 5S RNA's, suggesting that all prokaryotic 5S RNA's are similar in structure and function (Wrede & Erdmann, 1973). Because previous ^{1}H NMR of Gram-negative *E. coli* 5S RNA at 220 (Wong et al., 1972), 300 (Kearns & Wong, 1974), and 400 MHz (Burns et al., 1980) has already been extended to 500 MHz (Kime & Moore, 1983a,b; Kime et al., 1984; Kime, 1984a,b), we chose Gram-positive *B. subtilis* 5S RNA to test for universality of 5S RNA secondary structure between Gram-positive and Gram-negative bacteria.

Materials and Methods

Preparation of 5S RNA for NMR. *B. subtilis* 5S RNA was purified as described previously (Li et al., 1984a). The basic three-step procedure consisted of phenol/sodium dodecyl sulfate (SDS) extraction of whole cells, followed by DE-32 cellulose ion-exchange chromatography and large-scale gel filtration chromatography on Sephadex G-75. The 5S RNA preparation is homogeneous according to denaturing polyacrylamide gel electrophoresis, and only a few percent of the molecules showed nucleolytic nicks. For NMR experiments, Mg^{2+}-depleted 5S RNA samples were prepared by dialyzing first against 10 mM cacodylate, pH 7.0, 0.1 M NaCl, and 10 mM ethylenediaminetetraacetic acid (EDTA) followed by distilled water. After lyophilization, 16 mg of 5S RNA was dissolved in 0.4 mL of 10 mM cacodylate, pH 7.0, 0.1 M NaCl, and 1 mM EDTA, containing 5% D_2O to give a 1.0 mM 5S RNA solution. For Mg^{2+} titration experiments, aliquots of 0.25 M Mg^{2+} solution were added to give the specified Mg^{2+} concentrations.

NMR Spectroscopy. NMR spectra at 500-MHz were acquired with a Bruker AM-500 FT-NMR spectrometer with quadrature detection and four-phase cycling, without sample spinning, at 23 °C. Normal (i.e., non-NOE) spectra were obtained via a modified Redfield 214 pulse sequence (Redfield & Kunz, 1979) with a carrier frequency centered at 15.0 ppm (i.e., 5200 Hz from the H_2O resonance, with a total pulse duration of 196 μs), with an acquisition time of 0.3 s (16-bit analog-to-digital conversion to give 8K time-domain data points), and with a relaxation delay of 0.5 s between successive acquisitions. The durations of the "4" pulse and the final "2" pulse were adjusted to give maximal water suppression. When alternate delayed acquisition (Roth et al., 1980) was used, a 98-μs delay was introduced between the end of the 214 pulse and the data acquisition trigger. Usually, 1000–2000 scans were signal-averaged in the time domain and all spectra plotted without base-line correction. Chemical shifts were measured relative to the solvent H_2O peak and converted to a scale with DSS (2,2-dimethyl-2-silapentane-5-sulfonate) at zero, based on an independent calibration of H_2O relative to DSS.

For the NOE experiments, a peak of interest was selectively saturated by gated irradiation at that frequency at a decoupler power sufficient to reduce the peak intensity by 60–80% in 0.6 s. When two peaks were closely spaced, the peak of interest was irradiated slightly off-center on the side away from the adjacent peak, in order to reduce irradiation spillover to the neighboring peak. A 2-ms delay was inserted between the end of the preirradiation pulse and the start of the observation pulse sequence. An NOE difference spectrum was obtained by alternate addition and subtraction of 4K time-domain data sets, with simultaneously alternating on- and off-resonance irradiation (see below).

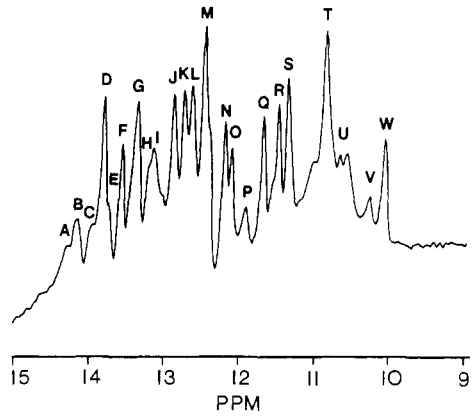

FIGURE 1: The 500-MHz low-field ^{1}H NMR spectrum of *B. subtilis* 5S RNA at 23 °C in 10 mM cacodylate, 0.1 M NaCl, 0.1 mM EDTA, pH 7.0, and 95:5 H_2O:D_2O, after complete removal of Mg^{2+} ion. The spectrum has been resolution-enhanced by Lorentzian-to-Gaussian conversion. The peaks are labeled from A to W for convenience in discussion.

Because the large H_2O signal is excited to a different degree by the on- and off-resonance irradiations, an NOE difference spectrum obtained by the above method will exhibit a curved base line. The problem is largely solved by an eight-step phase-cycling sequence: decoupler phase (0, 0, 0, 0, 0, 0, 0, 0); transmitter phase (0, 0, 180, 180, 90, 90, 270, 270); receiver phase (0, 180, 180, 0, 90, 270, 270, 90); decoupler frequency (on-, off-, on-, off-, on-, off-, on-, off-resonance). This scheme produces a time-domain NOE difference signal, suppresses the carrier "glitch" at zero frequency in quadrature mode, reduces the "image" peaks at negative frequencies, and corrects somewhat for the unequal excitation of H_2O produced by on- and off-resonance decoupling. This sequence performs well, even when the on-resonance decoupler frequency is relatively near to H_2O and the off-resonance frequency is far from H_2O (see Results).

A total of 16 000 time-domain transients (8000 on- and 8000 off-resonance) was typically required to distinguish small Overhauser enhancements from background noise in the transformed difference spectrum. Exponential apodization equivalent to frequency-domain line broadening of 5–10 Hz was applied to improve the signal-to-noise ratio in the final NOE difference spectrum.

Results and Discussion

Number of Base Pairs. At least 23 separate peaks are resolved in the low-field (9–15 ppm) region of the ^{1}H FT NMR spectrum of *B. subtilis* in the absence of Mg^{2+} (see Figure 1). By analogy to tRNA, most peaks in this spectral region correspond to hydrogen-bonded ring "imino" (i.e., N–H···N) proton resonances from Watson–Crick A·U or G·C base pairs or "wobble" G·U base pairs [e.g., Reid (1981)]. Of these, A·U and G·C contribute one NMR-visible imino proton each, and each G·U pair contributes two NMR-visible imino protons. The remaining hydrogen-bonded base pair protons exchange so rapidly with H_2O that their ^{1}H NMR signals are broadened below detectability.

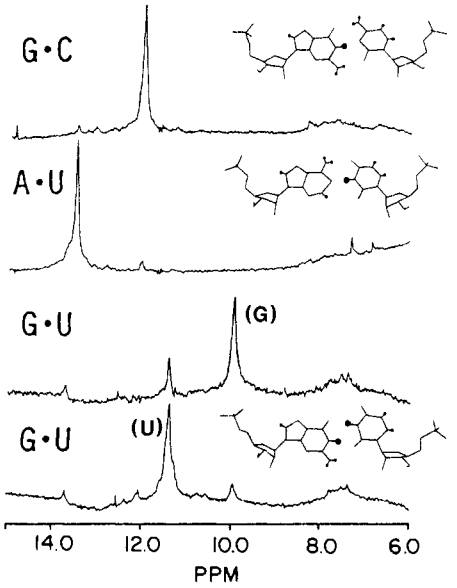

FIGURE 2: NOE difference spectra resulting from saturation of each of four low-field proton 500-MHz nuclear resonances of *B. subtilis* 5S RNA. The identity of each base pair proton shown as the large dot in each structure (i.e., A·U, G·C, or G·U) can be inferred from its characteristic NOE pattern (see text).

The reliability of determining the number of base pair protons by direct integration of the low-field ^1H NMR spectrum is questionable because of base-line curvature (due to the residual H_2O peak), the presence of partly melted base pairs whose resonances contribute less than unit intensity, and the nonflat power spectrum of the excitation pulse sequence (Redfield & Kunz, 1979). At best, the base pair proton number can be estimated from a simulated spectrum consisting of a sum of the minimum number of peaks (all of the same width and height) required to match the experimental spectrum (Reid, 1981). For the spectrum shown in Figure 1, such a simulation yields approximately 41 base pair protons. Since at least three of these are G·U pairs (see below), we conclude that at least 38 base protons are present, in good agreement with the FT-IR result (36 base pairs) for the same molecule in the same buffer (Chang et al., 1984) and with differential scanning calorimetry results (Chang & Marshall, 1986). It is worth noting that the NMR estimate represents a lower limit (since the intensity of some proton resonances may be reduced by chemical exchange with H_2O) and that the NMR value may include tertiary hydrogen-bonded protons that are not necessarily detected by FT-IR.

Peak Identification (A·U, G·C, or G·U) via Primary NOE's. Identification of low-field ^1H NMR resonances with specific base pair types (A·U, G·C, or G·U) is based on (homo)nuclear Overhauser enhancement (NOE) double-resonance experiments. The NOE experiment demonstrates the connection between spatially proximal protons by saturating the magnetization of one proton and observing subsequent magnetization transfer via through-space dipole–dipole coupling from that proton to a neighboring proton (Noggle & Schirmer, 1971). The NOE effect varies as r^{-6}, in which r is the separation between the two protons of interest. For a molecule the size of 5S RNA, the NOE effect is negative (i.e., leads to a decrease in the intensity of the proton coupled to an irradiated proton) and extends only over a range of a few angstroms (Bothner-by, 1979).

For example, the two imino protons in a G·U base pair are spatially very close together (ca. 2.5 Å), and both protons typically exchange relatively slowly ($<5\ s^{-1}$) with H_2O. Thus,

Table I: Base Pair Identification and Sequencing via Secondary NOE Connectivities

peak	base pair type	NOE connectivity to	inferred base pair sequences[a]	
D	A·U	K, R, W		
			R/W	G·U
G	G·C	O, Q, T		
			D	A·U
I	G·C	N, S		
			K	G·C
K	G·C	D, H		
N	G·U	I, S		
			Q/T	G·U
O	G·C	F, G, I		
			G	G·C
Q	G·U	G, T		
			O	G·C
R	G·U	D, W		
S	G·U	N	N/S	G·U
T	G·U	G, Q	I	G·C
W	G·U	D, R		

[a] Except in certain cases (see text), it is not possible to determine the parity of the base pair (e.g., G·C vs. C·G).

irradiation of either the G or U imino proton in a G·U pair produces a large NOE to its partner, as shown in the lower two spectra of Figure 2. As with tRNA's, the enhancement produced by irradiation of the G imino proton is larger than that by irradiation of the U imino proton (Hurd & Reid, 1979; Johnston & Redfield, 1981; Hare & Reid, 1982a). This characteristic NOE behavior makes for highly reliable identification of G·U resonances.

A·U and G·C resonances in the low-filed imino proton spectrum also exhibit characteristic NOE patterns (Sanchez et al., 1980; Hare & Reid, 1982b), as illustrated in the top two traces of Figure 2. Saturation of an A·U imino proton produces observable decreases in the intensity of a sharp resonance from the C2-H proton of adenine at ca. 7–8 ppm, as well as other broader resonances from the aromatic amino protons in the 7–9 ppm range. In contrast, saturation of a G·C imino proton elicits only broad NOE's to aromatic amino protons in the 7–9 ppm region. The above-listed primary NOE patterns provided identifications of the several low-field ^1H resonances of *B. subtilis* 5S RNA reported in Table I.

Short Sequences of Base Pairs via Secondary NOE's. As discussed above, "primary" NOE's from an irradiated base pair hydrogen-bond proton to other protons in the same base pair often suffice to identify the base pair type (A·U, G·C, G·U). However, the next-nearest NMR-visible neighbor proton may be a hydrogen-bond imino proton from the base pair immediately above or below the irradiated base pair imino proton. For example, the top two traces in Figure 2 clearly establish a connection between an A·U and a G·C, on the basis of the small interpair "secondary" NOE observed for either peak when the other is irradiated. Thus, by sequentially irradiating the imino protons that are connected by NOEs, one obtains a linear "base pair sequence", provided that all of the resonances in question are single and well-resolved (see below). When the RNA secondary and tertiary structures are known, NOE connectivities can be used to assign each resonance in the low-field ^1H NMR spectrum. For example, it has proved possible to identify and sequence all of the secondary and tertiary base pairs in $tRNA^{Phe}$ by ^1H NOE's (Johnston & Redfield, 1981; Heerschap et al., 1983a,b), to give a secondary structure in complete agreement with the X-ray structure.

In order to develop a complete 5S RNA secondary structure based upon NOE-determined base pair segments, it is necessary to begin from a well-resolved base pair imino proton

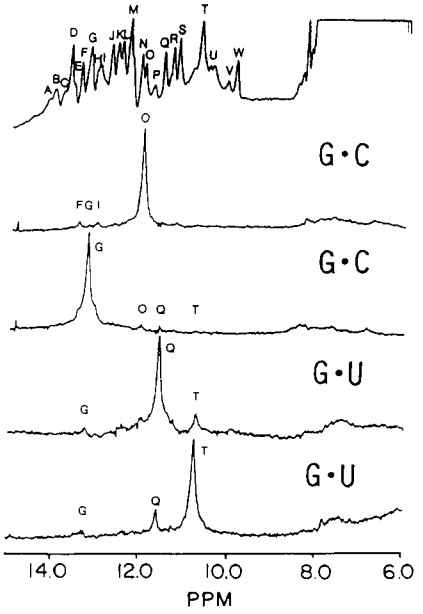

FIGURE 3: NOE difference spectra of resonances related to resonances Q/T. The topmost reference spectrum was obtained in the absence of preirradiation. The largest peak in each NOE spectrum is the preirradiated resonance; peaks to which saturation is transferred are labeled as in the reference spectrum. The probable identity of each irradiated resonance is denoated as A·U, G·C, or G·U, based on its primary NOE behavior (see text).

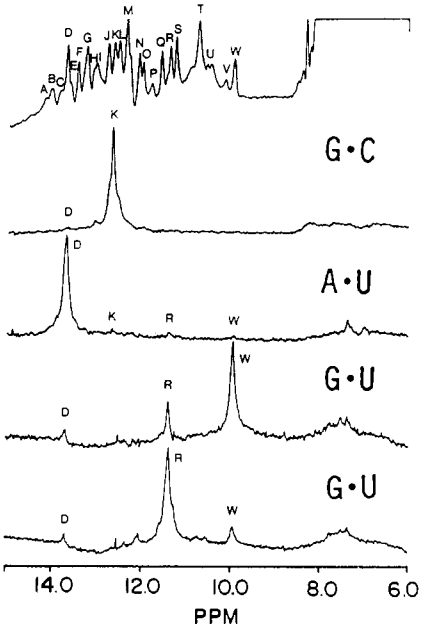

FIGURE 4: NOE difference spectra of resonances related to resonances R/W. Format is as for Figure 3.

resonance whose identity is unambiguous. G·U imino resonances form an especially attractive starting point for two reasons. First, their chemical shifts typically fall in the otherwise relatively uncrowded 10–12.5 ppm region of the ^{1}H NMR spectrum (see Figure 1). Second, saturation of either of the two imino protons of a G·U pair gives a larger NOE (≥15%) than for any other base pair type. Irradiation of all the resonances in the 10–12.5 ppm region produced large mutual NOE's for three pairs of resonances: Q and T (Figure 3), R and W (Figure 4), and N and S (Figure 5), which we therefore assign to be G·U hydrogen-bonded imino protons.

Primary and secondary NOE connectivities for resonances Q and T are shown in Figure 2, and the results are summarized in Table I. The two lowermost traces in Figure 2 show that

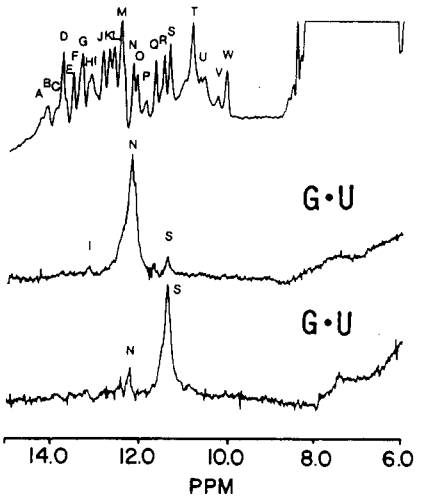

FIGURE 5: NOE difference spectra of resonances N/S. Format is as for Figure 3.

saturation either of peak Q at 11.7 ppm or of peak T at 10.9 ppm gives a large primary NOE to the other, confirming they comprise the two imino protons of a G·U base pair. A smaller secondary NOE to peak G is also visible. Saturation of peak G at 13.4 ppm produces only broad, diffuse NOE's in the aromatic region (characteristic of a G·C base pair imino proton), as well as the expected secondary NOE back to the G·U protons Q and T, and a small (presumably secondary) NOE to peak O at 12.1 ppm. Saturation of O in turn reveals an expected NOE back to G, as well as NOE's to peaks F and I. The NOE difference spectrum for peak O is again characteristic of a G·C pair, because no well-defined upfield is seen. Thus, we can trace NOE connectivity from peaks Q/T to G to O, corresponding to a "base pair sequence" of G·C-G·C-G·U or G·U-G·C-G·C. In general, the NMR NOE results provide neither the polarity of the base pair (e.g., G·C vs. C·G) nor the sequence (e.g., ABC vs. CBA).

The large mutual NOE's observed by saturation either of peak R at 11.5 ppm or of peak W at 10.1 ppm (see the two lowermost traces in Figure 4) identify these two resonances as the two hydrogen-bonded imino protons of a second G·U base pair. Peaks R/W show a secondary NOE connectivity to peak D at 13.8 ppm. Saturation of peak D gives the expected NOE's back to peaks R and W, as well as to peak K at 12.8 ppm, and a sharp NOE in the aromatic region that identifies peak D as an A·U base pair proton. Finally, saturation of peak K gives a barely detectable NOE back to peak D. Peak K is probably G·C because of its broad NOE in the aromatic region. Thus, an NOE connectivity from peaks R/W to D to K clearly establishes a G·U-A·U sequence, which can probably be extended to another G·C (see Table I).

A third proposed G·U base pair can be inferred from the large mutual NOE's produced by saturation of either peak N at 12.2 ppm or peak S at 11.4 ppm. A small secondary NOE is observed to peak I, which appears to be a G·C base pair proton. We speculate that the N/S pair could be either a terminal secondary base pair or a tertiary base pair.

Assignment of G·U Peaks Q and T as U6·G110 in the Terminal Stem. Two of the numerous proposed secondary structural models for 5S RNA have been adapted to the *B. subtilis* 5S RNA primary nucleotide sequence in Figure 6. These two models represent the two classes of secondary structures that can be adapted to virtually all of the known 5S RNA primary sequences. Moreover, application of the NOE strategy to an RNase-resistant fragment of *E. coli* 5S RNA has resulted in the identification of segments of the 5S

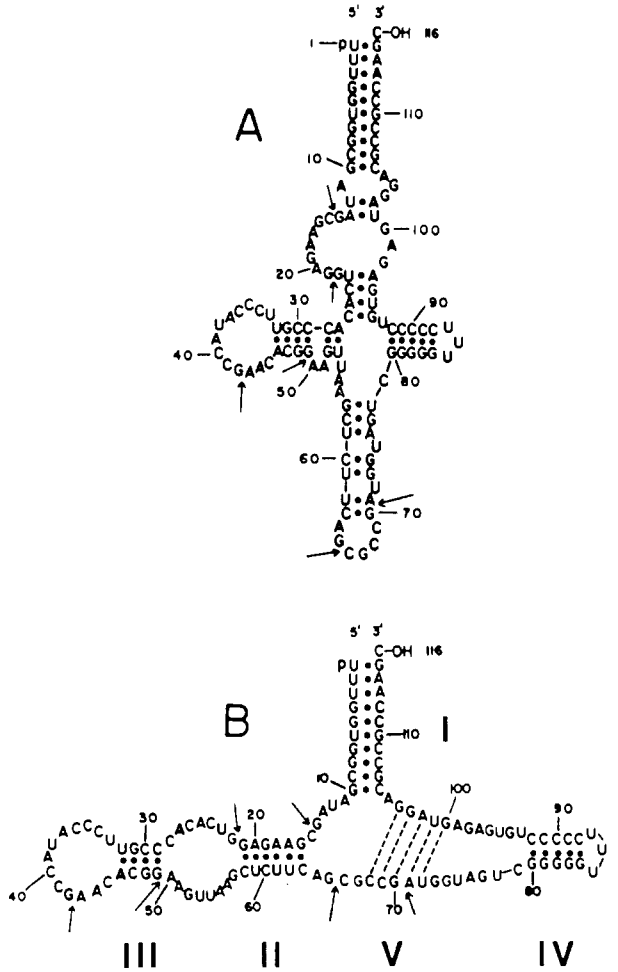

FIGURE 6: Two of the proposed secondary structural models adapted to the primary base sequence of *B. subtilis* 5S RNA: (A) Luoma and Marshall; (B) Fox and Woese model with an additional helix V depicted in dashed lines as proposed by De Wachter et al. (1982) and Delihas and Andersen (1982). Arrows denote sites of cleavage by RNase T1.

RNA secondary structure (Kime & Moore, 1983a,b; Kime et al., 1984), thereby eliminating most other proposed secondary structural models (Kime & Moore, 1983a). Inspection of the models reveals that only the terminal stem exhibits a base pair sequence that is compatible with the minisequence G·C-G·C-G·U (or G·U-G·C-G·C) inferred from NOE connectivities to peaks Q and T. Although the Q/T resonances can thus be assigned with some confidence to U6·G110, we are left with two possibilities for the other two peaks: O–G–Q/T, corresponding to base pairs G4·C112-G5·C111-U6·G110, or Q/T–G–O, corresponding to base pairs U6·G110-G7·C109-G8·C108. In either case, the identification of U6·G110 provides the most direct evidence to date for the terminal helical stem of 5S RNA from a Gram-positive organism.

Shifts Induced by a G·U Base Pair. The first reliable estimates of hydrogen-bond imino proton chemical shifts induced by neighboring base pairs in RNA were tabulated for A·U and G·C pairs in an 11-fold RNA A-helix by Arter and Schmidt (1976) on the basis of charge currents from each nucleotide base treated as a benzene ring (Giessner-Prette & Pullman, 1970). Although far from perfect, these estimates have proved quite helpful in reaching in a self-consistent set of assignments of all of these base pair imino proton resonances in $tRNA^{Phe}$ [e.g., Johnston & Redfield (1981)].

For G·U pairs, Geerdes and Hilbers (1979) have attempted to predict G·U resonance positions on the basis of chemical shifts induced *by* neighboring base pairs. However, since such estimates have been shown to err by as much as 0.5–0.7 ppm in some tRNA's (Reid, 1981), we cannot reliably use this method to assign G·U pairs in an RNA of unknown secondary base pair sequence.

Conversely, for the chemical shift induced *by* a G·U upon other base pair hydrogen-bond imino protons, previous authors have treated the G·U pair as if it were a G·C pair (Reid et al., 1979; Johnston & Redfield, 1981). However, careful scrutiny of the three-dimensional structure of yeast $tRNA^{Phe}$ reveals that the wobble G·U base pair exhibits substantially greater overlap with the base pair following it on the 3′-side of G than with the base pair preceding it on the 5′-side of G (Mizuno & Sundaralingam, 1978). Therefore, since a G·U pair distorts the RNA from its basic A-helix pattern, the validity of treating a G·U as a G·C in predicting ring current induced chemical shifts is questionable.

A semiempirical approach might be to compare the experimentally observed G·C (or A·U) chemical shift to the theoretical chemical shift obtained by subtracting from the unshifted G·C (or A·U) position the contributions from nearest and next-nearest G·C (or A·U) base pairs (Arter & Schmidt, 1976). The two problems with this approach are (a) the available data are necessarily incomplete, because of the scarcity of G·U pairs in RNA's of known secondary base pair sequence and because data involving terminal base pairs are less reliable, and (b) the distortion in the RNA A-helix produced by a G·U pair renders invalid any ring current calculations based on rigid A-helix geometry. In any case, such calculations suggest that the chemical shift induced by a G·U pair upon its next-nearest neighbors can be quite large (up to ±0.6 ppm) perhaps because the significant A-helix distortion produced by a G·U pair extends at least two base pairs distant from the G·U pair itself.

Proposed Assignments for a Second G·U Base Pair. Figure 4 shows the NOE connectivities from R/W to D to K. On the basis of their primary NOE patterns, peaks R/W represent the two hydrogen-bonded base pair imino protons from a G·U pair, and peaks D and K represent A·U and G·C imino protons, respectively. The NOE connectivities (see Table I) therefore suggest the presence of a base paired segment, G·U-A·U-G·C, in which the polarity of each pair (e.g., A·U vs. U·A) and the sense of the sequence (i.e., 5′ to 3′ or 3′ to 5′ on a given strand) cannot be inferred from Figure 4.

Assignment of the NOE-inferred base paired segment G·U-A·U-G·C depends upon the choice of a secondary structure for *B. subtilis* 5S RNA. The only such segment that occurs in the secondary base paired sequences common to the major models (see Figure 6) is G1·U115-U2·A114-U3·A113-G4·C112, provided that peak D is assumed to represent two consecutive A·U pairs. Indeed, saturation of peak D does appear to generate two intrapair NOE's at about 6.9 and 7.3 ppm, which could represent C2-H protons of two adenines. However, such an assignment would require that the G·U pair be the terminal G1·U115. A terminal G·U base pair is generally considered to be energetically unstable (Tinoco et al., 1973). Moreover, the fraying motion at a helical terminus is likely to be very rapid, rendering the two G·U imino protons unobservable due to fast exchange with solvent.

The remaining base paired segments of the Fox/Woese secondary structural model (Figure 6, bottom) do not contain a G·U-A·U-G·C sequence. However, recently de Wachter et al. (1982) and Delihas and Andersen (1982) have proposed variants of the Fox/Woese model that exhibit an additional helix V. In helix V, the sequence U72·G100-A71·U101-

G70·A102 could give an NOE connectivity matching our observed NOE results, provided that the NOE-connected resonance at K includes an A·G (rather than G·C) base pair. In support of this idea, the existence of an A·G base pair has been observed in a synthetic DNA fragment (Kan et al., 1983), at a resonant frequency (12.6 ppm) that is very close to the observed K resonance at 12.76 ppm. In the absence of more direct evidence, such an assignment must obviously be viewed as highly conjectural.

Alternatively, the remaining base paired segments of the Luoma/Marshall model (Figure 6, top) offer as a possible segment U61·G73-U62·A71-C63·G70. This assignment requires the questionable assumption that the one-base bulge at U72 not disrupt the RNA helical base stacking, so that dipolar coupling between the two imino protons of U61·G73 and the imino proton of U62·A71 be conserved.

A fourth possibility is that one or more of the NOE-connected base pairs represents a tertiary base pair. On the basis of the present evidence, we cannot logically choose between the above possibilities.

Implications of a Third G·U Base Pair. The large mutual NOE's between peaks N and S in Figure 5 offer strong evidence for the existence of a third G·U base pair in *B. subtilis* 5S RNA. This proposed G·U is NOE-connected to peak I, which appears to be a G·C base pair. The existence of a third G·U pair presents threee alternatives. First, the Fox/Woese model [as extended by de Wachter et al. (1982) and Delihas and Andersen (1982)] can account for a third G·U via pairing of positions 77–78 with 94–95. A problem with such an idea is the need for an interior loop, presumably stabilized by tertiary structure, or a non-Watson–Crick C·U pair at positions 79 and 93. Second, the Luoma/Marshall model could accommodate a third G·U pair at G57·U78, followed by a G·C pair at C58·G77. An objection to the latter assignment is that G57·U78 is located at a helical terminus and thus might be expected to fray too rapidly to produce NMR-observable peaks. A third alternative is that one of the three apparent G·U pairs is a tertiary base pair.

Effects of Mg^{2+} on the Low-Field ^{1}H NMR Spectrum of B. subtilis 5S RNA. The effects of various concentrations of Mg^{2+} upon the solution structure of *B. subtilis* 5S RNA are shown in Figure 7. The bottom spectrum corresponds to *B. subtilis* 5S RNA, which has been extensively dialyzed against 10 mM EDTA to remove Mg^{2+} as completely as possible. Dramatic changes in the low-field imino proton spectra are observed when the Mg^{2+} concentration increases from 2 to 20 mM, corresponding to an increase of the Mg^{2+}/RNA mole ratios from 5 to 50. Although little or no shifts were induced by $[Mg^{2+}] < 2$ mM, higher concentrations of Mg^{2+} induced shifts for most low-field resonances. For example, peak D shifts significantly upfield, whereas resonance F moves noticeably downfield, so that the two resonances eventually merge at a Mg^{2+}/RNA ratio of 50:1.

It is interesting to note that resonances O, G, and Q/T exhibit almost no Mg^{2+}-induced shifts, corroborating our previous assignment of peaks Q/T as U6·G110 in the terminal stem. Thus, even though there are major Mg^{2+}-induced conformational changes elsewhere in the *B. subtilis* 5S RNA secondary structure, the midregion of the terminal helix retains its shape.

Resonances R and W from the same G·U base pair appear to shift slightly upfield in the presence of Mg^{2+}. Resonances N and S from the same G·U base pair move significantly upfield on addition of Mg^{2+}, corroborating our previous speculation that N/S could be a terminal G·U or tertiary base pair (which would be expected to be especially sensitive to conformational changes).

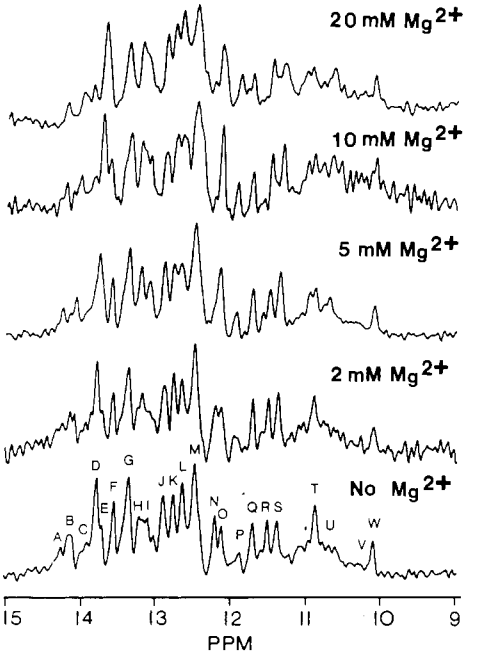

FIGURE 7: Effect of Mg^{2+} titration upon the 500-MHz low-field imino proton spectrum of *B. subtilis* 5S RNA. The lowermost spectrum is that of 0.4 mM *B. subtilis* 5S RNA in the complete absence of Mg^{2+} in a buffer containing 10 mM cacodylate, 0.1 M NaCl, and 1 mM EDTA, pH 7.0. Small aliquots of concentrated $MgCl_2$ solution were added to give the indicated concentrations. All spectra were plotted after Lorentzian-to-Gaussian resolution enhancement.

Finally, the addition of Mg^{2+} produces a variety of other changes in the low-field spectrum (e.g., note the additional spectral structure in the vicinity of peak T), whose interpretation must await additional identification and assignments of the remaining resonances.

Comparison between E. coli and B. subtilis 5S RNA. Our most definitive G·U assignment is for U6·G110 in the terminal stem of *B. subtilis* 5S RNA. Unfortunately, there is no corresponding G·U pair in the terminal stem of *E. coli* 5S RNA, although Kime and Moore (1983a) and Kime et al. (1984) have assigned several resonances in the terminal stem of *E. coli* 5S RNA.

Our assignment of a second G·U pair in *B. subtilis* 5S RNA (peaks R/W) depends upon a choice of secondary structural model, as noted above. For a modified Fox/Woese model (Figure 6), inspection of the secondary structures of both *E. coli* and *B. subtilis* 5S RNA's reveals that both 5S RNAs exhibit identical base pair sequences in helix V: i.e., the sequence C·G-C·G-G·A-A·U-U·G corresponding to base pairs 70·106 to 74·102 in *E. coli* 5S RNA, compared to base pairs 68·104 to 72·100 in *B. subtilis* 5S RNA. Kime et al. (1984) assigned the two imino protons of resonances P and R2 (their designation) at approximately 11.55 and 10.6 ppm to the U74·G102 pair of helix V in *E. coli* 5S RNA. By analogy, one might assign resonances R and W (our designation) at 11.5 and 10.1 ppm to the corresponding U72·G100 pair in *B. subtilis*.

One would expect that the identical sequence of helix V in the two 5S RNAs should produce very similar local structure and, hence, very similar chemical shifts for the imino protons of the two G·U pairs. We therefore suggest that the counterpart of our resonance W in *B. subtilis* 5S RNA should be Kime's et al. resonance S at 10.15 ppm for *E. coli* 5S RNA, rather than their initial assignment as resonance R2.

One-Sided NOE Produced by Saturation of 5S RNA G·U's. In accord with prior results for *E. coli* 5S RNA (Kime & Moore, 1983a) and the 3′-terminal colicin fragment of *E. coli* 16S RNA (Heus et al., 1983), we find that saturation of a G·U imino proton produces a one-sided NOE. More specifically, saturation of either G·U base pair imino proton produces saturation transfer to the imino proton of the nearest-neighbor base pair located on the 3′-side of G of the G·U base pair but not to the imino proton located on the 5′-side of G of the G·U base pair. It is known that a Watson–Crick base pair on the 3′-side of G of a G·U pair exhibits strikingly greater base-stacking overlap than a Watson–Crick base pair on the 5′-side of G of a G·U base pair (Mizuno & Sundaralingam, 1978). The one-sided NOE effect has been ascribed to that difference in base stacking (Heus et al., 1983). Nevertheless, it is puzzling that saturation of a tRNA G·U base pair imino proton produces a two-sided NOE: i.e., saturation transfer to imino protons of base pairs on both the 3′- and 5′-side of the G·U pair [e.g., Roy & Redfield (1983), Hare & Reid (1982b), and Heerschap et al. (1982)].

Implications of the Present Results. Two major differences between tRNA's and 5S RNA's are (a) tRNA's typically contain several chemically modified (e.g., methylated) bases, whereas 5S RNA's do not, and (b) 5S RNA's typically contain several G·U base pairs, which are rare in tRNA's. Thus, the modified bases, which provided convenient starting points for assignment of base pair proton spectra in tRNA's, are not available for 5S RNA. However, the present results show G·U base pairs in 5S RNA can readily be identified from their characteristic chemical shifts and large mutual nuclear Overhauser enhancements in the low-field proton 500-MHz spectral region. With these G·U resonances as starting points, short segments of NOE connectivity can then be established. It is worth noting that the identification and assignment even of a few short base paired segments may suffice to decide between proposed 5S RNA secondary structures.

It is obvious that the addition of Mg^{2+} ion produces major changes in the low-field imino proton spectrum of *B. subtilis* 5S RNA, implying major Mg^{2+}-induced conformational changes in regions other than the terminal helical stem. Recently, Kao and Crothers (1980) have proposed that certain conformational changes of 5S RNA modulated by salt ions or ribosomal proteins are needed to provide the necessary movement of ribosome relative to mRNA during protein biosynthesis. The presently observed effects of Mg^{2+} could provide a basis for such a conformational switch conjecture.

Future Prospects for NMR Determination of 5S RNA Secondary Structure. In the near future, it is unlikely that the highest available magnetic field for proton NMR will exceed ca. 14–17.5 T (ca. 600–750 MHz), and even that improvement would offer only a modest increase in chemical shift dispersion of the highly overlapped resonances in the 9–15 ppm hydrogen-bond proton region. Proton NMR characterization of purified RNA fragments obtained via specific enzymatic cleavage (Kime & Moore, 1983a; Kime, 1984a; Li & Marshall, 1986) offers a promising approach, provided that the fragments retain the structure of the intact 5S RNA molecule. Nitrogen-15 enrichment combined with ^1H–^{15}N two-dimensional FT NMR chemical shift correlation can aid in identification and assignment of ^1H NMR imino resonances (Griffey et al., 1983; Kime, 1984a,b). Differential chemical shifts induced by Mg^{2+} (Li & Marshall, 1986) or Na^+ (Lee and A. G. Marshall, unpublished results) can be very helpful in deciding how many protons correspond to a particular resonance. Site-specific spin-labeling (K. M. Lee and A. G. Marshall, unpublished results) produces proton NMR line broadening according to the electron–proton separation and can aid in identifying base pairs that are spatially near the spin-label. Comparison of 5S RNA spectra from species of slightly different primary sequence (S.-M. Chen and A. G. Marshall, unpublished results) can aid in identifying base pairs located at points of mutation. Finally, differential melting produced by heating any of the above samples can help in deciding which resonances correspond to protons in adjacent base pairs. We expect to report examples of each of these approaches shortly.

ACKNOWLEDGMENTS

We thank C. E. Cottrell of The Ohio State University Chemical Instrument Center for advice and assistance in optimizing NOE experiments.

REFERENCES

Arter, D. B., & Schmidt, P. G. (1976) *Nucleic Acids Res. 3*, 1437–1447.

Bothner-By, A. A. (1979) in *Biological Applications of Magnetic Resonance* (Shulman, R. G., Ed.) pp 177–219, Academic, New York.

Burkey, K. O., Marshall, A. G., & Alben, J. O. (1983) *Biochemistry 22*, 4223–4229.

Burns, P. D., Luoma, G. A., & Marshall, A. G. (1980) *Biochem. Biophys. Res. Commun. 96*, 805–811.

Chang, L.-H., & Marshall, A. G. (1986) *Biopolymers* (in press).

Chang, L.-H., Burkey, K. O., Alben, J. O., & Marshall, A. G. (1984) *Biochemistry 23*, 3659–3662.

Delihas, N., & Andersen, J. (1982) *Nucleic Acids Res. 10*, 7323–7344.

De Wachter, R., Chen, M.-W., & Vandenberghe, A. (1982) *Biochemie 64*, 311–329.

Erdmann, V. A. (1976) *Prog. Nucleic Acid Res. Mol. Biol. 18*, 45–90.

Erdmann, V. A., Wolters, J., Huysmans, E., & De Wachter, R. (1985) *Nucleic Acids Res. 13*, r105–r153.

Feigon, J., Denney, W. A., Leupin, W., & Kearns, D. R. (1983) *Biochemistry 22*, 5930–5942.

Fox, G. E., & Woese, C. R. (1975) *Nature (London) 256*, 505–507.

Garrett, R. A., Douthwaite, S., & Noller, H. F. (1981) *Trends Biochem. Sci. (Pers. Ed.) 6*, 137–139.

Geerdes, H. A. M., & Hilbers, C. W. (1979) *FEBS Lett. 107*, 125–128.

Giessner-Prette, C., & Pullman, B. (1970) *J. Theor. Biol. 27*, 87–95.

Griffey, R. H., Poulter, C. D., Bax, A., Hawkins, B. L., Yamaizumi, Z., & Nishimura, S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 5895–5897.

Hare, D. R., & Reid, B. R. (1982a) *Biochemistry 21*, 1835–1838.

Hare, D. R., & Reid, B. R. (1982b) *Biochemistry 21*, 5129–5135.

Hare, D. R., Wemmer, D. E., Chou, S.-H., Drobny, G., & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319–336.

Heerschap, A., Haasnoot, C. A. G., & Hilbers, C. W. (1982) *Nucleic Acids Res. 10*, 6981–7000.

Heerschap, A., Haasnoot, C. A. G., & Hilbers, C. W. (1983a) *Nucleic Acids Res. 11*, 4483–4499.

Heerschap, A., Haasnoot, C. A. G., & Hilbers, C. W. (1983b) *Nucleic Acids Res. 11*, 4501–4520.

Heus, H. A., Van Kimmenade, J. M. A., Van Knippenberg, P. H., Haasnoot, C. A. G., De Bruin, S. H., & Hilbers, C.

W. (1983) *J. Mol. Biol. 170*, 939–956.
Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., Marquisee, M., Merrill, S. H., Penswick, J. R., & Zamir, A. (1965) *Science (Washington, D.C.) 147*, 1462–1465.
Hori, H., & Osawa, S. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 381–385.
Hurd, R. E., & Reid, B. R. (1979) *Biochemistry 18*, 4017–4024.
Johnston, P. D., & Redfield, A. G. (1978) *Nucleic Acids Res. 5*, 3913–3927.
Johnston, P. D., & Redfield, A. G. (1981) *Biochemistry 20*, 1147–1156.
Kan, L.-S., Chandrasegaran, S., Pulford, S. M., & Miller, P. S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4263–4265.
Kao, T. H., & Crothers, D. M. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 3360–3364.
Kearns, D. R., & Wong, Y. P. (1974) *J. Mol. Biol. 87*, 755–774.
Kime, M. J., & Moore, P. B. (1983a) *Biochemistry 22*, 2615–2622.
Kime, M. J., & Moore, P. B. (1983b) *Biochemistry 22*, 2622–2629.
Kime, M. J. (1984a) *FEBS Lett. 173*, 342–346.
Kime, M. J. (1984b) *FEBS Lett. 175*, 259–262.
Kime, M. J., Gewirth, D. T., & Moore, P. B. (1984) *Biochemistry 23*, 3559–3568.
Ladner, J. E., Jack, A., Robertus, J. D., Brown, R. S., Rhodes, D., Clark, B. F. C., & Klug, A. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4414–4418.
Li, S.-J., & Marshall, A. G. (1985) *Biochemistry 24*, 4047–4052.
Li, S.-J., & Marshall, A. G. (1986) *Biochemistry* (in press).
Li, S.-J., Chang, L.-H., Chen, S.-M., & Marshall, A. G. (1984a) *Anal. Biochem. 138*, 465–471.
Li, S.-J., Burkey, K. O., Luoma, G. A., Alben, J. O., & Marshall, A. G. (1984b) *Biochemistry 23*, 3652–3658.
Luehrsen, K. R., & Fox, G. E. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 2150–2154.
Luoma, G. A., & Marshall, A. G. (1978a) *Proc. Natl. Acad. Sci. U.S.A. 75*, 4901–4905.
Luoma, G. A., & Marshall, A. G. (1978b) *J. Mol. Biol. 125*, 95–105.
Luoma, G. A., Burns, P. D., Bruce, R. E., & Marshall, A. G. (1980) *Biochemistry 19*, 5456–5462.
Luoma, G. A., Herring, F. G., & Marshall, A. G. (1982) *Biochemistry 21*, 6591–6598.
Marotta, C. A., Varricchio, F., Smith, I., Weissman, S. M., Sogin, M. L., & Pace, N. R. (1976) *J. Biol. Chem. 251*, 3122–3127.
Marshall, A. G., & Smith, J. L. (1980) *Biochemistry 19*, 5955–5959.
Mizuno, H., & Sundaralingam, M. (1978) *Nucleic Acids Res. 5*, 4451–4461.
Morikawa, K., Kawakami, M., & Takemura, S. (1982) *FEBS Lett. 145*, 194–196.
Nishikawa, K., & Takemura, S. (1974) *J. Biochem. (Tokyo) 76*, 935–947.
Noggle, J. H., & Schirmer, R. E. (1971) *The Nuclear Overhauser Effect: Chemical Applications*, Academic, New York.
Osterberg, R., Sjoberg, B., & Garrett, R. A. (1976) *Eur. J. Biochem. 68*, 481–487.
Patel, D. J., Pardi, A., & Itakura, K. (1982) *Science (Washington, D.C.) 216*, 581–580.
Pieler, T., & Erdmann, V. A. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4599–4603.
Redfield, A. G., & Kunz, S. D. (1979) in *NMR and Biochemistry* (Opella, S. J., & Lu, P., Eds.) pp 225–239, Marcel Dekker, New York.
Reid, B. R. (1981) *Annu. Rev. Biochem. 50*, 969–996.
Reid, B. R., McCollum, L., Ribeiro, N. S., Abbate, J., & Hurd, R. E. (1979) *Biochemistry 18*, 3996–4005.
Roth, K., Kimber, B. J., & Feeney, J. (1980) *J. Magn. Reson. 41*, 302–309.
Roy, S., & Redfield, A. G. (1983) *Biochemistry 22*, 1386–1390.
Salemink, P. J. M., Raue, H. A., Heerschap, A., Plant, R. J., & Hilbers, C. W. (1981) *Biochemistry 20*, 265–272.
Sanchez, V., Redfield, A. G., Johnston, P. D. & Tropp, J. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5659–5662.
Scheek, R. M., Boelens, R., Russo, N., van Boom, J. H., & Kaptein, R. (1984) *Biochemistry 23*, 1371–1376.
Schimmel, P. R., & Redfield, A. G. (1980) *Annu. Rev. Biophys. Bioeng. 9*, 181–221.
Studnicka, G. M., Eiserling, F. A., & Lake, J. A. (1981) *Nucleic Acids Res. 9*, 1885–1904.
Sussman, J. L., & Kim, S.-H. (1976) *Biochem. Biophys. Res. Commun. 68*, 89–96.
Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., & Gralla, J. (1973) *Nature (London), New Biol. 146*, 40–41.
Wong, Y. P., Kearns, D. R., Reid, B. R., & Shulman, R. G. (1972) *J. Mol. Biol. 72*, 741–749.
Wrede, P., & Erdmann, V. A. (1973) *FEBS Lett. 33*, 315–319.