Jovin, T. M., Englund, P. T., & Bertsch, L. L. (1969) *J. Biol. Chem. 24*, 2996–3007.
Joyce, C. M., & Grindley, N. D. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1830–1834.
Joyce, C. M., & Steitz, T. A. (1987) *Trends Biochem. Sci.* 288–292.
Klenow, H., & Henningson, I. (1970) *Proc. Natl. Acad. Sci. U.S.A. 65*, 168–175.
Kornberg, A. (1980) in *DNA Replication*, pp 104–105, Freeman, San Francisco.
Kuchta, R. D., Mizrahi, V., Benkovic, P. A., Johnson, K. A., & Benkovic, S. J. (1987) *Biochemistry 26*, 8410–8417.
Mizrahi, V., Benkovic, P. A., & Benkovic, S. J. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 231–235.
Moffatt, J. G. (1964) *Can. J. Chem. 42*, 599–604.
Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G., & Steitz, T. A. (1985) *Nature 313*, 762–766.
Onodera, M., Shiokawa, H., & Takagi, T. (1976) *J. Biochem. 79*, 195–201.
Setlow, P., Brutlag, D., & Kornberg, A. (1972) *J. Biol. Chem. 247*, 224–248.

# An NMR Study of the Helix V–Loop E Region of the 5S RNA from *Escherichia coli*†

P. Zhang and P. B. Moore*
*Department of Chemistry and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 05611*

*Received November 21, 1988; Revised Manuscript Received February 13, 1989*

ABSTRACT: Experiments are described that complete the assignment of the imino proton NMR spectrum of the fragment 1 domain from the 5S RNA of *Escherichia coli*. Most of the new assignments fall in the helix V–loop E portion of the molecule (bases 70–78 and 98–106), the region most sensitive to the binding of ribosomal protein L25. The spectroscopic data are incompatible with the standard, phylogenetically derived model for 5S RNA, which makes all the base pairs possible in loop E with the sequences aligned in parallel (C70–G106, C71–G105, etc.) [see Delihas et al. (1984) *Prog. Nucleic Acid Res. Mol. Biol. 31*, 161–190]. Furthermore, the alternative loop E model proposed for spinach chloroplast 5S RNA by Romby et al. [(1988) *Biochemistry 27*, 4721–4730] does not apply to the closely homologous 5S RNA from *E. coli*. The 5S RNAs from *E. coli* and spinach chloroplasts do not have the same secondary structures in solution despite their strong sequence homologies, and neither appears to conform to the standard model for 5S RNA in the loop E region.

Figure 1 shows the sequence of the 5S RNA from *Escherichia coli* drawn in the secondary structure derived for it by comparative sequence analysis. Like the larger rRNAs, it is a series of stem–loop structures, and its hydrogen-bonded stems tend to be short and to include non-Watson–Crick base pairs [see Delihas et al. (1984)]. We have been studying this molecule by nuclear magnetic resonance (NMR),[1] concentrating on the portion that resists attack by RNase A, fragment 1 (bases 1–11 and 69–120). Fragment 1 has a three-dimensional structure similar to that of the same sequences in intact 5S RNA, and it binds ribosomal protein L25, the only one of the three 5S binding proteins whose binding site it fully contains (Kime & Moore, 1983a–c).

Fragment 1 includes the helix IV–helix V stem of 5S RNA. Helix IV (bases 79–85 and 91–97) is normal double helix [see Kime and Moore (1983b)], but the secondary structure of the rest of the stem is less obvious. It has two G-C juxtapositions (G106–C70 and G105–C71) at its proximal end, but as one proceeds toward helix IV aligning the strands as shown in Figure 1, only two conventional base pairs can be made. For this reason, there have been differences of opinion about the length of helix V and the size of the unpaired region between helices IV and V (loop E). Nevertheless, helix V and loop E resist enzymatic attack as though they were helical (Kjems et al., 1985), and some archaebacterial 5S RNAs have sequences that permit nearly perfect Watson–Crick pairing in this region (Stahl et al., 1981). Furthermore, the bases of loop E contribute several resonances to the imino proton NMR spectrum of fragment 1, which are particularly conspicuous in complexes of fragment 1 with L25, consistent with the view that loop E is significantly structured (Gewirth et al., 1987). Thus, the helix V–loop E region is shown in Figure 1 heavily hydrogen bonded.

Below we report the results of investigations of $^{15}N$-labeled fragment 1 samples in $Ca^{2+}$-containing buffers at low temperature. Under these conditions, resonances and NOEs can be detected in the free RNA that are not seen in $Mg^{2+}$ at room temperature, the conditions of most of our previous experiments. Most of these "new" resonances represent loop E protons that are also detectable in the spectrum of the fragment 1–L25 complex in $Mg^{2+}$ at room temperature. The new data enable us to complete the assignment of the imino proton spectrum of fragment 1 we began several years ago (Kime & Moore, 1983b). Spectroscopic results on several mutant RNAs

†This work was supported by grants from the National Science Foundation (DMB-860823) and the National Institutes of Health (AI-09167 and GM-22778).

*Address correspondence to this author at the Department of Chemistry, Yale University, 225 Prospect St., New Haven, CT 06511.

[1] Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; poly(U), poly(uridylic acid).

0006-2960/89/0428-4607$01.50/0 

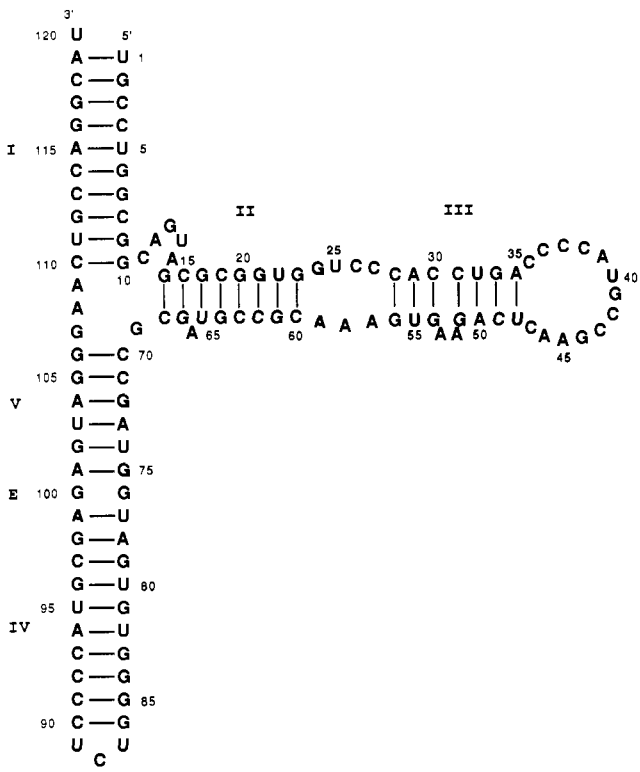


FIGURE 1: Secondary structure of the 5S RNA from *E. coli*. The sequence of the 5S RNA from *E. coli* is shown drawn in the three-stem secondary structure suggested for 5S RNA from sequence comparisons. The version shown here is that proposed by Delihas et al. (1984). Helical regions are designated by roman numerals. The only loop of interest here is denoted by an upper case E.

are presented in support of these assignments.

These new assignments permit us to evaluate the validity of the standard model of *E. coli* 5S RNA in the loop E region and to comment on the applicability to *E. coli* 5S RNA of the alternative model for loop E proposed recently by Romby et al. (1988). Neither is compatible with our data. It is possible that prokaryotic 5S RNAs in solution do not have the same secondary structures in the loop E region despite their strong sequence homologies.

## Materials and Methods

*Construction of Mutant Plasmids.* A plasmid (M13mp9) carrying a gene for a 5S RNA double mutant (A72, A67) was given to us by Kathy Triman (U. C. Santa Cruz). Mutants of pDG07 RNA were created by changing position 102 from G to A, and position 74 from U to C, both by site-directed mutagenesis in M13mp18 (BRL) (Gewirth & Moore, 1987; Gewirth, 1988). The A102 and C74 mutants were selected by using the method of Kunkel (1985), and confirmed by sequencing. In all three cases, *Hin*dIII fragments containing 5S or pDG07 RNA genes were excised from the corresponding double-stranded M13 and used to replace the *Hin*dIII fragment in pKK5-1 that contains the rrnB 5S gene (Brosius, 1984).

*Preparation of RNAs and Proteins.* The methods used for producing 5S RNA from cells containing pKK5-1 and for making $^{15}N$-labeled 5S RNA have been described elsewhere (Moore et al., 1988). The RNA products of mutant plasmids were produced the same way. Fragment 1 was made from intact 5S RNA by digestion with RNase A (Kime & Moore, 1983b). Digestions were done at 4 °C for 45 min at an enzyme to RNA ratio of 1:100. Mutant pDG07 RNAs were purified by HPLC on a Nucleogen DEAE 500-10 column using the conditions described previously (Gewirth & Moore, 1987). The method used for preparing ribosomal protein L25 may be found elsewhere (Moore et al., 1988). The $^{15}N$-labeled fragment 1 molecules used here were only partially labeled. Isotopic hybrids were prepared by partial reconstitution (Gewirth et al., 1987).

*Spectroscopic Samples and Methods.* Samples were prepared for spectroscopy by dialysis either into 0.1 M KCl, 4 mM $MgCl_2$, and 5 mM cacodylate, pH 7.0, or into 0.1 M KCl, 4 mM $CaCl_2$, and 5 mM cacodylate, pH 7.0, at RNA concentrations between 0.06 and 0.15 mM. In cases where RNA complexes with ribosomal protein L25 were to be studied, renatured L25 were added in a 1.2-fold molar excess. Samples were brought to their final concentrations (1 mM or higher) by ultrafiltration using Centricon-10 ultrafilters (Amicon). The solutions were made 5% in $D_2O$ for the spectrometer lock, and a small amount of dioxane was added as a chemical shift standard. It is assumed that the chemical shift of dioxane is 3.741 ppm relative to the methyl resonance of DSS at all temperatures.

Spectra were obtained in the Fourier-transform mode on the 490-MHz and 500-MHz NMR spectrometers of the Yale University Chemical Instrumentation Center. $^1H$-observe, $^{15}N$-decouple experiments were performed by using a Cryomagnet System, Inc., HR-44 wide-band-decouple, proton-observe, 5-mm probe. Imino proton spectra were obtained by using the twin-pulse method for water suppression (Kime & Moore, 1983a). Spectra were taken in 8K blocks with a spectral width of 15 000 Hz, and the offset at about 15 ppm. Nuclear Overhauser effects (NOEs) were obtained by the one-dimensional difference method. On- and off-resonance spectra were collected in an interleaved manner. Resonances were preirradiated for 0.2 s at a decoupler power level adjusted to give 50% saturation.

$^{15}N$–$^1H$ chemical shift correlations were measured by the difference decoupling method [see Griffey et al. (1985)]. Samples were irradiated at $^{15}N$ frequencies at low power (0.1 W) during acquisition of $^1H$ spectra. On- and off-resonance spectra were acquired sequentially. Off-resonance spectra were subtracted from on-resonance spectra to obtain the difference spectra shown below.

## Results

*Imino Proton Spectra in $Ca^{2+}$ at Low Temperature.* Figure 2 shows how the imino proton spectrum of fragment 1 in 4 mM $Ca^{2+}$ varies with temperature below 303 K. Downfield of 11.8 ppm, it remains constant as the temperature falls below 303 K, except for the expected increases in line width. Upfield of 11.8 ppm, other changes are evident. First, the doublet of sharp resonances at 11.7 ppm (O and P1) collapses to a singlet due to small changes in chemical shift. (Correspondences between resonances under different conditions were verified by NOE experiments.) Second, a resonance at 11.6 ppm which is broad and weak at 303 K (P2) narrows, and the intensity observable at the chemical shift of P2 increases due to the appearance of a new resonance, designated P3. The components of the unassigned, multiproton resonance at 10.5 ppm (R) alter their chemical shifts, leading to the generation of several "new" resonances having chemical shifts between 10.2 and 11.7 ppm, and the intensity in the R region also increases. These resonances are designated R4, R8, (R1b, R2), and R1a. Similar effects are seen in buffers containing $Mg^{2+}$ instead of $Ca^{2+}$ when the temperature falls, but they are less exaggerated [see Figure 8 in Kime and Moore (1983b)].

We have made a synthetic helix IV–helix V stem loop by transcribing an appropriate DNA template with T7 RNA polymerase in vitro. This RNA's sequence includes all of helix

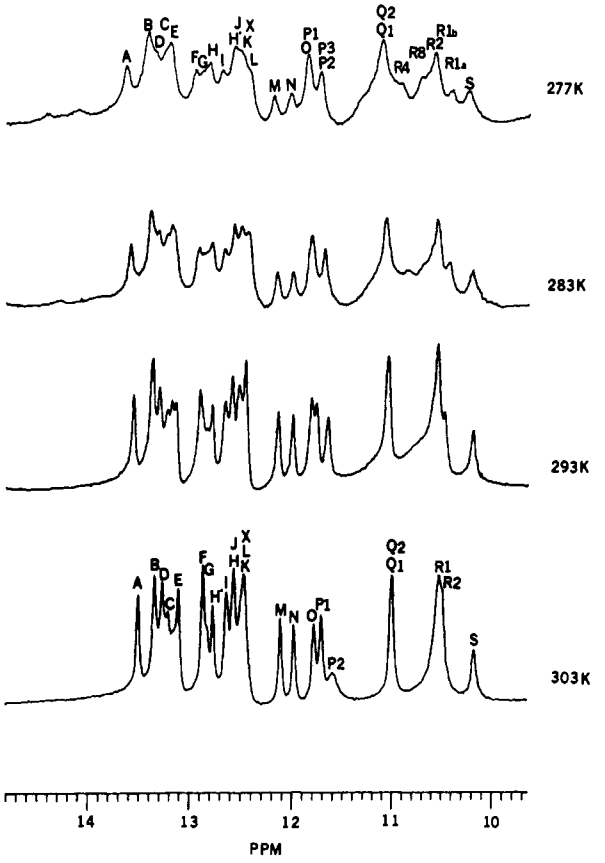


FIGURE 2: Imino proton spectrum of fragment 1 as a function of temperature in $Ca^{2+}$. A sample of fragment 1 was prepared at a concentration of about 1 mM in the $Ca^{2+}$-containing buffer described under Materials and Methods. Downfield spectra were then acquired at the temperatures shown. Resonances in the spectra taken at 303 and 277 K are identified by letters according to the standard convention (Kime & Moore, 1983b; Gewirth et al., 1987).

IV and all but G106–C70 of helix V. Its termini include a few bases not found in the 5S RNA sequence to ensure efficient transcription, but its imino proton spectrum is virtually identical with that of the corresponding sequences in intact 5S RNA at room temperature. At low temperatures in $Ca^{2+}$, its spectrum shows five well-resolved resonances between 11 and 10 ppm (data not shown). The unimpressive shoulders in the low-temperature spectrum of intact fragment 1 in that chemical shift range are real, and they come from the helix V region.

Temperature effects in the downfield spectrum of the fragment 1–L25 complex are less dramatic than in the free RNA. As the temperature decreases, some of the resonances between 10 and 11 ppm change their chemical shifts somewhat and resonances broaden, as usual, but no new resonance intensity appears (see Table I).

Since virtually every imino proton in the fragment 1–L25 complex gives a resonance at room temperature (Gewirth et al., 1987), most, if not all, of the resonances that become observable in the spectrum of the free RNA as the temperature falls must correspond to resonances in the fragment 1–L25 spectrum at room temperature. The purpose of the experiments described below is to determine what those correspondences are and to assign the resonances in both spectra that were previously unassigned.

*Identification of UN3 Resonances Originating in Strand III.* The experiments described below make use of fragment 1 molecules segmentally labeled with $^{15}N$. The nucleolytic digestion necessary to make fragment 1 from 5S RNA yields a product which is a complex of three oligonucleotides. Its

Table I: Imino Proton Chemical Shifts for the Fragment 1–L25 Complex in $Ca^{2+}$ at 303 and 277 K[a]

| resonance | assignment | chemical shift (ppm) at 303 K | chemical shift (ppm) at 277 K |
|---|---|---|---|
| J | G2 | 12.57 | 12.51 |
| C | G117 | 13.21 | 13.21 |
| F | G116 | 12.86 | 12.95 |
| B | U5 | 13.35 | 13.43 |
| M | G6 | 12.10 | 12.16 |
| E | G7 | 13.11 | 13.21 |
| H | G111 | 12.76 | 12.78 |
| P2 | U110 | 11.69 | 11.66 |
| R2 | G9 | 10.50 | 10.54 |
| T | G10 | – | 13.21 |
| L | G106 | 12.47 | 12.46 |
| K | G105 | 12.47 | 12.46 |
| R1b/5 | G72 | 10.40 | 10.44 |
| X | U103 | 12.57 | 12.60 |
| R1a/6 | G102 | 10.16 | 9.97 |
| R8/8 | U74 | 9.83 | 9.95 |
| P3/3 | G75 | 11.13 | 11.23 |
| S/7 | G76 | 10.04 | 10.06 |
| H′ | U77 | 12.76 | 12.86 |
| R4/4 | G98 or G100 | 10.73 | 10.84 |
| D | G79 | 13.11 | 13.17 |
| O | U80 | 11.78 | 11.92 |
| Q1 | G96 | 11.13 | 11.21 |
| P1 | U95 | 11.52 | 11.66 |
| Q2 | G81 | 11.05 | 11.12 |
| A | U82 | 13.50 | 13.62 |
| N | G83 | 11.97 | 11.99 |
| I | G84 | 12.57 | 12.62 |
| G | G85 | 12.82 | 12.84 |
| R9/9 | U87 | – | 11.00 |

[a] Chemical shifts are given for the imino protons in the spectrum of the fragment 1–L25 complex in 0.1 M KCl, 4 mM $CaCl_2$, and 10 mM cacodylate, pH 7.0, at two different temperatures. The identities of resonances were established by NOE experiments and follow the conventions established earlier (Kime & Moore, 1983b,c). Entries marked "–" correspond to resonances that are not observable under the condition specified. The top block of assignments corresponds to helix I, the middle block to helix V–loop E, and the bottom block to helix IV. Data supporting the assignments in helices I and IV may be found elsewhere [see Gewirth et al. (1987) and references cited therein]. Most of the assignments in helix V–loop E are described in this paper.

component oligonucleotides are known as strand II (bases 89–120), strand III (bases 69–87), and strand IV (bases 1–11) (see Figure 1). Strand III dissociates from the other two in buffers that lack $Mg^{2+}$ and reassociates when $Mg^{2+}$ is added back. We take advantage of this phenomenon to make isotopic hybrids of fragment 1 in which either strand III or strands II and IV are $^{15}N$ labeled (Moore et al., 1988).

All imino proton resonances originating from an $^{15}N$-labeled oligonucleotides are 80–90-Hz doublets due to coupling of the protons with the $^{15}N$'s to which they are covalently bonded. These doublets can be decoupled by applying energy to the sample at $^{15}N$ frequencies. A decoupled spectrum minus an undecoupled spectrum, i.e., a difference decoupling spectrum, will have a positive peak of weight 2 flanked by two negative peaks of weight 1 for every resonance affected by the radio-frequency energy applied at $^{15}N$ frequencies. The UN3 $^{15}N$ resonances are tightly clustered in the $^{15}N$ domain and do not overlap with the cluster for GN1 $^{15}N$'s (Gonnella et al., 1982). Thus, measurement of the $^{15}N$ frequencies at which different imino proton resonances decouple identifies them as to base type.

Figure 3 shows some difference decoupling spectra taken at 277 K in $Ca^{2+}$- on $^{15}N$-labeled fragment molecules, with L25 (right side) and without L25 (left side). The fragment molecules in question are $^{15}N$ labeled in strand III, and the

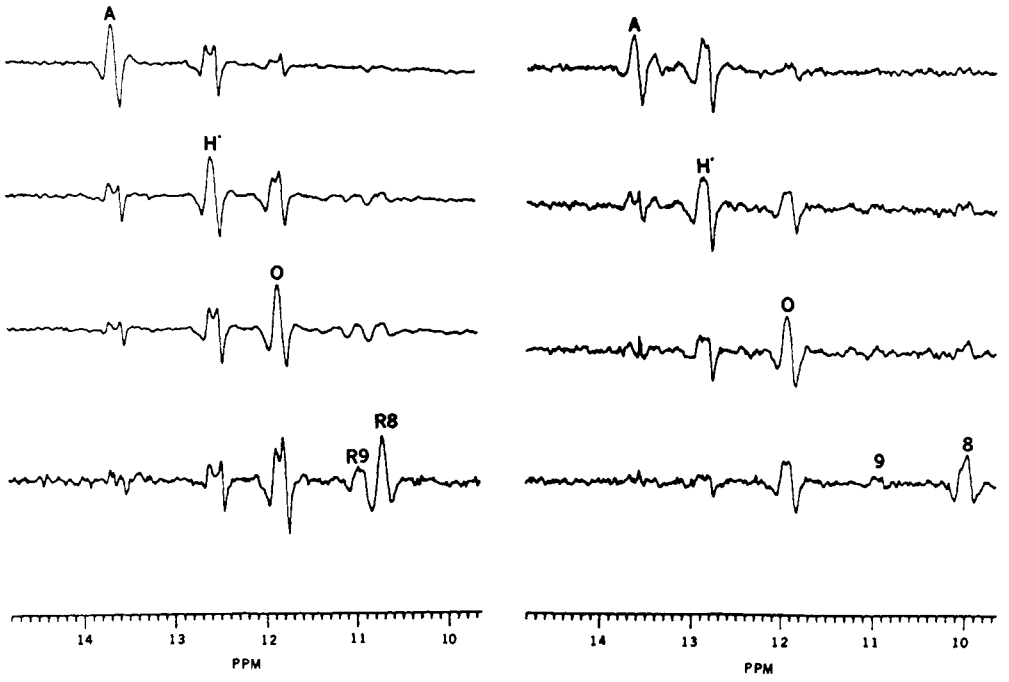


FIGURE 3: UN3 difference decoupling spectra for fragment 1 and the fragment 1–L25 complex $^{15}N$ labeled in strand III. Difference decoupling spectra were obtained as described under Materials and Methods on samples in $Ca^{2+}$ at 277 K. The left-hand spectra were given by fragment 1 $^{15}N$ labeled in strand III (bases 69–87). The right-hand series of spectra were obtained from the fragment 1–L25 complex again labeled with $^{15}N$ in strand III. Difference spectra are shown for four different settings of the $^{15}N$ decoupler frequency, all in the region characteristic of UN3 nitrogens. The resonances most perfectly decoupled are identified in each spectrum. The $^{15}N$ frequency settings are the same as in the left-hand series of spectra.

decoupling frequencies explored are in the UN3 region. Spectra on the two sides of the figure at the same vertical level were taken at the same $^{15}N$ decoupling frequency. Five UN3 signals are seen in the spectrum of both the free RNA and the RNA complex with L25. Four of them are quite intense (A, H′, O, and R8), and the fifth is a weak, broad signal at about 11 ppm labeled R9 in the spectrum of the RNA and 9 in the spectrum of the complex. The correspondence of the $^{1}H$ and $^{15}N$ chemical shifts of the five UN3 resonances are obvious.

[Resonances in the spectrum of fragment 1 are identified by letter (Kime et al., 1983b; Gewrith et al., 1987), and resonances in the spectrum of the fragment 1–L25 complex have both letter and number designations (Kime & Moore, 1983c). Numbered resonances are resonances in the spectrum of the complex whose correspondents in the spectrum of fragment 1 were unknown when they were first observed. Resonances with the same (letter) name in the two spectra represent the same protons. In cases where subsequent experiments have established the correspondences between numbered and lettered resonances, resonances are often referred as to letter name/number name, e.g., S/7 (Gewrith et al., 1987).]

These data raise several points. First, the UN3 spectra for the protein–RNA complex shown here are the same as those obtained for similarly labeled samples at room temperature in $Mg^{2+}$ except for modest differences in chemical shift, and the very weak signal at 11 ppm, labeled 9 or R9 (Gewirth et al., 1987; Gewirth, 1988). Second, neither the strong UN3 signal at 10.7 ppm, designated R8, nor the weak 11 ppm signal is observed in the spectrum of fragment 1 at room temperature in either $Mg^{2+}$ or $Ca^{2+}$ [see Figure 6 of Gewirth et al. (1987)]. Proton exchange must be so fast at room temperature at both positions that their resonances cannot be observed. Third, because of their similarities in chemical shift in both the $^{1}H$ and $^{15}N$ frequency domains, it is likely that R8 in the free RNA and 8 in the protein complex represent the same proton, as do R9 and 9. Finally, there are five U's in strand III. UN3 resonances A and O were assigned by nuclear Overhauser methods long ago (Kime & Moore, 1983b); they represent U82 and U80, respectively. Resonances H′, R8/8, and R9/9 must represent U74, U77, and U87.

*Identification of Strand III GN1 Resonances.* The top three spectra in Figure 4 show some difference decoupling spectra obtained on the same samples, and under the same conditions as for Figure 3 except that the $^{15}N$ decoupling frequencies explored are in the GN1 range. The resonances upfield of 12 ppm are of particular interest. In the free RNA in $Mg^{2+}$ at 303 K, only resonances Q2 and S can be observed in such experiments. In $Ca^{2+}$ at 277 K, however, two new resonances appear, P3 and R1b. On the basis of their chemical shifts in both the $^{1}H$ and $^{15}N$ domains, P3 must correspond to resonance 3 in the complex spectrum, and R1b to resonance 5. In difference spectra comparing normal fragment 1 and fragment 1 $^{15}N$ labeled in strand III in $Ca^{2+}$ at room temperature, a weak, broad resonance can be seen at the R position that must represent R1b since R1b is the only resonance split by strand III labeling that has a chemical shift anywhere near R (data not shown).

The $^{15}N/^{1}H$ spectra obtained with the fragment–L25 complex under these conditions differ from those taken at room temperature in $Mg^{2+}$ in one respect only. There is a clear signal for resonance 5 at low temperature that is not seen at high temperature. Resonance 5 must therefore represent a strand III GN1 proton. The strand III G residues whose imino protons have not been assigned previously are 72, 75, 76, and 86, and we know from prior data that S/7 must be either G75 or G76 (Gewirth et al., 1987). P3/3 and R1b/5 must account for all but one of the three that will remain unassigned once S/7 has been put in place.

*Classification of Strand II and IV Resonances at Low Temperature in $Ca^{2+}$.* There are seven U residues in the strand II–IV moiety of fragment 1: U1, U5, U89, U95, U103, U111, and U120. Four of them give resonances that have been

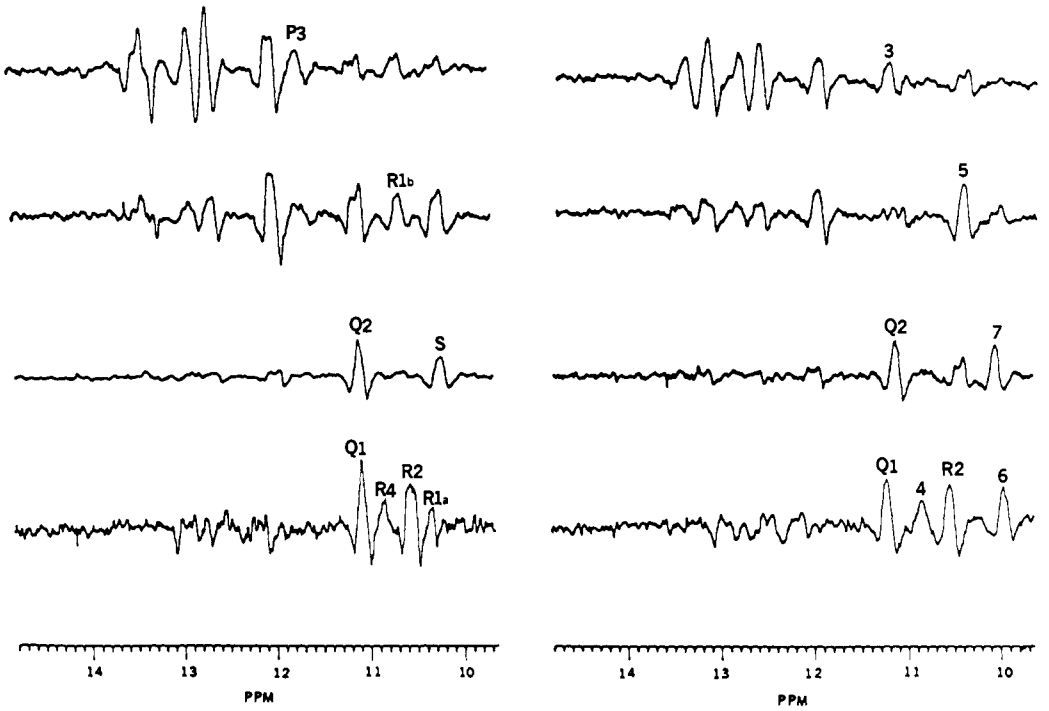


FIGURE 4: GN1 difference decoupling spectra for fragment 1 and the fragment 1–L25 complex $^{15}N$ labeled in strand III and strands II and IV. The top three spectra shown were obtained from the same samples, and under the same conditions as the spectra shown in Figure 3. Spectra were taken with the $^{15}N$ decoupler set at frequencies characteristic of GN1 nitrogens. The left-hand series of spectra were obtained with fragment 1, and the right-hand series were obtained with the fragment 1–L25 complex. The $^{15}N$ decoupling frequencies are the same in spectra at the same horizontal level. The bottom spectrum was obtained by using a sample $^{15}N$ labeled in strands II and IV with the decoupler set at GN1 frequencies.

assigned previously; resonances B, X, P1, and P2 correspond to bases U5, U103, U95, and U111, respectively (Gewirth et al., 1987). All give signals in the UN3-decoupled differences spectra of fragment 1 labeled with $^{15}N$ in strands II and IV, both in the presence and in the absence of L25, in $Ca^{2+}$ at low temperature (data not shown). Nothing is seen that might be ascribed to U1, U89, or U120.

The low-temperature, $Ca^{2+}$, GN1-decoupled spectra of fragment 1 $^{15}N$ labeled in strands II and IV are more interesting. The regions below 12 ppm in the spectra of both the free RNA and the complex are similar but too complex to resolve. Nine assigned resonances belonging to strands II and IV are to be found in that region (Gewirth et al., 1987). Upfield of 12 ppm in the free RNA spectrum, there are only two resonances belonging to strands II and IV that decouple at GN1 frequencies at room temperature, resonances Q1 and R2. At low temperature, however, there are four. The two "new" ones are designated R4 and R1a (see Figure 4, bottom spectrum). The corresponding GN1-decoupled spectra for the L25 complex are essentially the same as at room temperature. On the basis of chemical shift comparisons, there can be little doubt that R4 corresponds to resonance 4, and R1a is the same as 6.

*Nuclear Overhauser Effects in Helix V.* As we have just seen, the $^{15}N$ difference decoupling data identify six resonances in fragment 1 that are visible only at low temperatures in the presence of $Ca^{2+}$, correlate five of the six with resonances detectable both at high and at low temperatures in the fragment 1–L25 complex, and assign all with respect to base type and strand of origin. NOE data are the only additional information available to assist in assigning these resonances. The NOEs that have been observed involving resonances 3–9 and their free RNA equivalents are listed in Table II.

*Assignment of Resonances 3–9.* The NOE "walking" method for assigning imino proton resonances assumes that a base whose imino proton gives an NOE to the imino proton of a given base pair will normally be an immediate neighbor of one of the members of that base pair in the molecule's sequence [see Patel et al. (1987)]. It is the fitting of the observed pattern of NOEs to the sequence that results in assignments. What is certain about the helix V region is that it is not conventional double helix. Thus, there is no reason that imino proton resonances related by NOEs must represent adjacent bases in the sequence. We will nevertheless interpret our data that way and rely on the fit of the data to the sequence to convince the reader that the assignments proposed are correct.

Table II: Nuclear Overhauser Effects Involving Numbered Resonances[a]

| resonance | NOE | conditions of observation |
|---|---|---|
| P3/3 | R1a/6 | cmplx, $Ca^{2+}$, 303 K |
| R4/4 | none | |
| R1b/5 | none | |
| R1a/6 | X | cmplx, $Mg^{2+}$, 303 K; RNA, $Ca^{2+}$, 303 K |
| R1a/6 | P3/3 | cmplx, $Ca^{2+}$, 303 K |
| S/7 | H′ | RNA, cmplx, $Ca^{2+}$, $Mg^{2+}$, 303 K |
| R8/8 | X | cmpx, $Ca^{2+}$, $Mg^{2+}$, 303 K |
| R9/9 | none | |

[a] Resonances are identified by both their free RNA and their RNA–protein complex names, when they are different. For each NOE observed, the table lists the type of sample in which it is seen (cmplx = L25–fragment 1 complex; RNA = free RNA), the divalent cations that must be present, and the highest temperature at which it has been observed. The NOEs to H′ and X have all been seen in the reverse of the sense listed in the table.

The key to the assignments in this region is resonance X. On the basis of earlier $^{15}N$ data, it is assigned to the imino proton of U103. Resonance 6 in the spectrum of the complex and its correspondent in the spectrum of the free RNA, R1a, give NOEs to resonance X. Since R1a/6 represents a strand II GN1 proton, it is likely to be G102, as we argued previously (Gewirth et al., 1987).

Resonance 6 gives a strong, reciprocal NOE to resonance 3, which represents a strand III GN1. The G on strand III

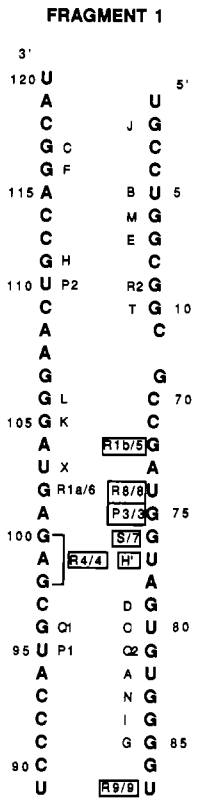


FIGURE 5: Assignments in fragment 1. The assignments for imino proton resonances in the fragment 1 spectrum are shown on a model for the molecule's secondary structure. The boxed-in assignments are substantiated in this paper. Evidence for the unboxed assignments may be found elsewhere [see Gewirth et al. (1987) and references cited therein].

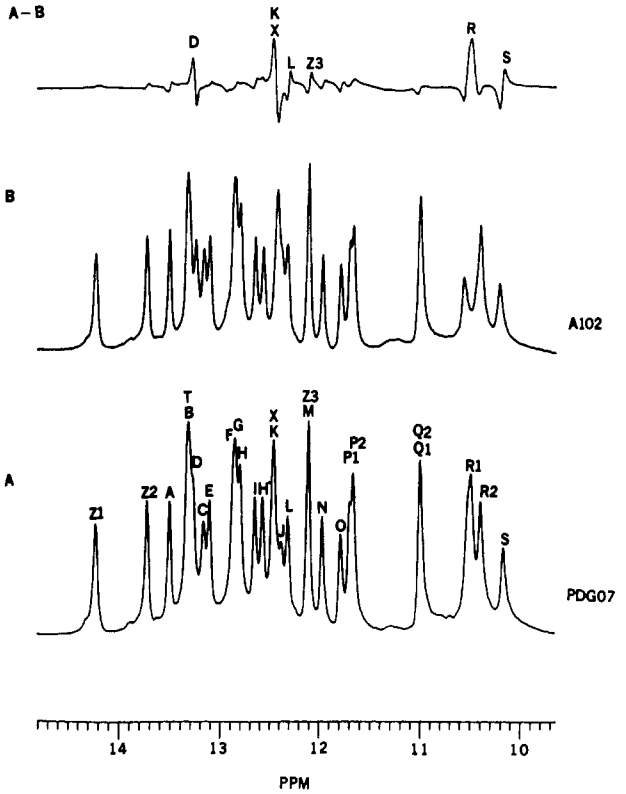


FIGURE 6: Comparison of the downfield spectrum of pDG07 RNA and a mutant containing an A at position 102. All spectra were obtained in $Ca^{2+}$ at 303 K. The bottom spectrum is that of pDG07 RNA. All its resonances are identified according to the standard convention (Kime & Moore, 1983b; Gewirth et al., 1987; Gewirth & Moore, 1987). The middle spectrum is that of the mutant. The top spectrum is the difference between the two spectra (pDG07 – mutant).

nearest to G102 is G75, when the sequence is aligned as shown in Figure 1. It has been recognized for some time that resonance S/7 represents either G75 or G76 (Gewirth et al., 1987). If G75 is resonance P3/3, G76 must be S/7. S/7 gives a strong reciprocal NOE to resonance H′, a strand III UN1 resonance. It follows that H′ represents the N3 proton of U77. The assignment of resonance H′ to U75 leaves only one assignment possible for R8/8, namely, U74. The 8 to X NOE we have observed would appear to confirm this scheme.

R1b/5 is a strand III GN1. At this point, there are only two GN1's that remain unassigned to strand III, G72 and G86. We suggest that the N1 proton of G72 is responsible for resonance R1b/5 but must rely on mutant data to prove it (see below). Within helix V, there are two GN1s on strand II that could give rise to R4/4, G98 and G100. The data do not permit us to choose between these alternatives.

One strand III UN3 resonance remains to be placed, R9/9. The characteristics of R9/9 are (1) that it is unique to low-temperature spectra, (2) that it is the broadest, least well-defined of all the UN3 resonances originating in that strand, and (3) that its chemical shift is 11 ppm. These characteristics are appropriate for the imino proton of a U residue that makes no hydrogen bonds, and we believe that R9/9 should be assigned to U87.

*Resonances K and L in Fragment 1.* Resonances K and L represent base pairs G106–C70 and G105–C71 (Gewirth & Moore, 1987). There are no interpretable NOEs to K or L in the spectra of fragment 1 or pDG07 RNA that can be used to assign K and L more specifically. The thermal melting behavior of these resonances is suggestive, however. Resonance K melts at a lower temperature than resonance L in pDG07–A102 RNA (data not shown). Since G105–C71 is adjacent to loop E, which is only marginally stable, it should melt before G106–C70. Thus, L is likely to represent the imino proton of G106, and K the imino proton of G105. Figure 5 summarizes the assignments in fragment 1.

*Fragment 1 Mutants.* Several sequence variants of fragment 1 have been prepared to test the assignments just proposed. The first is a mutant of pDG07 RNA that has an A at position 102 instead of a G. pDG07 RNA is the product of a gene for 5S RNA from which the sequences that code for the helix II–III arm of the parent molecule have been deleted. The sequence of pDG07 RNA is the same as that of fragment 1 except that G10 is connected to G70 by a three-base insert (5′UUC3′). Its NMR spectrum is similar to that of fragment 1, and it binds protein L25 normally (Gewirth & Moore, 1987).

Figure 6 compares the downfield spectra of the G102 mutant and "wild-type" pDG07 RNA in $Ca^{2+}$ at 303 K. Under these conditions, a substantial amount of intensity is lost from R1, as expected, and resonances S, D, K, L, and X all change their chemical shifts. When the same pair of RNAs are compared in $Mg^{2+}$, the loss of R1 intensity is less marked because R1a is weak under those conditions, resonance X disappears, and there are chemical shift changes for resonances D, H′, K, L, Z3, and S. If S and H′ represented G75 and U74, respectively, one would anticipate that H′ might well be lost in such a mutant, and S lost or dramatically altered in chemical shift. The fact that both are perturbed in only a secondary way suggests that S and H′ are correctly assigned in the current scheme.

Fragment 1 was prepared from a mutant of 5S RNA, which has an A instead of a G at position 72, and its downfield

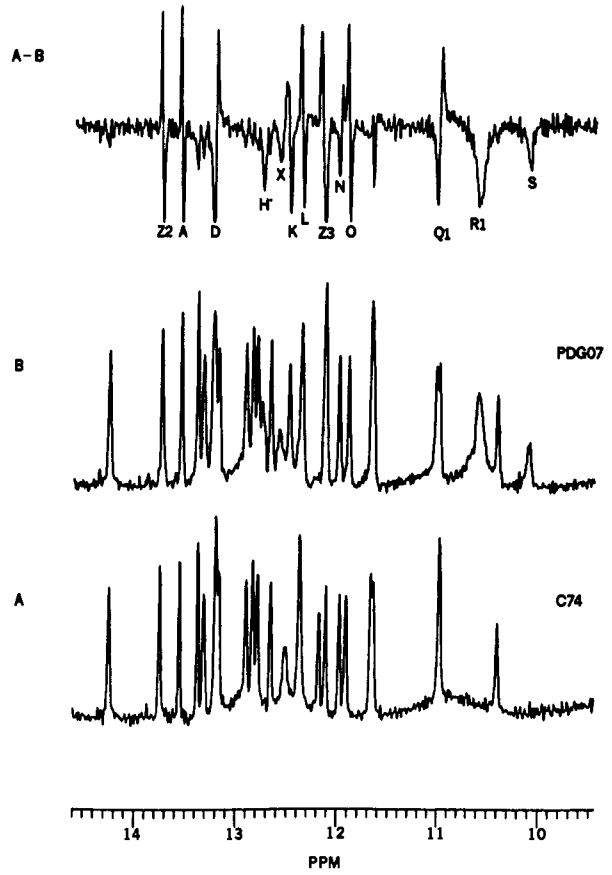


FIGURE 7: Comparison of the downfield spectrum of pDG07 RNA and a mutant containing a C at position 74. Spectra were obtained at 303 K in $Mg^{2+}$. The bottom spectrum (A) is that of the mutant while the middle spectrum (B) is that of wild-type pDG07 RNA. Both spectra are strongly resolution enhanced to make comparison easier. Information on the naming and assignment of pDG07 RNA imino proton resonances may be found in Gewirth and Moore (1987). The difference between the two spectra (A – B) is the top spectrum in the figure.

spectrum was compared with that of wild-type fragment 1 in $Ca^{2+}$ at room temperature. The mutant's spectrum showed a clear loss of intensity at the chemical shifts of R1, K, and L, and some small chemical shift changes on the part of several other resonances (data not shown). Since R1b/5 is detectable under these conditions, and the bases responsible for L and K are near position 72, this result strongly supports the assignment of R1b/5 to G72.

Another pDG07 derivative of interest has a C at position 74 instead of a U. The standard model for 5S RNA (Figure 1) suggests that U74 (R8/8) should make a wobble base pair with G102, the G we assign to R1a/6. If this were true, replacement of U74 with a C should stabilize the loop E–helix V region, since GU and GC base pairs have closely similar geometries and GC's are more stable than GU's. Figure 7 shows the result. The spectrum of the C74 mutant includes none of the resonances characteristic of loop E, resonances H', X, and S. The normal structure of loop E is incompatible with Watson–Crick base pairing at the U74–G102 position.

*Effect of L25 Binding on Chemical Shifts in the Downfield Spectrum of Fragment 1.* We have pointed out in the past that the binding of L25 to fragment 1 (or 5S RNA) alters its imino proton spectrum. The chemical shifts of many resonances change, indicating that the three-dimensional structure of the molecule is perturbed when the protein binds (Kime & Moore, 1983c). Furthermore, the rate of solvent exchange of many imino protons is reduced when L25 binds (Leontis & Moore, 1986), which presumably is the cause of the appearance of the numbered resonances. Now that we know which loop E resonances are detectable in free fragment 1, and what their chemical shifts are, we can assess the effect of protein binding on loop E. Figure 8 shows the result for fragment 1 and fragment 1/L25 complex in $Ca^{2+}$ at 277 K.

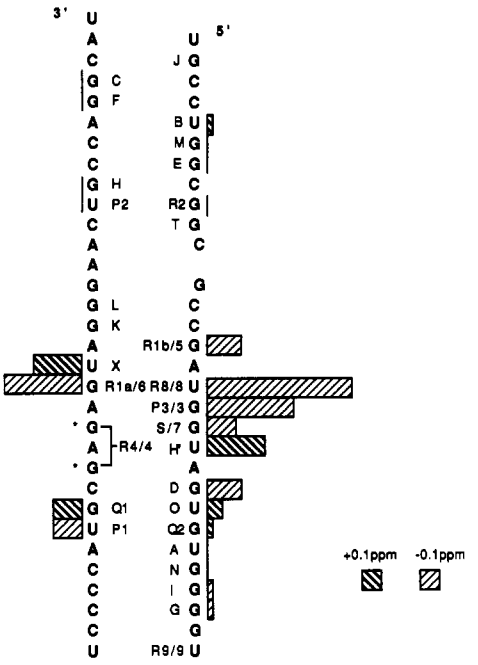


FIGURE 8: Chemical shift changes in the imino proton spectrum of fragment 1 induced by L25 binding in $Ca^{2+}$ at 277 K. The chemical shifts and assignments of the imino proton spectra of fragment 1 and the fragment 1–L25 complex were determined in $Ca^{2+}$ at 277 K. The figure shows the changes in chemical shift due to protein binding as chemical $shift_{complex}$ – chemical $shift_{free\ RNA}$.

## Discussion

Fundamental for arguments we have used to assign imino proton resonances in the helix V–loop E region is the assumption that they reflect secondary structure interactions, not tertiary interactions. Although we cannot prove this assumption is correct, the evidence we have strongly supports it. The characteristic resonances of this region are H', X, and S. All have been unambiguously identified in the spectra of intact 5S RNA (Jarema & Moore, 1986), fragment 1 (Kime, 1984), and pDG07 RNA (Gewirth & Moore, 1987). If tertiary interactions are involved in helix V–loop E, they cannot involve the helix II–helix III arm, which is present in 5S RNA, but absent from the other two species. The observation that pDG07 RNA's spectrum is normal in this regard makes it unlikely that helix I bases interact with loop E because helix I is joined to the base of helix V so as to make the two a single, coaxial helix in pDG07 RNA. Given additionally the existence of interstrand NOEs within the helix V–loop E region, tertiary interactions appear most unlikely.

The assignments we have made in the helix V–loop E region permit us to comment on the model for that region that has been proposed by Romby et al. (1988). Their model is based on the chemical reactivity of bases in spinach chloroplast 5S RNA and envisions a structure for helix V–loop E based on secondary interactions only. A most striking feature of the Romby model is its suggestion that loop E contains nothing but purine-purine base pairs and that its three pyrimidines are turned away from the helix axis. When transferred to *E. coli*, the model implies that the imino protons of U74, U77, and U104 all face the solvent (Figure 9). Were this the case, they should not contribute resonances to the downfield spectrum of 5S RNA at neutral pH and room temperature, due to rapid

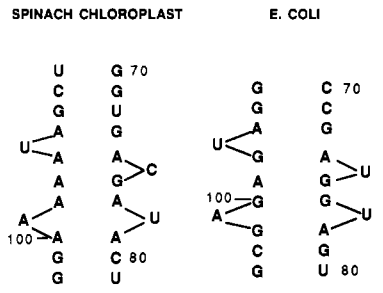


FIGURE 9: Schematic of the Romby model for helix V and loop E in chloroplast and *E. coli* 5S RNAs. The secondary structure proposed by Romby et al. (1988) for spinach chloroplast 5S RNA in the loop E region is shown on the left. All the bases in line vertically are proposed to pair with the bases opposite them. All the pyrimidines bulge outward and are not base paired. When the Romby model is transferred to the sequence of *E. coli* 5S RNA, the result is a structure like the one shown on the right.

exchange. The resonances of the imino protons of U77 and U104 are nevertheless detectable under these conditions, and U74's resonance can be seen in $Ca^{2+}$ at low temperature.

Furthermore, two loop E NOEs are inconsistent with the Romby model: H′ to S/7 and R1a/6 to X. The Romby model points G76 and G103 toward the helix axis while it directs U77 and U104 outward, as already noted. Because the distance between imino protons of G76 and U77 greatly exceeds 5 Å, the limit for measurable NOEs, an S/7 to H′ NOE would be impossible if the *E. coli* molecule had the Romby structure. The model is incompatible with the R1a/6 to X NOE for the same reason. Thus, it is clear that the Romby model does not describe 5S RNA from *E. coli*.

Does the inapplicability of the Romby model to *E. coli* 5S RNA indicate that it is incorrect for 5S RNAs in general, or only that is is incorrect for *E. coli*? We would be uncertain about the answer to this question had J. Egebjerg and R. Garrett of the University of Aarhus not made the results of their chemical modification studies of *E. coli* 5S RNA available to us (Egebjerg, 1987; Egebjerg and Garrett, personal communication).

The data of Egebjerg and Garrett indicate that the imino protons of U's in loops, e.g., U87, are solvent accessible, i.e. reactive, as expected, but do not place U74, U77, and U104 in that category, consistent with the NMR results. Indeed the pattern of nucleotide reactivities in *E. coli* observed by Garrett and co-workers is quite different from what Romby and his colleagues observe in spinach chloroplast 5S RNA. It is certain that the structure of loop E is not the same in the two RNAs, and it remains plausible that the Romby model is correct for spinach chloroplast 5S RNA.

The Romby model differs from the standard model for 5S RNA (Figure 1) in the loop E–helix V region, and the structure of *E. coli* 5S RNA does not conform to the Romby model. Does *E. coli* 5S RNA have the standard structure? The chemical shifts of the resonances in loop E and the pattern of NOE connectivities observed make it certain that the answer to this question is negative.

In the standard model, U77 (H′) and U103 (X) make Watson–Crick base pairs with A99 and A73, respectively. The imino protons of both should have chemical shifts characteristic of AU base pairs and give strong (>15%) NOEs to the H2 protons of the A's to which they are hydrogen bonded. The fact is that the chemical shifts of H′ and X are about 1 ppm upfield of normal AU imino proton resonances. Both give weak (<5%) NOEs to aromatic resonances in $Mg^{2+}$ which become stronger (~10%) in $Ca^{2+}$ (Kime and Moore, unpublished observations; Leontis, 1986). The identities of the protons responsible for these aromatic NOEs are unknown, but it is clear that the proton related to X is not an AH2. Its chemical shift is 7.69 ppm in $Mg^{2+}$ and 7.72 ppm in $Ca^{2+}$, and the chemical shifts of the AH2's in strand III are 8.12 and 6.95 ppm (Rycyna and Moore, unpublished results). Thus, neither X nor H′ are the UN3 resonances of normal AU base pairs. (It is interesting that Egebjerg and Garrett find the N1 of A73 to be unusually reactive, which it would not be were it paired with U104 in the usual way.)

There is a second imino aromatic NOE in the loop E region, which is conspicuous in $Ca^{2+}$, but which can sometimes be seen in $Mg^{2+}$ as well, the one between S (G76) and a resonance at 7.98 ppm (Gewirth et al., 1987). Under some conditions, this NOE is fully as strong as the UN1–AH2 NOE of an AU base pair, suggesting that the distance between the N1 proton of G76 and the aromatic proton in question must be less than 3.0 Å under some conditions. Deuterium labeling experiments have shown that the 7.98 ppm resonance is unlikely to come from the strand III side of loop E (Rycyna and Moore, unpublished results) and suggest that it may represent an AH2 proton (Gewirth et al., 1987). The bases most likely to be involved would appear to be A99 and A101.

It is hard to imagine how the aromatic NOE of G76 could appear unless the edge of G76 were in close contact with the edge of another base. Could it be that G76 forms an AG base pair with A99 or A101? The hydrogen bonds would have to run from G06 to AN6, and from GN1 to AN1 in order to account for an imino NOE to AH2. A(anti)-G(anti) base pairs of this kind have been examined in DNAs (Patel et al., 1984). The GN1's in DNA A-G base pairs have chemical shifts of about 13 ppm, a normal value for a base-paired GN1 proton. The chemical shift of S, however, is 10 ppm, which strongly suggests that it is not base paired at all. Furthermore, the chemical shift of S is insensitive to the changes in conditions that control the strength of its NOE to the aromatic at 7.98 ppm, which again is hard to rationalize. Clearly, if G76 is paired with an A, the geometry of the pair must be unusual.

The result obtained with the C74 mutant confirms that the heavily base-paired model for loop E shown in Figure 1 is incorrect for *E. coli* 5S RNA in solution. This mutation replaces a GU in the standard model with a GC, which should have stabilized it, since GC's are more strongly hydrogen bonded than GU's. Yet, the mutation obviously destabilized the normal structure, and, surprisingly, the molecule did not adopt a structure like the one in Figure 1; no new imino proton resonances replaced the ones that were lost. A more radical mutant, one that replaces A101 with a U, has a similar effect on the molecule (Gewirth, 1988).

Since 5S RNAs in *E. coli* and chloroplasts have similar sequences and function interchangeably in reconstituted ribosomes (Vogel et al., 1984), one would think that they should have closely similar secondary structures. Furthermore, one might have thought that both would conform to the 5S secondary structure deduced from sequence comparisons. Neither expectation is fulfilled. A resolution to these conflicts can be obtained if one accepts the idea that the physiologically relevant secondary structure for loop E appears only when 5S RNA becomes incorporated into the ribosome.

This hypothesis is not entirely comfortable, however. It is likely that L25 interacts with the same region of 5S RNA in the ribosome that it does in solution (Herold & Nierhaus, 1987; Stoeffler & Stoeffler, 1984). Since L25 binding does not appear to alter the structure of loop E dramatically, one could plausibly suggest that the loop E structure in *E. coli* 5S

RNA in vitro is close to the one that exists in the ribosome, even though it is not the standard one.

Clearly, sequence comparisons are a powerful means for determining secondary structures, but apparently there are limits. In the loop E, the problem is the determination of the organization of a region that has so many noncanonical base juxtapositions that the question of whether it has a secondary structure at all was unclear for a long time. In this case, sequence comparison seems to be failing.

The chemical shifts and thermal sensitivities of resonances like R1b/5, R1a/6, and R8/8 are not much different from those of the imino protons of nucleotides in solution (Young & Kallenbach, 1978; van den Hoogen et al., 1988; P. Zhang, unpublished data). It is likely that many of these bases are only slightly protected from solvent exchange, and may well not be hydrogen bonded at all. Given that the chemical shifts of resonances X, H′, and S/7 are also anomalously high, it is hard to escape the conclusion that there are very few, if any, conventional base pairs in loop E.

In closing, we note that the region of 5S RNA most perturbed by the binding of protein L25 starts with the GU's at the base of helix IV and runs through loop E to the bottom of helix V. Loop E, helix IV, and helix V all lie within the region identified as the L25 binding site by chemical and enzymatic protection experiments (Douthwaite et al., 1982; Huber & Wool, 1984; Kime & Moore, 1983c). In $Mg^{2+}$, all of helix IV is noticeably more affected by protein binding than helix I (Kime & Moore, 1983c), but in $Ca^{2+}$, the distal parts of helix IV are no more responsive than helix I. In thinking about the response of loop E to L25, it is important to bear in mind that all loop E resonances and some of its NOEs can be seen in the free RNA. L25 binding thus alters the structure of loop E in an evolutionary way, not a revolutionary one.

## Acknowledgments

We thank Drs. Jan Egebjerg and Roger Garrett for making their chemical protection data available to us prior to publication, and we are grateful to Drs. Kathy Triman and Harry Noller for their gift of 5S RNA mutants. Joyce Sherman and Alex Szewczak assisted in the production of the mutant RNAs described here, and Peter Demou of the Yale Chemical Instrumentation Center helped us obtain $^{15}N/^{1}H$ chemical shift correlation data.

## References

Brosius, J. (1984) *Gene 27*, 161–172.

Delihas, N., Anderson, J., & Singhal, R. P. (1984) *Prog. Nucleic Acid Res. Mol. Biol. 31*, 161–190.

Douthwaite, S., Christensen, A., & Garrett, R. R. (1982) *Biochemistry 21*, 2313.

Egebjerg, J. (1987) Ph.D. Thesis, Aarhus University.

Gewirth, D. T. (1988) Ph.D. Thesis, Yale University.

Gewirth, D. T., & Moore, P. B. (1987) *Biochemistry 26*, 5657–5665.

Gewirth, D. T., Abo, S. R., Leontis, N. B., & Moore, P. B. (1987) *Biochemistry 26*, 5213–5220.

Gonnella, N. C., Birdseye, T. R., Nee, M., & Roberts, J. D. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4834–4837.

Griffey, R. H., Redfield, A. G., Loomis, R. E., & Dahlquist, F. W. (1985) *Biochemistry 24*, 817–822.

Herold, M., & Nierhaus, K. H. (1987) *J. Biol. Chem. 262*, 8826–8833.

Huber, P. W., & Wool, I. G. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 322.

Jarema, M., & Moore, P. B. (1986) in *3D Structure and Dynamics of RNA* (van Knippenberg, P. H., & Hilbers, C. W., Eds.) pp 175–189, Plenum Press, New York.

Kime, M. J. (1984) *FEBS Lett. 173*, 342–346.

Kime, M. J., & Moore, P. B. (1982) *Nucleic Acids Res. 10*, 4973–4983.

Kime, M. J., & Moore, P. B. (1983a) *FEBS Lett. 153*, 199–203.

Kime, M. J., & Moore, P. B. (1983b) *Biochemistry 22*, 2615–2622.

Kime, M. J., & Moore, P. B. (1983c) *Biochemistry 22*, 2622–2629.

Kime, M. J., Ratcliffe, R. G., Moore, P. B., & Williams, R. J. P. (1981) *Eur. J. Biochem. 116*, 269–276.

Kjems, J., Olesen, S. O., & Garrett, R. A. (1985) *Biochemistry 24*, 241–250.

Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488–492.

Leontis, N. B. (1986) Ph.D. Thesis, Yale University.

Leontis, N. B., & Moore, P. B. (1986) *Biochemistry 25*, 5736–5744.

Leontis, N. B., Ghosh, P., & Moore, P. B. (1986) *Biochemistry 25*, 7389–7392.

Moore, P. B., Abo, S., Freeborn, B., Gewrith, D. T., Leontis, N. B., & Sun, G. (1988) *Methods Enzymol. 164*, 158–174.

Patel, P. J., Kozlowski, S. A., Ikuta, S., & Itakura, K. (1984) *Biochemistry 23*, 3207–3217.

Patel, D. J., Shapiro, L., & Hare, D. (1987) *Q. Rev. Biophys. 20*, 36–113.

Romby, P., Westhof, E., Toukifimpa, R., Mache, R., Ebel, J.-P., Ehresmann, C., & Ehresmann, B. (1988) *Biochemistry 27*, 4721–4730.

Stahl, D. A., Leuhrsen, K. R., Woese, C. R., & Pace, N. R. (1981) *Nucleic Acids Res. 9*, 6129–6137.

Stoeffler, G., & Stoeffler-Meilicke, M. (1984) *Annu. Rev. Biophys. Bioeng. 13*, 303–330.

van den Hoogen, Y. Th., van Beuzekom, A. A., de Vroom, E., van der Marel, G. A., van Boom, J., & Altona, C. (1988) *Nucleic Acids Res. 16*, 5013–5030.

Vogel, D. W., Hartmann, R. K., Bartsch, M., Subramanian, A. R., Kleinow, W., O'Brien, T. W., Pieler, T., & Erdmann, V. A. (1984) *FEBS Lett. 169*, 67–72.

Young, P. R., & Kallenbach, N. R. (1978) *J. Mol. Biol. 126*, 467–479.