Structural Characterization of a Six-Nucleotide RNA Hairpin Loop Found in *Escherichia coli*, r(UUAAGU)[†,‡]

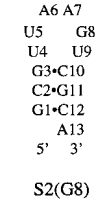
Hong Zhang,[§] Matthew A. Fountain,[||] and Thomas R. Krugh*[,§]

Department of Chemistry, University of Rochester, Rochester, New York 14627-0216, and Departments of Chemistry and Biology, State University of New York College at Fredonia, Fredonia, New York 14063

Received June 14, 2001

ABSTRACT: The binding region of the *Escherichia coli* S2 ribosomal protein contains a conserved UUAAGU hairpin loop. The structure of the hairpin formed by the oligomer r(GCGU4U5A6A7G8U9CGCA), which has an r(UUAAGU) hairpin loop, was determined by NMR and molecular modeling techniques as part of a study aimed at characterizing the structure and thermodynamics of RNA hairpin loops. Thermodynamic data obtained from melting curves for this RNA oligomer show that it forms a hairpin in solution with the following parameters: $\Delta H° = -42.8 \pm 2.2$ kcal/mol, $\Delta S° = -127.6 \pm 6.5$ eu, and $\Delta G°_{37} = -3.3 \pm 0.2$ kcal/mol. Two-dimensional NOESY WATERGATE spectra show an NOE between U imino protons, which suggests that U4 and U9 form a hydrogen bonded U·U pair. The U5(H2′) proton shows NOEs to both the A6(H8) proton and the A7(H8) proton, which is consistent with formation of a "U" turn between nucleotides U5 and A6. An NOE between the A7(H2) proton and the U9(H4′) proton shows the proximity of the A7 base to the U9 sugar, which is consistent with the structure determined for the six-nucleotide loop. In addition to having a hydrogen-bonded U·U pair as the first mismatch and a U turn, the r(UUAAGU) loop has the G8 base protruding into the solvent. The solution structure of the r(UUAAGU) loop is essentially identical to the structure of an identical loop found in the crystal structure of the 30S ribosomal subunit where the guanine in the loop is involved in tertiary interactions with RNA bases from adjacent regions [Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000) *Nature 407*, 327−339]. The similarity of the solution and solid-state structures of this hairpin loop suggests that formation of this hairpin may facilitate folding of 16S RNA.

RNAs are intimately associated with a wide variety of biological functions (*1*−*3*), and understanding the molecular basis of these functions requires knowledge of the three-dimensional structure of RNA. Exciting progress in both X-ray and NMR structural studies of RNA is continuing (*4*−*17* and references therein). The hairpin loop is one of the most commonly observed RNA structural motifs (*18*−*20*). We report the use of NMR spectroscopy and molecular modeling to derive the solution structure of an RNA hairpin r(GCGUUAAGUCGCA) (Figure 1) which contains the six-nucleotide loop r(UUAAGU) that is highly conserved in 16S rRNA and has been shown to be important for the binding of the ribosomal protein S2 in *Escherichia coli* (*21*). This hairpin has G as the eighth nucleotide, and we call this the S2(G8)[1] hairpin (Figure 1). The G·C closing base pair and the six-nucleotide loop region in the S2(G8) hairpin are identical to a hairpin loop region in 16S RNA from *E. coli* (shown within the box in Figure 2). The crystal structure of the 30S ribosomal subunit recently reported (*16*) contains an r(UUAAGU) loop, and as shown below, we find excellent agreement between the solution structure of the hairpin loop and the corresponding loop in the crystal structure.



S2(G8)

FIGURE 1: Sequence of the S2(G8) RNA hairpin, r(GCGUUAAGUCGCA), studied by NMR and molecular modeling.

In an earlier study, we (*22*), as well as the Draper lab (*23*), reported the structure of another six-nucleotide hairpin loop, r(GUAAUA), in which G is opposite A as the first mismatch. A crystal structure of a protein−RNA complex containing this loop was recently reported (*24*). The predominant conformer of our solution structure (*22*) for the GUAAUA hairpin loop has essentially the same structure as the corresponding hairpin loop observed in the protein−RNA complex (*24*).

In addition to direct structural studies, methods have been developed to predict the folding of RNAs (*25*−*30*). Current secondary structure prediction models evaluate the stability of an RNA hairpin by considering the number of bases in the loop (*25*, *31*), the stacking of the first mismatch in the loop (*32*), and the nature of the closing base pair (*33*−*35*). On the basis of these studies, the r(UUAAGU) loop is

[†] This publication was supported by NIH Grant GM53826 (T.R.K.).
[‡] Protein Data Bank entry 1HS2.
* To whom correspondence should be addressed. Phone: (716) 275-4224. Fax: (716) 473-6889. E-mail: krugh@chem.rochester.edu.
[§] University of Rochester.
[||] State University of New York College at Fredonia.
[1] Abbreviations: S2(G8), r(GCGUUAAGUCGCA); EDTA, ethylenediaminetetraacetic acid; T_m, melting temperature in degrees Celsius.

10.1021/bi011226x CCC: $20.00
Published on Web 07/26/2001

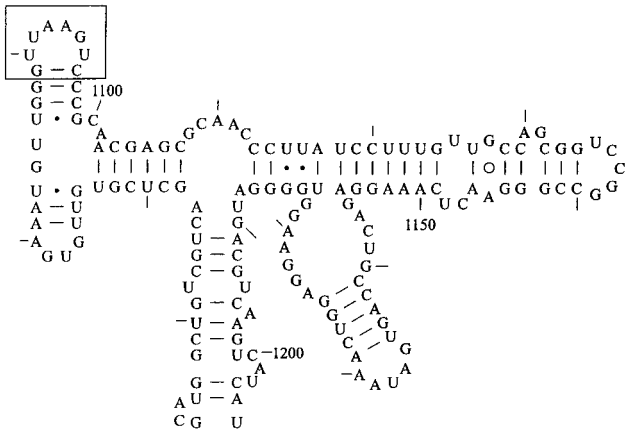


FIGURE 2: Schematic drawing of the secondary structure of the ribosomal protein S2 binding region of *E. coli*.

expected to be a stable loop, which is verified by melting studies of the S2(G8) hairpin.

MATERIALS AND METHODS

RNA Preparation and Purification. The RNA sequence was synthesized using the phosphoramadite method (*36*) with an Applied Biosystems model 392 DNA/RNA synthesizer. Protecting groups were removed using the procedure described by Glen Research (*37*). The RNA oligonucleotide was cleaved from the solid support, and base labile protecting groups were removed using a 1:3 ethanolic/ammonium hydroxide solution. The silyl protecting groups were removed using tetrabutylammonium fluoride at 55 °C. The oligomer was desalted using a Sephadex PD-10 column and then purified by TLC using Si500F TLC plates developed in a 55:35:10 1-propanol/ammonium hydroxide/water solution. Oligonucleotides were further purified using preparative HPLC (Hypersil ODS reverse-phase column, 250 mm × 10 mm) with 0.01 M phosphate buffer as the aqueous phase and methanol as the organic phase. A solvent gradient of 0 to 20% organic phase over the course of 20 min was used for the final purification step, and to assess purity. The oligomer was then dialyzed overnight to remove salt. For NMR studies in D_2O, the oligonucleotide was dissolved in 0.1 M NaCl, 0.01 M sodium phosphate, and 0.5 mM EDTA (pH 7.0). The sample was lyophilized to dryness twice with 99.9% D_2O and then once with 99.996% D_2O. The concentration of the RNA oligonucleotide in D_2O was 3 mM. The concentration of the RNA oligonucleotide for NMR experiments in water was 1 mM with a 0.1 M NaCl, 0.01 M sodium phosphate, and 0.5 mM EDTA (pH 5.0) buffer and a 9:1 H_2O:D_2O ratio.

Thermodynamics Measurements. Thermodynamic parameters were measured using the 0.1 M NaCl, 0.01 M sodium phosphate, and 0.5 mM EDTA (pH 7.0) buffer. Oligomer single-strand concentrations, C_T, were calculated from high-temperature absorbances using single-strand extinction coefficients. Oligomer concentrations were varied over a 20-fold range. Absorbance versus temperature melting curves were measured at 280 nm from 95 to 15 °C with a cooling rate of 1 °C/min on a Beckman DU 640 spectrometer. Thermodynamics parameters for hairpin formation were determined using the Meltwin 3.0 program (*38*) in which the melting profiles were fit to a two-state model (*39*).

NMR Spectroscopy. NMR spectra were collected on a Varian INOVA 500 MHz spectrometer. NMR data were processed with VNMR (Varian) and Felix (Molecular Simulations Incorporated). One-dimensional exchangeable proton spectra were recorded using a binomial 1331 solvent suppression pulse sequence (*40*). Two-dimensional spectra were recorded in the phase sensitive mode using the States−Haberkorn method (*41*). Spectra of nonexchangeable protons were recorded at 30 °C and of exchangeable protons at 0 °C. The magnitude of the residual HDO peak was reduced by low-power presaturation. Phosphorus spectra were referenced to the internal phosphate buffer at pH 7.0. Spin−lattice relaxation rates were determined in D_2O using the inverse-recovery method.

NOESY spectra of nonexchangeable protons were recorded at 60, 100, 120, 150, and 600 ms at 30 °C. Spectra were recorded with 256 t_1 increments, each having 2K data points with 32 scans per fid and a recycle delay time of 3 s. The data were zero filled to 2K data points in the t_1 dimension to yield a 2K × 2K matrix. The data were apodized with a shifted sine-bell squared function in both directions. A NOESY WATERGATE pulse sequence (*42*) with a 175 ms mixing time and spectral width of 10 000 Hz was used to record spectra in 90% H_2O solutions.

DQF-COSY spectra were recorded with spectral widths of 2000 and 5000 Hz at 30 °C with 512 t_1 increments, each having 4K data points with 32 or 64 scans per fid and a recycle delay time of 2.5 s. The data were apodized with a phase-shifted sine-bell squared function in both directions.

TOCSY spectra were recorded with 256 t_1 increments, each having 2K data points with 64 transients per fid with a recycle delay time of 3 s and a spin lock mixing time of 80 ms. The data were zero filled to 2K data points in the t_1 dimension to yield a 2K × 2K matrix. The data were apodized with a phase-shifted sine-bell squared function in both directions.

$^1H-^{31}P$ HETCOR spectra were acquired using the pulse sequence described by Sklenar et al. (*43*) with spectra widths of 3000 Hz for ^1H and 1000 Hz for ^{31}P, a 2.5 s recycle delay time, and a total of 256 t_1 increments with 2K points. The data were apodized with a phase-shifted sine-bell squared function in both directions.

Restraints. NOE-based estimates of distances between nonexchangeable protons were derived from NOESY cross-peak volumes at 100, 120, and 150 ms using an isolated spin pair approximation. Cross-peak volumes were integrated using Felix, and the buildup rates were compared to H5−H6 (2.45 Å) cross-peaks and scaled by $1/r^6$. Distance restraints were added to maintain Watson−Crick base pairs in the hairpin stem.

Structure Calculations. Structure calculations, visualization, and analysis were performed using Insight II and Discover (Molecular Simulations Inc.) running on an SGI O2 or SGI Octane workstation. The AMBER force field (*44*, *45*) was used for all calculations with a 16 Å cutoff for nonbonding interactions. Force constants for both distance and dihedral restraints were set to 50 kcal/Å^2 and 50 kcal mol^{-1} rad^{-2}, respectively. The RNA molecule was placed in a periodic boundary conditions box (40 Å × 30 Å × 42 Å), and the remaining space was filled with water molecules by using a soak command. The first stage of the calculations involved restrained energy minimization using steepest

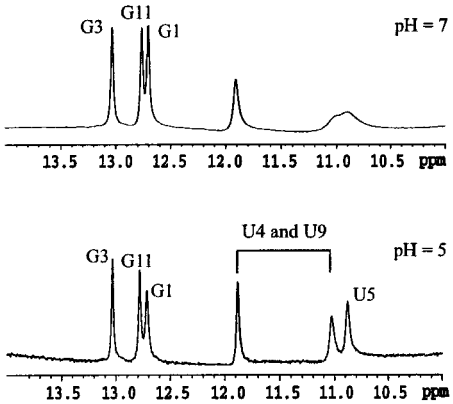


FIGURE 3: Imino proton region of the ^{1}H NMR spectrum of the hairpin r(GCGUUAAGUCGCA) at 0 °C in a 90% H_2O/10% D_2O mixture recorded with a 1331 binomial pulse sequence. The top spectrum was recorded at pH 7, and the bottom spectrum was recorded at pH 5.

descents and conjugate gradient methods. The restrained energy minimization was then followed by restrained molecular dynamics. The molecular dynamics simulations began with heating the RNA for 10 ps at both 300 and 400 K and then for 3 ps at 500 K. This was followed by 3 ps at 400 and 300 K. The new structure was then subjected to restrained energy minimization using steepest descents and conjugate gradient methods.

RESULTS

Thermodynamics of Hairpin Formation. The UV absorbance as a function of temperature for the S2(G8) hairpin exhibits a characteristic melting profile (not shown). The melting temperature of the oligomer (62.7 °C) is independent of oligomer concentration (data not shown) which indicates formation of a hairpin. The melting curve shows that the S2(G8) hairpin is stable at temperatures used for NMR structural studies (≤30 °C). The thermodynamic parameters for hairpin formation are as follows: $\Delta H^\circ = -42.8 \pm 2.2$ kcal/mol, $\Delta S^\circ = -127.6 \pm 6.5$ eu, $\Delta G^\circ_{37} = -3.3 \pm 0.2$ kcal/mol.

Exchangeable Proton Assignments. The imino proton regions of the one-dimensional ^{1}H NMR spectra of the S2(G8) hairpin at pH 7 and 5 are shown in Figure 3. Although the S2(G8) hairpin (Figure 1) has seven imino protons, only six imino proton resonances are observed at either pH 5 or 7. The spectra were recorded at several pH values, yet no evidence was obtained for a seventh resonance in the 9−17 ppm range. We assign the three upfield resonances observed in Figure 3 to the U4, U5, and U9 imino protons in the S2(G8) hairpin. The resonances at 11.0 and 11.9 ppm are assigned to the U4 and U9 imino protons, respectively, in part because they exhibit an NOE to one another in a two-dimensional NOESY WATERGATE spectrum at pH 5 (peaks b and c in Figure S1 of the Supporting Information). Although these resonances (11.0 and 11.9 ppm) could not be specifically assigned to the U9 and U4 imino protons, respectively, these chemical shifts and the observation of an NOE between these resonances are consistent with formation of a U·U mismatch base pair (*46*). NOE data (presented below) for both exchangeable and nonexchangeable protons show that the U4 and U9 bases are stacked on the G3·C10 base pair, which is consistent with formation of a U4·U9 base pair. A distance restraint with a lower bound of 1.5 Å and an upper bound of 4.0 Å between the U4(NH3) and U9(NH3) imino protons was used in structural calculations.

The resonance at 11.9 ppm also exhibits an NOE to the G3 imino proton resonance (cross-peak a in Figure S1), showing that this imino proton is stacked on the closing G·C Watson−Crick base pair. The remaining resonance at 10.9 ppm does not exhibit any strong NOEs in the WATERGATE spectrum (other than an exchange cross-peak to the solvent) and is tentatively assigned to the U5(NH3) imino proton. We are unable to locate a resonance for the G8(NH1) imino proton.

The imino proton resonances for the three G·C base pairs in the S2(G8) hairpin stem were assigned by standard procedures (*47*−*49*). These resonances exhibit cross-peaks to C amino protons in the two-dimensional NOESY WATERGATE spectrum in H_2O (Figure S1). The C amino protons were assigned on the basis of cross-peaks observed to their C(H5) protons. The assignment of the resonance at 12.7 ppm to the G1(NH1) proton was verified by observation of temperature-dependent behavior characteristic of an imino proton at the end of a helix (*50*). The G11(NH1) resonance at 12.8 ppm exhibits cross-peaks to both G3(NH1) at 13.1 ppm and G1(NH1) at 12.7 ppm. The NOEs observed in the WATERGATE spectrum verified the presence of Watson−Crick hydrogen bonding for the G·C pairs in the stem of the hairpin, and thus, distance restraints were incorporated to maintain G·C pairing.

Nonexchangeable Protons. One-dimensional proton spectra at temperatures ranging from 10 to 40 °C are shown in Figure S2. The U9(H6) proton moved upfield and the C10(H6) proton downfield as the temperature was lowered from 40 to 10 °C. These two resonances exhibit a 0.1 ppm change in chemical shift over this temperature range. These relatively modest changes in chemical shift suggest that there are no significant structural changes in this temperature range.

The base to H1′ and H5 region of a 600 ms NOESY spectrum is shown in Figure 4. Assignments of nonexchangeable protons were made using standard procedures (*47*, *49*, *51*) and are reported in Table S1. Because of spectral overlap, some chemical shift assignments were made at 20 °C.

In the 100 ms NOESY spectrum collected at 30 °C, all of the base to H1′ cross-peaks were weaker than the cytosine and uridine H5−H6 cross-peaks, indicating that all glycosidic torsion angles are anti. The G8(H1′) proton exhibits a doublet pattern with a $J_{H1'-H2'}$ of 7.5 Hz, as shown in Figure 5. This 7.5 Hz coupling shows that the G8 sugar is predominantly, although not exclusively, in the C2′-endo (S type) conformation. Thus, a dihedral restraint with a lower bound of 130° and an upper bound of 170° was used for the G8(H1′−C1′−C2′−H2′) dihedral angle.

The terminal A13 sugar exhibits a resolvable doublet with a $J_{H1'-H2'}$ of 2 Hz (Figure 5), as anticipated for a 3′-terminal sugar. The value of $J_{H1'-H2'}$ for U9 was not determined due to overlap with other sugar resonances. All other sugar H1′ resonances in the S2(G8) hairpin have $J_{H1'-H2'}$ values of <2 Hz and thus have C3′-endo (N type) conformations.

The H2 protons of A6, A7, and A13 were distinguished from other base protons by their characteristic long T_1 relaxation times (data not shown). The A13(H2) proton exhibits a cross-strand NOE to the G1(H1′) proton (cross-peak b in Figure 4), consistent with having the A13 base

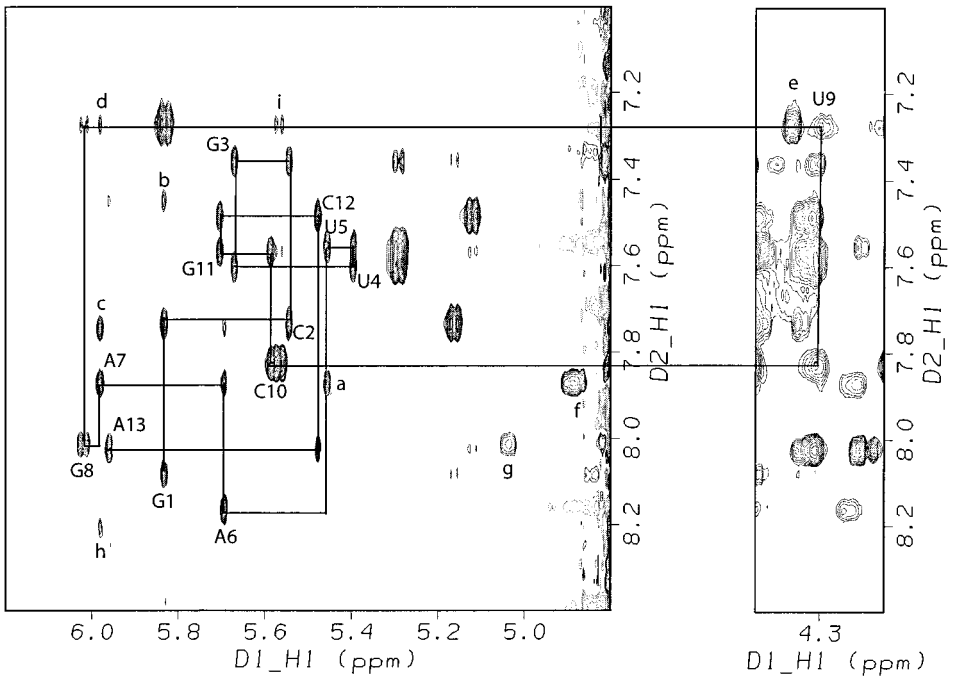


FIGURE 4: Base to H1′ and H5 region of a 600 ms NOESY spectrum of the S2(G8) hairpin, r(GCGUUAAGUCGCA), recorded at 30 °C in 0.1 M NaCl, 0.01 M sodium phosphate, and 0.5 mM EDTA (pH 7.0). The base to H1′ NOESY walk is traced with solid lines. The base labels represent intranucleotide base to H1′ NOE connectivities. The labeled cross-peaks are as follows: (a) U5(H1′) to A7(H8), (b) G1-(H1′) to A13(H2), (c) A6(H2) to A7(H1′), (d) A7(H1′) to U9(H6), (e) G8(H4′) to U9(H6), (f) A7(H3′) to A7(H8), (g) G8(H2′) to G8(H8), (h) A7(H2) to A7(H1′), and (i) C10(H5) to U9(H6).

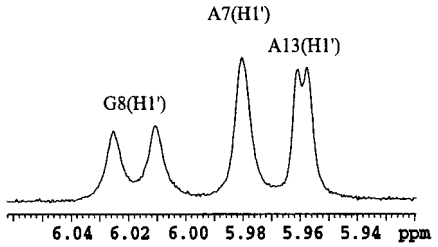


FIGURE 5: Portion of the one-dimensional spectrum of the hairpin r(GCGUUAAGUCGCA) recorded at 30 °C in D_2O showing the H1′ resonances of G8, A7, and A13.

stacked on the G1·C12 base pair. The 3′-dangling A is present to increase the melting temperature of the RNA hairpin; it also sharpens the imino proton resonance of the terminal base pair (*52*).

The A6(H2) and A7(H2) protons were distinguished from one another by NOEs to their own H1′ sugar protons. The A6(H2) proton shows NOEs to its own H1′ and to the A7-(H1′) proton (cross-peak c in Figure 4) which positions the A6 base stacked on the A7 base. The presence of a cross-peak between the A6(H2) and A7(H2) protons (data not shown) confirms the assignment of the these resonances and is consistent with stacking of the A6 and A7 bases.

NOEs That Characterize the Structure. The U5(H1′) proton exhibits an NOE to the A7(H8) proton (cross-peak a in Figure 4), and the U5(H2′) proton shows a strong NOE to the A7(H8) proton (data not shown). These NOEs are consistent with the turning point of the hairpin located between U5 and A6.

The A7(H2) proton shows an unusual $N + 2$ NOE to the U9(H4′) proton (Figure S3), suggesting that the G8 base and sugar protrude into the solvent, allowing the A7 base to stack on the U9 sugar. A weak NOE between the A7(H1′) and U9(H6) protons (cross-peak d in Figure 4), as well as A7-(H2′) and U9(H6), A7(H3′) and U9(H6), and A7(H4′) and U9(H6) NOEs (cross-peaks k, g, and j in Figure S4), provides supporting evidence for the presence of a bulged G8 nucleotide. Although these NOEs may arise from spin diffusion, and were not used to calculate distance restraints, their presence is consistent with having G8 protruding into the solvent. The U9(H6) proton shows an NOE to the C10-(H6) proton (cross-peak a in Figure S4), but does not show an NOE to the G8(H8) proton. The G8(H8) proton does not exhibit NOEs to any A7 protons, or to any U9 protons. The G8(H2′) proton shows a weak NOE to the U9(H6) proton (cross-peak f in Figure S4), while the G8(H4′) proton exhibits a relatively strong NOE to the U9(H6) proton (cross-peak e in Figure 4 and cross-peak f in Figure S4). These NOEs are consistent with the conformation of the loop for the A7-G8-U9 region that is shown below.

The U4(H8) proton exhibits NOEs (Figure 4) that are consistent with having the U4 base stacked with the G3 base, as opposed to having the U4 base protrude into the solvent. The U9(H1′) proton exhibits an NOE to the C10(H6) proton, and the C10(H5) proton exhibits an NOE to the U9(H6) proton (cross-peak i in Figure 4), which indicates stacking of the U9 base with the C10 base. Stacking of the U4 and U9 bases on the G3·C10 base pair, along with the NOE observed between two U(NH3) protons, provides strong evidence for the formation of a U4·U9 pair.

Structural Characterization of the r(GCGUUAAGUCG-CA) Hairpin. Thirty structures were calculated using the energy minimization and molecular dynamics protocol described in Materials and Methods. The distance and dihedral angle restraints used in all of the calculations are listed in Tables S2 and S3, along with observed distance and dihedral angles in the average structure. Figure 6 shows

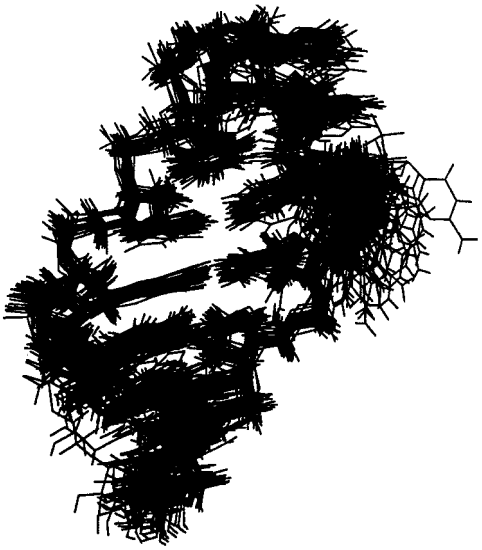


FIGURE 6: Superposition of 30 structures of the S2(G8) hairpin obtained from molecular dynamics simulations.

a superposition of the 30 structures. The coordinates of the 30 structures generated from molecular dynamics were averaged, and the resulting structure was energy-minimized to provide the average structure shown in Figure 7. The coordinates of the average structure are available from the Protein Data Bank as entry 1HS2.

The rmsd values for the 30 structures are given in Table 1. The extrahelical G8 nucleotide and the dangling A13 nucleotide are less well-defined than other nucleotides, as expected. The average structure satisfies all of the distance restraints (i.e., there are no distance violations larger than 0.1 Å), and all but one of the dihedral restraints are satisfied. The G8(H1′−C1′−C2′−H2′) dihedral angle of 100° in the average structure does not satisfy the 130−170° restraint used in the calculations. This restraint is based on the observation of the coupling constant $J_{H1'-H2'}$ of 7.5 Hz for the G8 sugar which we interpret as being consistent with a C2′-endo conformation predominating for the G8 sugar. Inspection of the G8 sugar in each of the 30 calculated structures showed that the G8 sugar adopted a C3′-endo conformation in 18 of the 30 structures, and a C2′-endo conformation in the remaining 12 structures. We conclude that both experimental data and MD simulations indicate a mixture of sugar conformations for G8, and note that the distribution of the C2′-endo and C3′-endo populations found in the MD simulations differs somewhat from that expected from the observation of a $J_{H1'-H2'}$ of 7.5 Hz for G8.

The salient features of the r(GCGUUAAGUCGCA) hairpin structure (Figure 7) include a "U" turn of the loop between U5 and A6, the external location of the G8 base, stacking of the A7 base on the U9 sugar, stacking of U5 on the U4·U9 mismatch, and formation of a U·U base pair as the first mismatch. Analysis of the 30 calculated structures shows that the U4 and U9 bases form a U·U pair involving two hydrogen bonds (Figures 8) in 28 of the 30 structures. In each of these 28 structures, the U4(NH3) proton forms a hydrogen bond to the U9(O6) oxygen, while the U9(NH3) proton forms the corresponding hydrogen bond to the U4-(O6) oxygen. One of the 30 structures has a U·U pair with only one hydrogen bond between U4 and U9, and one structure did not exhibit hydrogen bonding between U4 and U9.

In the average structure, the U5(NH3) imino proton forms a hydrogen bond to a phosphate oxygen atom of the A7-G8 linkage (i.e., the $N + 3$ G8 phosphate), as expected in a U turn (*20*). The A6 and A7 bases are stacked on one another, with one surface of the A6 base exposed to solvent.

DISCUSSION

Several previously observed RNA structural motifs characterize the r(GCGUUAAGUCGCA) hairpin loop: formation of a U·U mismatch, formation of a uridine turn, and a bulged loop nucleotide.

U·U Mismatch. The U·U base pair can adopt several different hydrogen bonding patterns (*53*). The cis wobble base pairing observed in the S2(G8) hairpin is similar to U·U base pairing in the tandem internal loop, 5′-r(CGCUUGCG)-3′ (*46*). This kind of pairing uses the Waston−Crick faces of both pyrimidines to form two hydrogen bonds. A novel trans Hoogsteen pairing U·U mismatch has been observed by Wahl et al. (*54*) in a crystal structure of the RNA fragment 5′-r(UUCGCG)-3′ which has a 5′-terminal UU overhang forming a Hoogsteen-like U·U self-base pair with the overhang of an adjacent duplex. The trans Hoogsteen

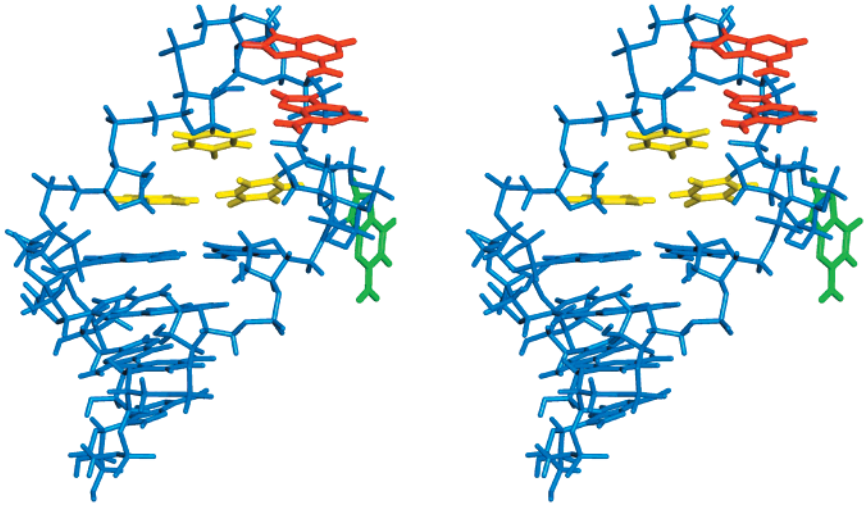


FIGURE 7: Stereo image of the average structure of the S2(G8) hairpin, r(GCGUUAAGUCGCA). The stem bases are blue, and the U4, U5, and U9 bases are yellow. The A6 and A7 bases are red, and the G8 base is green.

Table 1: Number of Restraints and Average rmsds (Å) for the r(GCGUUAAGUCGCA) Hairpin

nucleotide	no. of NOEs (intra/inter)	no. of dihedral restraints	rmsd[a]
G1	4/12	3	1.00
C2	6/16	5	0.74
G3	5/16	3	0.69
U4	4/9	0	0.68
U5	5/6	0	0.72
A6	5/5	0	1.03
A7	8/10	0	0.88
G8	5/2	1	2.16
U9	8/11	0	0.78
C10	5/11	1	0.66
G11	2/11	4	0.53
C12	5/13	4	0.66
A13	5/4	0	1.19
total[b]	67/126	15	0.90
total[c]	57/120	14	0.76

[a] The rmsd values are calculated from the 30 structures generated from molecular dynamics as described in Structure Calculations. [b] These values include all 13 nucleotides. [c] These values exclude G8 and A13.

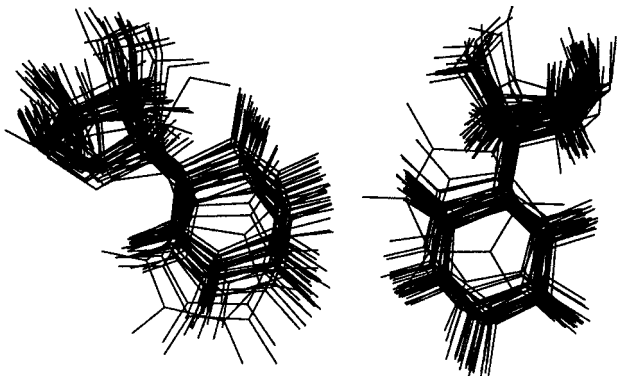


FIGURE 8: Superposition of the U·U mismatch in the 30 structures generated for the S2(G8) hairpin, r(GCGUUAAGUCGCA), with the U·U mismatch in the average structure.

base pairing uses the Waston−Crick face of one pyrimidine and the Hoogsteen face of another with a single hydrogen bond between N3H and O. There have been other reported U·U mismatch base pairs. For example, in a GAAA tetraloop receptor RNA (*55*), a U·U mismatch containing a single imino to carbonyl hydrogen bond is observed. The variability of the U·U mismatch raises the interesting possibility that other hairpin loops with U·U first mismatches may adopt hydrogen bonding patterns other than the cis wobble base pairing found in the S2(G8) hairpin (Figures 7 and 8).

Bulged G8 Nucleotide. The location of the G8 base is a prominent feature of the solution structure of the r(UUAAGU) loop. The lack of interaction of the G8 base with the other bases of the loop suggests that the nature of this base may not be important in determining the overall structure of the loop. We (H. Zhang, M J. Culyba, H. Volkman, and T. R. Krugh, in preparation) have studied the structures of related hairpins and found that hairpins containing A, C, and U bases at the eighth position have structures that are similar to the S2(G8) hairpin. Of course, the eighth base will be important if this base is involved in tertiary interactions when this loop is a part of a larger RNA structure.

Upfield Shifting of U9 Sugar Protons. As shown in Figure 7, the A7 base is stacked on the U9 sugar. The ring current anisotropy of the A7 base results in unusual upfield chemical shifts of the U9(H1′) resonance (4.34 ppm) and the U9(H4′) resonance (3.39 ppm) as shown in Table S1. These ring current effects are consistent with the structure of the S2(G8) hairpin (for examples of ring current shifts in other contexts, see refs *56*−*58*).

FIGURE 9: Superposition of the S2(G8) hairpin, r(GCGUUAAGUCGCA), which is shown in green, with a portion of the ribosomal subunit (*16*) which is shown in yellow. Nucleotides 1087−1098, with the sequence r(GGGUUAAGUCCC), from the ribosomal subunit are shown.

Uridine-Turn Motif. Uridine-turn motifs are common in RNA hairpin loops and have been observed in the anticodon and TψC loops in tRNA, in the hammerhead ribozyme, and in other RNA structures (*18*, *20*, *57*, *59*, *60*). The structure of the S2(G8) hairpin reported here (with the loop sequence U4-U5-A6-A7-G8-U9) has a sharp reversal in the direction of the backbone (the so-called U turn) between the U5 and A6 nucleotides, as did the r(GUAAUU) hairpin loop reported previously (*22*, *23*). In both the r(UUAAGU) loop and the r(GUAAUU) loop, the U5 base stacks on the first mismatch pair and the U5(NH3) imino proton is hydrogen bonded to a phosphate oxygen connecting nucleotides 7 and 8.

Comparison with the 30S Ribosomal Subunit Crystal Structure. The 30S ribosomal subunit contains a hairpin loop with the same six nucleotides as in S2(G8), along with the same closing base pair (Figures 1 and 2). The recently determined crystal structure published by Wimberly et al. (*16*) provides the opportunity to compare the solution structure of the S2(G8) hairpin with a hairpin of the same sequence found in the crystal structure of the 30S ribosomal subunit. Figure 9 shows the average structure of the S2(G8) hairpin (green), superimposed with the equivalent hairpin loop (bases 1090−1095) from the 30S ribosomal subunit (yellow). The similarity of the structures of the two hairpin loops is apparent. We noted above that in solution the G8 base protrudes into the solvent, which suggests that it would be available for tertiary interactions when the UUAAGU loop is a part of a larger RNA structure. In the 30S ribosomal subunit, the corresponding base (G1094) is involved in tertiary interactions, and has proximal neighbors that include nucleotides 1068−1070, 1084, and 1105−1107.

The similarity of the structures of the hairpin loop in solution and in the crystal suggests that formation of this hairpin may play a role in folding 16S RNA. However, it is known that the structures of some RNA motifs are different in solution than when incorporated as a domain in a larger RNA or RNA−protein complex (*61*, *62*). Thus, while it may be premature to conclude that the UUAAGU loop has special significance as an RNA motif, the identification of such motifs will enhance the ability to predict RNA structure based on sequence.

SUPPORTING INFORMATION AVAILABLE

Chemical shift assignments of the S2(G8) hairpin (Table S1), distance restraints used for the calculations (Table S2), dihedral restraints used for the calculations (Table S3), a plot of imino proton NOEs observed in a two-dimensional WATERGATE spectrum (Figure S1), the base region of the ^{1}H one-dimensional spectrum of S2(G8) at various temperatures (Figure S2), a portion of the two-dimensional NOESY spectrum showing the unusual cross-peak between nucleotides A7 and U9 (Figure S3), and a portion of a two-dimensional NOESY D_2O spectrum showing cross-peaks between the U9(H6) proton and other protons (Figure S4). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Cech, T. R. (1990) *Annu. Rev. Biochem. 59*, 543−568.
2. Altman, S. (1990) *J. Biol. Chem. 265*, 20053−20056.
3. Walter, P., and Johnson, A. E. (1994) *Annu. Rev. Cell Biol. 10*, 87−119.
4. Tinoco, I. (1996) *J. Phys. Chem. 100*, 13311−13322.
5. Doudna, J. A., and Cate, J. H. (1997) *Curr. Opin. Struct. Biol. 7*, 310−316.
6. Moore, P. B. (1998) *Annu. Rev. Biophys. Biomol. Struct. 27*, 35−58.
7. Pappalardo, L., Kerwood, D. J., Pelczer, I., and Borer, P. N. (1998) *J. Mol. Biol. 282*, 801−818.
8. Holbrook, S. R. (1998) *Mol. Model. Nucleic Acids 682*, 56−76.
9. Hoogstraten, C. G., Legault, P., and Pardi, A. (1998) *J. Mol. Biol. 284*, 337−350.
10. Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N., and Noller, H. F. (1999) *Science 285*, 2095−2104.
11. Zimmermann, G. R., Wick, C. L., Shields, T. P., Jenison, R. D., and Pardi, A. (2000) *RNA 6*, 659−667.
12. Butcher, S. E., Allain, F. H. T., and Feigon, J. (1999) *Nat. Struct. Biol. 6*, 212−216.
13. Hermann, T., and Patel, D. J. (1999) *J. Mol. Biol. 294*, 829−849.
14. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) *Science 289*, 905−920.
15. Mollova, E., and Pardi, A. (2000) *Curr. Opin. Struct. Biol. 10*, 298−302.
16. Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000) *Nature 407*, 327−339.
17. Agalarov, S. C., Prasad, G. S., Funke, P. M., Stout, C. D., and Williamson, J. R. (2000) *Science 288*, 107−112.
18. Moore, P. B. (1999) *Annu. Rev. Biochem. 68*, 287−300.
19. Turner, D. H. (2000) in *Nucleic Acids: Structures, Properties, and Functions* (Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Eds.) pp 259−334, University Science Books, Sausalito, CA.
20. Gutell, R. R., Cannone, J. J., Konings, D., and Gautheret, D. (2000) *J. Mol. Biol. 300*, 791−803.
21. Mueller, F., and Brimacombe, R. (1997) *J. Mol. Biol. 271*, 545−565.
22. Fountain, M. A., Serra, M. J., Krugh, T. R., and Turner, D. H. (1996) *Biochemistry 35*, 6539−6548.
23. Huang, S. G., Wang, Y. X., and Draper, D. E. (1996) *J. Mol. Biol. 258*, 308−321.
24. Conn, G. L., Draper, D. E., Lattman, E. E., and Gittis, A. G. (1999) *Science 284*, 1171−1174.
25. Turner, D. H., Sugimoto, N., and Freier, S. M. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 167−192.
26. Zuker, M. (1989) *Science 244*, 48−52.
27. Michel, F., and Westhof, E. (1990) *J. Mol. Biol. 216*, 585−610.
28. Gautheret, D., and Cedergren, R. (1993) *FASEB J. 7*, 97−105.
29. Mathews, D. H., Andre, T. C., Kim, J., Turner, D. H., and Zuker, M. (1998) *Mol. Model. Nucleic Acids 682*, 246−257.
30. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) *J. Mol. Biol. 288*, 911−940.
31. Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373−9377.
32. Jaeger, J. A., Turner, D. H., and Zuker, M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7706−7710.
33. Serra, M. J., Lyttle, M. H., Axenson, T. J., Schadt, C. A., and Turner, D. H. (1993) *Nucleic Acids Res. 21*, 3845−3849.
34. Serra, M. J., Barnes, T. W., Betschart, K., Gutierrez, M. J., Sprouse, K. J., Riley, C. K., Stewart, L., and Temel, R. E. (1997) *Biochemistry 36*, 4844−4851.
35. Dale, T., Smith, R., and Serra, M. J. (2000) *RNA 6*, 608−615.
36. Usman, N., Ogilvie, K. K., Jiang, M. Y., and Cedergren, R. J. (1987) *J. Am. Chem. Soc. 109*, 7845−7854.
37. GlenResearch. http://glenres.com/ProductFiles/Technical/f_T-B_RNA.HTML.
38. McDowell, J. (1996) *Biochemistry 35*, 14077−14089.
39. Petersheim, M., and Turner, D. H. (1983) *Biochemistry 22*, 264−268.
40. Hore, P. J. (1983) *J. Magn. Reson. 54*, 539−542.
41. States, P. J., Haberkorn, R. A., and Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286−292.
42. Lippens, G., Dhalluin, C., and Wieruszeski, J. M. (1995) *J. Biomol. NMR 5*, 327−331.
43. Sklenar, V., Miyashiro, H., Zon, G., Miles, H. T., and Bax, A. (1986) *FEBS Lett. 208*, 94−98.
44. Weiner, S. J., Kollman, P. A., Nguyen, D. T., and Case, D. A. (1986) *J. Comput. Chem. 7*, 230−252.
45. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) *J. Am. Chem. Soc. 117*, 5179−5197.
46. SantaLucia, J., Kierzek, R., and Turner, D. H. (1991) *Biochemistry 30*, 8242−8251.
47. Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, New York.
48. Varani, G., Cheong, C. J., and Tinoco, I. (1991) *Biochemistry 30*, 3280−3289.
49. Varani, G., Aboulela, F., and Allain, F. H. T. (1996) *Prog. Nucl. Magn. Reson. Spectrosc. 29*, 51−127.
50. Jaeger, J. A., and Tinoco, I. (1993) *Biochemistry 32*, 12522−12530.
51. Varani, G., and Tinoco, I. (1991) *Q. Rev. Biophys. 24*, 479−532.
52. Burkard, M. E., Kierzek, R., and Turner, D. H. (1999) *J. Mol. Biol. 290*, 967−982.
53. Leontis, N. B., and Westhof, E. (1998) *Q. Rev. Biophys. 31*, 399−455.
54. Wahl, M. C., Rao, S. T., and Sundaralingam, M. (1996) *Nat. Struct. Biol. 3*, 24−31.
55. Butcher, S. E., Dieckmann, T., and Feigon, J. (1997) *EMBO J. 16*, 7490−7499.
56. Giessner-Prettre, C., and Pullman, B. (1976) *Bioc. Biop. Res. Comm. 70*, 578−581.
57. Jucker, F. M., and Pardi, A. (1995) *RNA 1*, 219−222.

58. Fountain, M. A., and Krugh, T. R. (1995) *Biochemistry 34*, 3152−3161.
59. Quigley, G. J., and Rich, A. (1976) *Science 194*, 796−806.
60. Pley, H. W., Flaherty, K. M., and McKay, D. B. (1994) *Nature 372*, 68−74.
61. Wu, M., and Tinoco, I. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 11555−11560.
62. Williamson, J. R. (2000) *Nat. Struct. Biol. 7*, 834−837.

BI011226X