# NMR Reveals the Absence of Hydrogen Bonding in Adjacent UU and AG Mismatches in an Isolated Internal Loop from Ribosomal RNA[†,‡]

Neelaabh Shankar,[§] Tianbing Xia,[⊥] Scott D. Kennedy,[§] Thomas R. Krugh,[#] David H. Mathews,[§] and Douglas H. Turner*[,#,||]

*Department of Biochemistry and Biophysics, Center for Pediatric Biomedical Research and Department of Pediatrics, School of Medicine and Dentistry, University of Rochester, Rochester, New York 14642, Department of Molecular and Cell Biology, University of Texas at Dallas, Richardson Texas 75083, and Department of Chemistry, University of Rochester, Rochester, New York 14627*

*Received April 27, 2007; Revised Manuscript Received July 24, 2007*

ABSTRACT: NMR studies provide insights into structural features of internal loops. These insights can be combined with thermodynamic studies to generate models for predicting structure and energetics. The tandem mismatch internal loop, $^{5'G\mathbf{UG}3'}_{3'C\mathbf{UA}C5'}$, has been studied by NMR. The NMR structure reveals an internal loop with no hydrogen bonding between the loop bases and with the G in the AG mismatch flipped out of the helix. The sequence of this internal loop is highly conserved in rRNA. The loop is located in the large ribosomal subunit and is part of a conserved 58-nt fragment that is the binding domain of ribosomal protein L11. Structural comparisons between variants of this internal loop in crystal structures of the 58-nt domain complexed with L11 protein and of the large ribosomal subunit (LSU) suggest that this thermodynamically destabilizing internal loop is partially preorganized for tertiary interactions and for binding L11. A model for predicting the base pairing and free energy of 2 × 2 nucleotide internal loops with a purine−purine mismatch next to a pyrimidine−pyrimidine mismatch is proposed on the basis of the present NMR structure and previously reported thermodynamics.

RNA sequences are becoming available at a tremendous rate, and new functions ranging from catalysis (*1*, *2*) to gene regulation (*3*−*6*) are being discovered. Prediction and determination of RNA secondary structures is a natural extension toward application of the enormous amount of sequence information. Understanding the structural and functional characteristics of various RNA sequences is a step toward the correct prediction of secondary and tertiary structure. Predictions on the basis of free energy minimization facilitate determination of secondary structure (*7*−*15*). It is important to understand the sequence dependence of stability and structure of non-Watson−Crick regions of RNA secondary structure motifs in order to improve the accuracy of structure prediction (*16*).

The thermodynamic stabilities of 2 × 2 nucleotide internal loops are very sequence dependent (*17*, *18*). In order to predict the thermodynamics of 2 × 2 nucleotide sequence non-symmetric internal loops, the current model uses averages of the loop free energies of 2 × 2 sequence symmetric internal loops with the same mismatches (*19*). Thermodynamics of sequence non-symmetric tandem mismatches (*20*, *21*), however, suggest that differences in base size and hydrogen bonding of the two mismatches are major factors in determining stability. This likely reflects the local backbone distortions required to accommodate the internal loop. Therefore, it is important to understand the structural basis for the sequence-specific energetics of sequence non-symmetric 2 × 2 internal loops in order to improve models for predicting structure and energetics.

A previous NMR study (*22*) on symmetric tandem AG mismatches closed by CG base pairs, $^{5'C\mathbf{AG}3'}_{3'G\mathbf{GA}C5'}$, showed that the AG mismatches have the imino-hydrogen bonded conformation shown in Figure 1. Imino proton NMR (*18*, *23*) and X-ray crystal structures (*24*, *25*) show that tandem UU mismatches also have strong hydrogen bonds involving imino protons. The $^{5'\mathbf{AG}3'}_{3'\mathbf{GA}5'}$ and $^{5'\mathbf{UU}3'}_{3'\mathbf{UU}5'}$ motifs are thermodynamically stabilizing (*17*, *18*, *21*, *26*), but the $^{5'G\mathbf{AU}G3'}_{3'C\mathbf{GU}C5'}$ and $^{5'G\mathbf{UG}G3'}_{3'C\mathbf{UA}C5'}$ internal loops, with two differently sized mismatches, are thermodynamically destabilizing (*21*).

A 1D imino proton NMR spectrum for the internal loop $^{5'G\mathbf{AU}G3'}_{3'C\mathbf{GU}C5'}$ (*21*) provided no evidence for hydrogen bonding in the AG and UU mismatches. Because of severe spectral overlap in the imino proton region, the structures of the mismatches in a related motif, $^{5'G\mathbf{UG}G3'}_{3'C\mathbf{UA}C5'}$, could not be inferred from the 1D spectrum of a duplex model system (*21*). To provide additional structure−energetics−function relationships, this article presents a 2D NMR study of the internal loop, $^{5'G\mathbf{UG}G3'}_{3'C\mathbf{UA}C5'}$. Two questions addressed are (a) Is the structure of the loop consistent with models used for predicting stabilities of 2 × 2 internal loops? (b) Is the internal loop preorganized for tertiary interactions?

[†] This work was supported by NIH Grants GM22939 (to D.H.T.) and GM076485 (to D.H.M.).

[‡] The coordinate 2jse was deposited in the Protein Data Bank.

* To whom correspondence should be addressed. Phone: (585) 275-3207. Fax: (585) 276-0205. E-mail: turner@chem.rochester.edu.

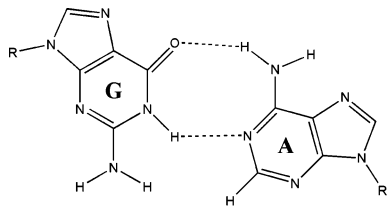
[§] Department of Biochemistry and Biophysics, University of Rochester.

[||] Center for Pediatric Biomedical Research and Department of Pediatrics, University of Rochester.

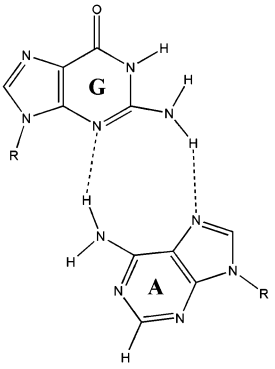
[⊥] University of Texas at Dallas.

[#] Department of Chemistry, University of Rochester.

10.1021/bi700802s CCC: $37.00 
Published on Web 10/11/2007




Imino GA pair




Sheared GA pair

FIGURE 1: Imino and sheared hydrogen bonded conformations of GA pairs.

The $^{5'\mathbf{UG}3'}_{3'\mathbf{UA}5'}$ motif in 2 × 2 internal loops occurs relatively often in natural RNA. For example, the $^{5'\mathbf{UG}3'}_{3'\mathbf{UA}5'}$ motif occurs 14 times in 398 RNA secondary structures determined by sequence comparison (*21*). Most of the natural occurrences of the internal loop, $^{5'\mathbf{GUGG}3'}_{3'\mathbf{CUAC}5'}$, are a variant corresponding to part of helix A of a conserved 58-nt fragment (Figure 2) of large subunit ribosomal RNA (LSU[1] rRNA) that is the recognition site for ribosomal protein L11 and for thiostrepton antibiotics (*27*−*29*). The 58-nt domain contains an extensive set of RNA tertiary (*30*−*33*) and quaternary interactions with the conserved protein L11 (*33*, *34*). Presumably, this thermodynamically unfavorable internal loop motif has been selected during evolution for structural and functional roles because of a unique arrangement of hydrogen-bonding groups and backbone conformation required for tertiary folding and molecular recognition.

In the NMR structure of helix A taken from *Escherichia coli* (Figure 2a), where the natural sequence in the internal loop is $^{5'\mathbf{GUUG}3'}_{3'\mathbf{CUAC}5'}$ (*35*), there is a one-hydrogen bond UU mismatch flanked by Watson−Crick GC and AU pairs. Two crystal structures of RNA−protein complexes between the 58-nt rRNA domain and protein L11 have been solved (*36*, *37*). One is the *E. coli* rRNA domain with a U1061 → A mutation in the internal loop region of helix A (Figure 2a) complexed with the C-terminal domain of L11 (*36*), and the other is the *Thermotoga maritima* rRNA domain without the mutation, complexed with the whole L11 protein (*37*). The structures of the RNA in the two complexes are essentially the same, and the structures of the corresponding protein components are also the same and in good agreement with previous structures of the components (*38*−*40*) and with structures of LSUs (*41*, *42*) and a whole ribosome (*43*).

The internal loop chosen for this study, $^{5'^{4}\ \mathbf{G}\mathbf{U}\mathbf{G}^{7}3'}_{3'_{19}\mathbf{C}\mathbf{U}\mathbf{A}\mathbf{C}_{16}5'}$, occurs in the LSU rRNA of *Sulfolobus acidocaldarius* (Figure 2a), *Giardia intestinalis*, and in mitochondrial rRNA of *Xenopus laevis* and *Zea mays* (*44*). Its sequence is similar to the equivalent internal loop in the 58-nt domains of *E. coli* and *T. maritima* (Figure 2a and b). The structural information provides a way to understand the basis of the observed thermodynamics of sequence non-symmetric tandem mismatches and provides insight into the importance of local interactions in determining three-dimensional structures.

## MATERIALS AND METHODS

*Oligoribonucleotide Design, Synthesis, and Purification.* The RNA sequence, 5′GGAGUGGCCGAAAGGCAUCUCC3′, was designed to form the hairpin structure shown in Figure 2d, which contains the 2 × 2 internal loop, $^{5'\mathbf{GUGG}3'}_{3'\mathbf{CUAC}5'}$. The GAAA tetraloop was chosen because it has been studied by NMR (*45*, *46*), and it helps ensure correct secondary structure formation. An AU pair was included in the stem to serve as an NMR reporter for the formation of the intended stem. The RNA was synthesized on an Applied Biosystems 392 DNA/RNA synthesizer using standard phosphoramidite chemistry (*47*, *48*). Base-protecting groups were removed by incubation in 2 mL ammonia/ethanol solution (3:1, v/v) at 55 °C overnight (*49*), and the solid support was removed by filtration. The filtrate was lyophilized, and silyl protecting groups were removed by incubation in a 9:1 (v/v) TEA-3HF (triethylamine trihydrofluoride)/DMF (*N*,*N*-dimethyl formamide) mixture. RNA was precipitated with 2-butanol, lyophilized, and desalted using a GE PD-10 Sephadex gel filtration column. The eluted sample was purified on a preparative denaturing PAGE gel with 8 M urea and 20% acrylamide. The appropriate band on the gel was identified by UV shadowing and excised. RNA was eluted from the gel slice with a Schleicher and Schuel electroeluter and desalted. The purity of the sample was confirmed by gel electrophoresis after 5′-$^{32}$P labeling with T4 polynucleotide kinase. All the samples were >95% pure.

The $^{13}$C and $^{15}$N labeled samples were prepared with an Epicenter Ampliscribe T7 Flash transcription kit, labeled rNTPs (Silantes GmbH, München, Germany) and synthetic DNA oligonucleotides (Integrated DNA Technologies, Inc.). RNA was purified and analyzed for purity with methods similar to those described above.

*NMR Sample Preparation.* Sample preparation was similar to that described in Shankar et al. (*50*) with minor modifications. The RNA samples were dialyzed overnight against 1 L of filtered autoclaved water in a Gibco Life Technologies microdialysis system with a 1000 MWCO Spectro-por dialysis membrane. After dialysis, the sample was lyophilized and dissolved in 250−300 μL of NMR buffer (80 mM NaCl, 10 mM sodium phosphate, and 0.5 mM $Na_2EDTA$ at pH 6). For exchangeable proton spectra, the sample was reconstituted in 90% $H_2O$ with 10% $D_2O$. For nonexchangeable proton spectra, $D_2O$ exchange was done by three repetitions of lyophilization with 99.996% $D_2O$ as the solvent, and the sample was finally dissolved in 300 μL of $D_2O$ from

[1] Abbreviations: DQF-COSY, double-quantum-filtered correlation spectroscopy; HETCOR, heteronuclear correlation; HSQC, heteronuclear single-quantum coherence; LSU, large ribosomal subunit; MWCO, molecular weight cutoff; N, any nucleotide, i.e., A, C, G, or U; NMR, nuclear magnetic resonance spectroscopy; NOESY, nuclear Overhauser spectroscopy; R, any purine, i.e., G or A; RDC, residual dipolar coupling; TOCSY, total correlation spectroscopy; Y, any pyrimidine, i.e., U or C.

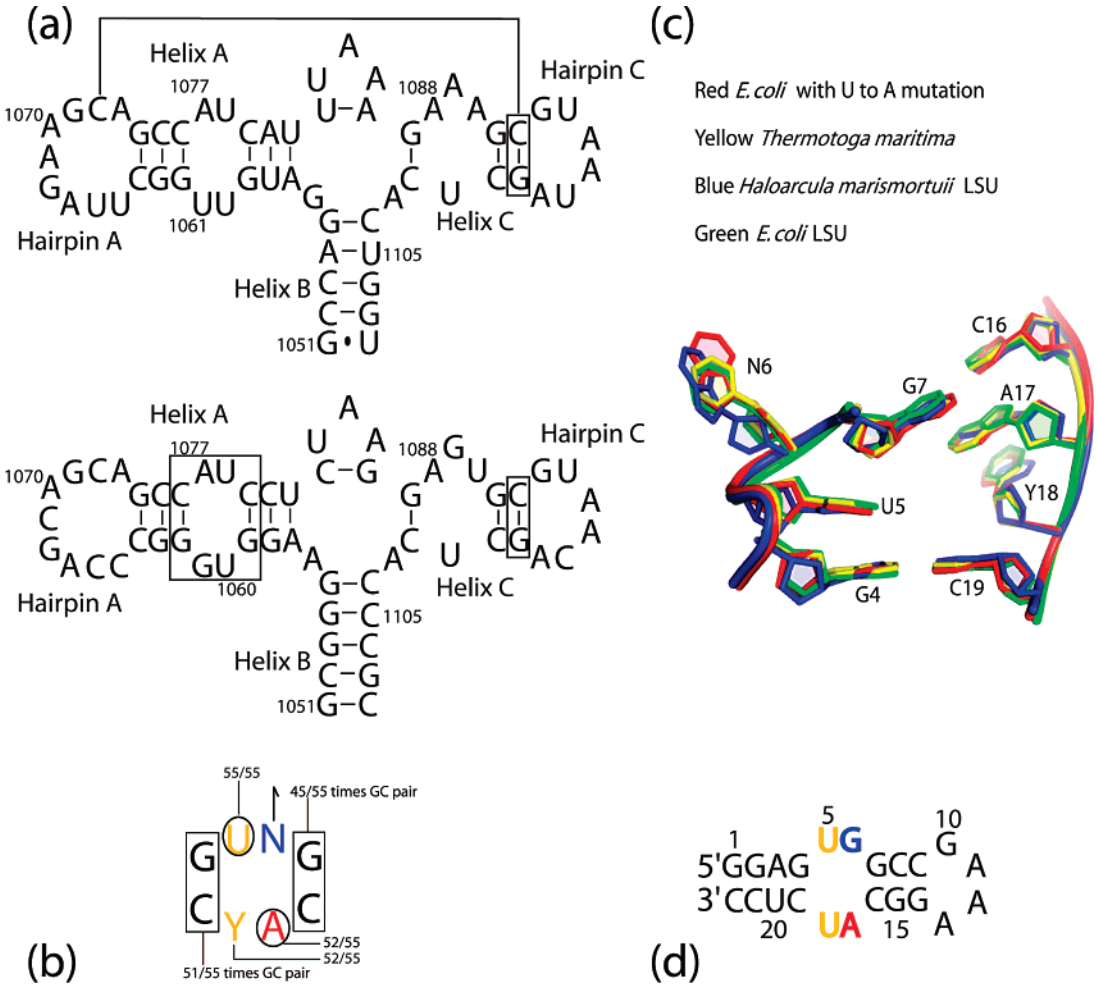



FIGURE 2: (a) Secondary structure of the conserved 58-nt fragment of the large ribosomal subunit (LSU). This region is the binding site for L11 ribosomal protein and the thiostrepton family of antibiotics. *E. coli* numbering has been used. The top sequence is from *E. coli*. Conn et al. (*36*) used a U1061 to A mutation for crystal structure determination. The lower sequence is from *S. acidocaldarius*; the highly conserved internal loop is boxed. (b) Sequence conservation of the internal loop is shown for 55 sequences with 2 × 2 nucleotide internal loops taken from Robin Gutell's website [http://www.rna.icmb.utexas.edu]. The circled residues and the sequence and orientation of the closing pairs are highly conserved (*30*). N is any nucleotide, and Y is a pyrimidine. The G7C16 pair is an AU pair in some organisms. (c) A heavy atom overlap of internal loops from mutated *E. coli* ( $^{5'1059}$**GUAG**$^{1062}3'$ shown in red (*36*), $^{5'1059}$**GUUG**$^{1062}3'$ shown in green (*43*)), *Thermotoga maritima* ( $^{5'1059}$**GUUG**$^{1062}3'$ shown in yellow (*37*)), and *Haloarcula marismortui* ( $^{5'1163}$**GUGA**$^{1166}3'$ shown in blue (*41*)). The residues are numbered as shown in (d), the hairpin studied by NMR.

Cambridge Isotope Laboratories. The hairpin concentration was approximately 2.5 mM for the unlabeled sample and 0.6 mM for one $^{13}C$ and $^{15}N$ labeled sample.

Residual dipolar couplings (RDCs) were measured with $^{13}C$-$^{15}N$ labeled RNA. The sample used for RDC measurements was partially aligned with 6.5% C12E5/hexanol ($r = 0.96$) (*51*). Buffer conditions and RNA concentrations were the same before and after addition of alignment media (0.1 mM $^{13}C$ and $^{15}N$ labeled RNA in 325 $\mu$L of NMR buffer at pH 6). The $D_2O$ resonance in the aligned RNA sample showed ~33 Hz splitting at pH 6 and 35 °C with a line width of ~1.5 Hz.

*NMR Spectroscopy.* NMR spectra were collected on Varian Inova 500 and 600 MHz spectrometers. Proton spectra were referenced to known temperature-dependent chemical shifts of $H_2O$ or HDO relative to 3-(trimethylsilyl) tetradeutero sodium propionate (TSP). All spectra were processed with NMRpipe (*52*), and resonance assignments were made with Sparky (*53*).

*Distance, Dihedral Angle, and RDC Restraint Generation.* Distance restraints were obtained from 100 and 200 ms mixing time NOESY spectra at 25 and 35 °C and 100 ms SNOESY spectra at 10 and 25 °C. The box integration method in the Sparky software (*53*) was used for cross-peak volume calculation. The volumes were converted to CNS distance restraint input by a C++ program, using $1/r^6$ scaling relative to the average of pyrimidine H5−H6 cross-peak volumes (2.45 Å) and a two-spin approximation. Error limits of ±30% (100 ms mixing time) and ±40% (200 ms mixing time and SNOESY spectra) were assigned to the NOE-derived distances to allow for relaxation, spin diffusion, baseline distortions, water exchange, and noise. Restraints for overlapped peaks were estimated or discarded in case of extreme overlap. No restraints were used from H5′ and H5″ protons. A total of 234 interproton distance restraints and 104 dihedral angle restraints were used to model the 22-mer hairpin. Consistent with NMR data, the seven Watson−Crick pairs and the sheared GA pair in the tetraloop were subjected to 44 hydrogen bond distance restraints (1.8−2.5 Å between hydrogen and acceptor, 2.7−3.5 Å between the heavy atoms). Weak planarity restraints (5 kcal/mol Å$^2$) were also applied to the Watson−Crick pairs.

Because of severe overlap in DQF-COSY and HETCOR spectra, torsion angle restraints were not determined from the direct measurement of coupling constants but were set to wide ranges. No torsion angle restraints were used for the internal loop and the hairpin-loop regions, except for $\delta$ and $\chi$. Backbone dihedral angles for the Watson−Crick stem residues were loosely restrained: $\alpha$ (0 ± 120°), $\beta$ (180 ± 30°), $\gamma$ (60 ± 30°), $\delta$ (85 ± 30°), $\epsilon$ (−140 ± 40°), and $\zeta$ (0 ± 120°). For all residues, the glycosidic torsion angles were restrained in *anti* conformation, $\chi$ (−120 ± 90°) because the intensity of the intranucleotide cross-peak between H1′−H8/H6 was not comparable to the H5−H6 cross-peaks. G6 was also modeled without the $\chi$ restraint. The resulting models were not substantially different. H1′−H2′ scalar couplings in TOCSY (Supporting Information) and DQF-COSY spectra were used to identify residues in C3′ or C2′ *endo* sugar pucker. U5, G6, A17, and U18 from the internal loop and all residues of the GAAA tetraloop had weak H1′−H2′ cross-peaks indicating conformational sampling between C3′ and C2′ *endo* sugar puckers. These residues were restrained to cover both the C3′ and C2′ *endo* conformation with $\delta$ (122.5 ± 67.5°). The remaining residues were restrained to be in C3′ *endo* conformation, $\delta$ (85 ± 30°). See Supporting Information for complete tables of distance, dihedral angle, and endocyclic sugar torsion angle restraints.

$^1$H-$^{13}$C couplings were measured in isotropic and partially aligned samples from 2D $^1$H-$^{13}$C CT HSQC spectra. Residual dipolar couplings were determined from the difference in $^1$H-$^{13}$C couplings between the isotropic and partially aligned spectra (Isotropic−Aligned). A total of 33 RDC restraints were used to model the structure (Supporting Information). On the basis of two independent individual measurements of peak positions, the errors for RDC measurements were estimated to be ±2 Hz.

*Simulated Annealing and Structure Calculation.* Structures restrained by NOE distances and dihedral restraints were generated with CNS version 1.1 (*54*). An updated CNS version that improves the susceptibility anisotropy (RDC) protocol was used for the calculation of structures using experimental RDCs (*55*). Initial starting structures were built with Insight II (Biosym). The structures were calculated with the following protocol using implicit solvent: (1) high-temperature dynamics at 3000 K in torsion angle space for 4 ps with NOE and dihedral scale factors of 150 kcal/mol Å$^2$ and 25 kcal/mol rad$^2$, respectively; (2) simulated annealing in torsion angle space for 37.5 ps with slow cooling from 1000 to 0 K (25,000 steps) with NOE and dihedral scale factors of 150 kcal/mol Å$^2$ and 200 kcal/mol rad$^2$, respectively; (3) simulated annealing for 37.5 ps in Cartesian space with slow cooling from 1000 to 0 K (100,000 steps) with the NOE and dihedral angle scale factors constant at 150 kcal/mol Å$^2$ and 200 kcal/mol rad$^2$, respectively; the van der Waals factor was linearly increased from 1 to 4; and (4) Powell energy minimization was applied with van der Waals and electrostatic scale factors set to 1 for all atoms. A total of 20 structures were calculated with and without the RDC restraints, out of which 10 lowest energy structures without distance and dihedral violations were chosen for further analysis. Incorporation of RDC restraints in structure modeling was carried out with a protocol similar to that used by Lukavsky et al. (*56*). Initial values of the axial component ($D_a$) and the rhombicity (R) of the principal alignment tensor were estimated from analysis of the powder pattern distribution of measured RDCs. Lowest energy structures calculated for a pair of Da and R values were then analyzed by PALES (*57*, *58*) to predict the components of the alignment tensor. The averages of the predicted $D_a$ and R values were then used to calculate 20 new structures, and this process repeated until the average values of $D_a$ and R in subsequent cycles varied less than ± 0.4 Hz and ± 0.02, respectively.

All of the 10 accepted structures agreed with experimentally derived NOE and dihedral restraints within 0.2 Å and 5°, respectively. Pymol (*59*) was used to visualize and analyze the structures.

*Optical Melting Experiments.* The 22-mer oligonucleotide was lyophilized and dissolved in melt buffer (1.0 M NaCl, 20 mM sodium cacodylate, and 0.5 mM $Na_2EDTA$ at pH 6 and 7), in magnesium buffer (150 mM KCl, 5 mM $MgCl_2$, and 20 mM sodium cacodylate at pH 7), and in NMR buffer (80 mM NaCl, 0.5 mM EDTA, and 10 mM phosphate at pH 6 and 7). Strand concentrations for melting curves in the buffers ranged from 5 $\mu$M to 3 mM. Extinction coefficients were predicted from those of dinucleotide monophosphates and nucleosides (*60*, *61*). Absorbance versus temperature melting curves were acquired at 280 nm at heating rates of 1 °C/min with a Beckman Coulter DU640C spectrophotometer having a Peltier temperature controller cooled with flowing water. A mutant of the 22-mer, 5′GGAGUIGCCGAAAGGCAUCUCC3′, where G6 is replaced by inosine to give an IA mismatch, was also melted in NMR buffer at 1 mM strand concentration.

## RESULTS

*Sequence and Structure Conservation.* Analysis of 55 secondary structures of LSUs with 2 × 2 nucleotide internal loops in helix A (*44*) revealed that this internal loop sequence is highly conserved (Figure 2a and b). U1060 and A1077 are universally conserved and involved in tertiary interactions. Mutation of either of these bases destabilizes the RNA tertiary structure and reduces protein binding affinity (*28*, *30*). Position 1061 (shown as N in Figure 2b) is not conserved. In *E. coli* and many other eubacteria, nucleotide 1061 is a U, while in many archaebacteria and most eukaryotes, there is a G or A at position 1061 (*62*, *63*). The closing base pairs and their orientation in these internal loops are also highly conserved. Moreover, the structure is highly conserved in crystals. Figure 2c shows structures from the 58-nucleotide domain of LSU rRNA of *E. coli* ($^{5'1059}$**GUAG**$^{1062}$3′ / $^{3'1079}$**CUAC**$_{1076}$5′, PDB entry 1QA6, with U1061 → A mutation), *Thermotoga maritima* ($^{5'1059}$**GUUG**$^{1062}$3′ / $^{3'1079}$**CUAC**$_{1076}$5′, PDB entry 1MMS) (*36*, *37*), and from the entire LSU of *E. coli* ($^{5'1059}$**GUUG**$^{1062}$3′ / $^{3'1079}$**CUAC**$_{1076}$5′, PDB entry 2AWB) and *H. marismortui* ($^{5'1163}$**GUGA**$^{11663}$′ / $^{3'1183}$**CCA**U$_{1180}$5′, PDB entry 1JJ2) (*41*, *43*). There is no hydrogen bonding between the bases within the internal loop, but the bases help maintain the functional structure. It is believed that the L11 ribosomal protein locks helix A in the peculiar conformation to allow intercalation of A1088 to pair with U1060. NMR studies of the 1057−1081 portion of helix A from *E. coli* show that this internal loop forms a UU imino pair and an AU Watson−Crick pair when not involved in protein contacts (*35*). This structure is different from the structure in the

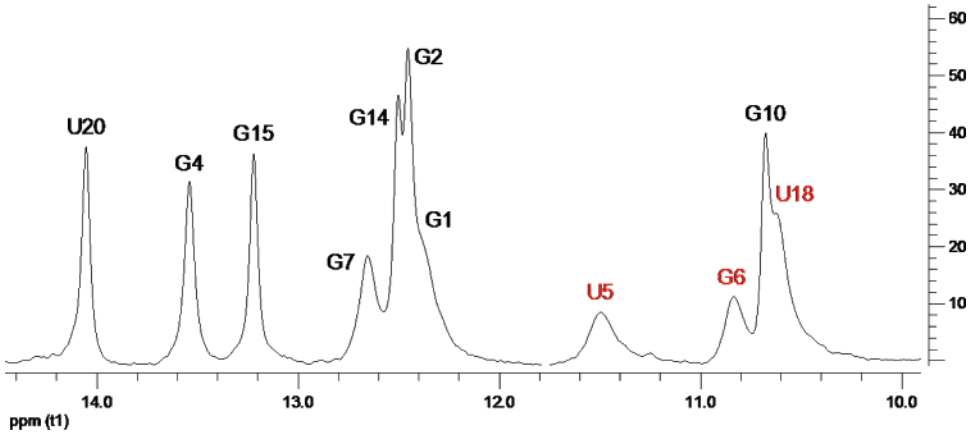



FIGURE 3: One-dimensional spectrum of the hairpin (Figure 2d) at 10 °C at pH 6. The assignments are shown above the peaks. The assignments in red are tentative. One-dimensional spectra at temperatures ranging from 5 °C to 55 °C are shown in the Supporting Information.

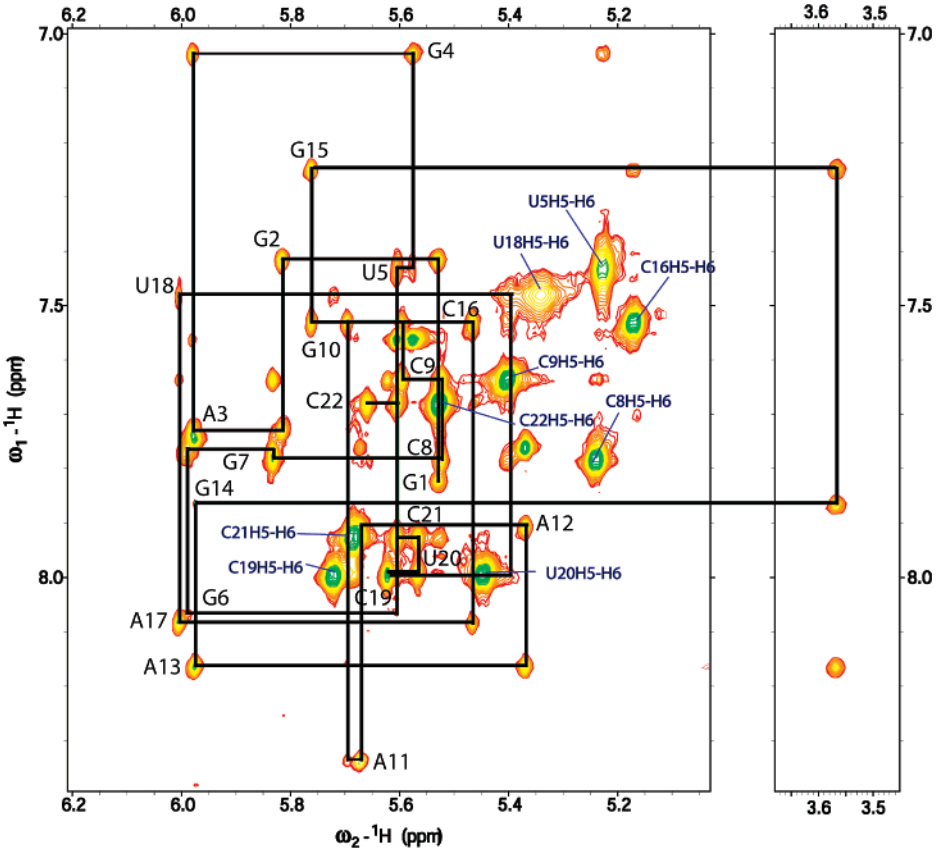



FIGURE 4: NOESY walk region (H8/H6 H1′/H5) of the hairpin (Figure 2d) NOESY spectrum at 100 ms mixing time, 35 °C, and pH 6. Black lines trace the NOESY walk for the hairpin; black labels along the horizontal lines mark the H8/H6 resonance for the residue.

crystals complexed with L11 protein and supports the role of L11 in rearranging the base pairing within the internal loop for A1088 insertion.

*NMR Spectra of Exchangeable and Nonexchangeable Protons.* Exchangeable protons for the hairpin shown in Figure 2d were assigned using 1D (Figure 3), 2D NOESY from 0 to 25 °C with mixing times of 100 and 150 ms using an S-shaped excitation pulse (*64*) for water suppression, and NHSQC spectra. The presence of seven imino proton resonances between 12 and 14.5 ppm indicates the formation of seven Watson−Crick base pairs including one AU pair (Figure 3). The resonance at 10.67 ppm is consistent with the sheared GA pair in the tetraloop (*45*). The three broader resonances between 10 and 12 ppm are from the imino protons of the two U bases and the G in the internal loop.

Nonexchangeable protons were assigned by standard procedures (*65*, *66*). NOESY spectra of the sample in $D_2O$ were acquired at 15, 25, and 30 °C with mixing times of 100, 200, and 400 ms. The NOESY walk (H8/H6/H2 to H1′/H5) region of a 100 ms spectrum is shown in Figure 4. The pyrimidine H5−H6 cross-peaks are easily identified from their high volume, splitting pattern, and presence in the same region of DQF-COSY and TOCSY spectra (data not shown). $^{1}$H-$^{13}$C HSQC spectra at 35 °C confirmed the H8/H6/H2/H1′ assignments. Assignments of sugar resonances were based on NOESY, DQF-COSY, TOCSY, and $^{1}$H-$^{31}$P HETCOR spectra. The AH2 protons were identified from a $^{1}$H-$^{13}$C HSQC spectrum. Proton chemical shift assignments are

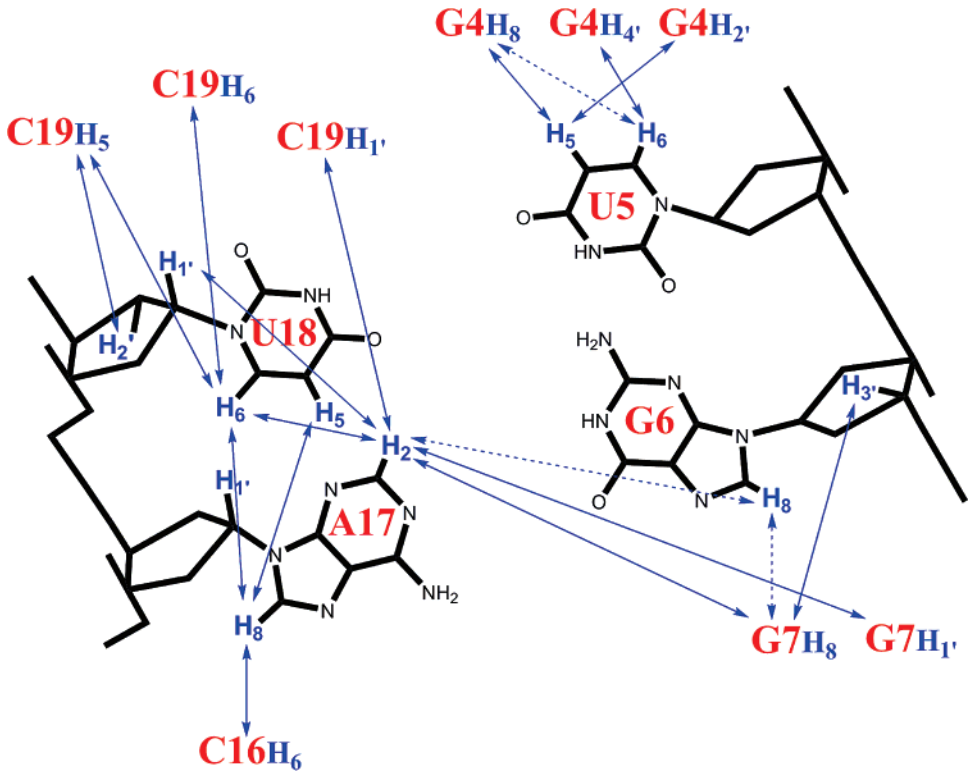



FIGURE 5: Important interstrand and intrastrand NOE restraints for the 2 × 2 nucleotide internal loop shown as solid blue arrows. Very weak NOE restraints (lower limit >3.5 Å and/or upper limit <7 Å) are shown as dotted blue arrows. The typical NOESY walk restraints and less important hydrogens are omitted for the sake of clarity.

shown in Supporting Information. Riboses for loop residues U5, U18, and A17 are primarily in the C3′-*endo* conformation with a small population in the C2′-*endo* form as exhibited by H1′−H2′ scalar couplings. The G6H1′−H2′ cross-peak has the largest coupling.

*Structure Determination with and without Residual Dipolar Couplings.* A total of 20 structures were calculated with a simulated annealing protocol both with and without the RDC restraints. Some important distance restraints for the internal loop are shown in Figure 5. All calculated structures converged to satisfy the distance and dihedral angle restraints within 0.2 Å and 5°, respectively. Ten structures were accepted on the basis of lowest energy and no distance or dihedral angle violations. Nuchemics (*67*) was used to predict chemical shift values for the 10 accepted structures. The predicted H1′, H5, H6, H8, and H2 chemical shifts for the loop residues (Supporting Information) agree well with the observed chemical shifts. To improve the long range order, the structures were calculated again with 33 experimentally derived RDCs. The axial component ($D_a = 21$ Hz) and the rhombicity ($R = 0.44$) of the principal alignment tensor were optimized as described in Materials and Methods. The experimentally observed RDCs agree well with the average of calculated RDCs for 10 accepted structures (Figure 6). The superimposed structures calculated with and without RDCs show significant improvement of long range order (Figure 7). Supporting Information lists the RMSD against an average calculated structure for each residue. The RMSDs show marked improvements over the structures calculated without RDC restraints. The two stem regions of 7 Watson−Crick base pairs form two segments of helix close to A-form. Figure 8 shows the hairpin observed from both the major and minor grooves. No structure has base pairing

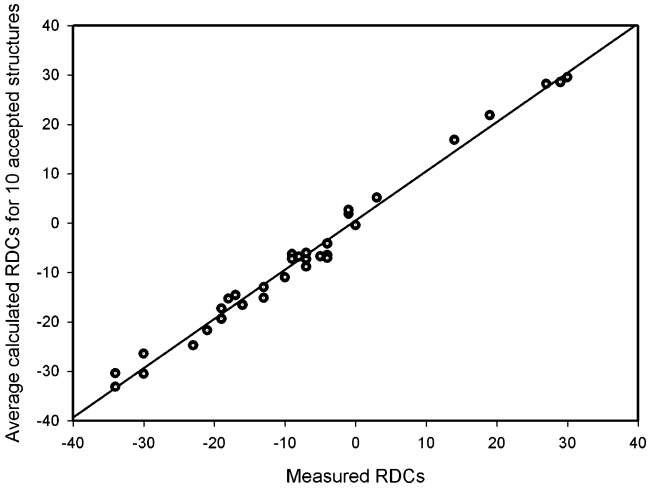



FIGURE 6: Plot of 33 measured RDCs against the average of calculated RDCs for accepted structures of the hairpin ($R^2 = 0.98$).

within the internal loop. Forcing a sheared GA or imino GA base pair resulted in NOE violations. The G6 base shows conformational dynamics, and its position could not be restrained on the basis of the experimental data. The chemical shifts, connectivity pattern, restraints, and structure for the tetraloop region agree with those previously reported (*45*, *46*).

*UV and NMR Denaturation Profiles.* As expected, the oligonucleotide, 5′GGAGUGGCCGAAAGGCAUCUCC3′ (Figure 2d), has concentration independent thermal denaturation profiles (data not shown) in all buffers tested (see Materials and Methods), consistent with hairpin formation. The curves are biphasic, with the first transition at about 64 °C and the second at about 75 °C (Figure 9).

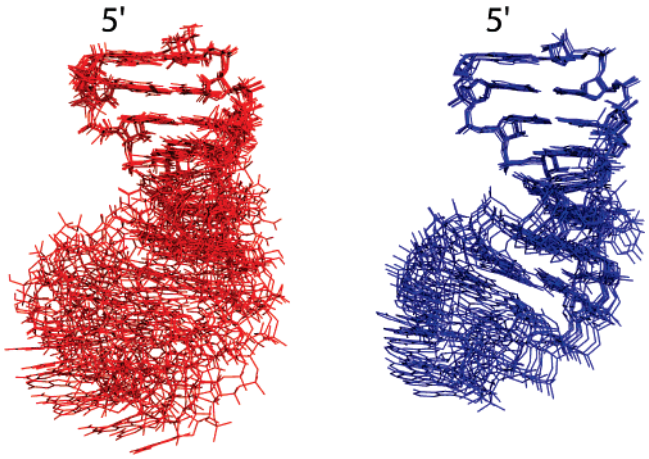



FIGURE 7: Ten accepted structures calculated without RDC restraints (red) and with 33 RDC restraints (blue) are shown with the top four Watson−Crick pairs superimposed. The RMSD for each residue against an average structure for both calculations (with and without RDCs) is listed in Supporting Information.

In NMR melting experiments, the 5′ stem imino proton resonances disappear at lower temperatures than the hairpin stem resonances (Supporting Information). Evidently, the lower and higher temperature transitions correspond to melting of the 5′ stem and the tetraloop hairpin, respectively. The temperature for most of the NMR experiments used for structure determination was 35 °C, which is below the melting temperature of the first transition (Figure 9).

A G6 → inosine substitution has little effect on the UV melting profile, except to reduce the hyperchromicity (Figure 9). Replacing G by inosine in a tandem GA internal loop can stabilize a duplex when the G forms an imino-hydrogen bonded or sheared base pair and presumably results in imino hydrogen-bonded I−A pairs in both cases (*68*). Inosine, however, destabilizes a sheared GA if the sheared conformation is in a GCAA tetraloop (*69*). Because the oligonucleotides with GA and IA mismatches studied here have essentially identical melting behavior (Figure 9), the melting experiments are consistent with the NMR structure, which shows no hydrogen bonding between G and A.

*G6 and A17 Do Not Form any Recognizable Base Pair, and G6 May Be Flipped Out of the Helix.* G6 and A17 residues within the internal loop do not form any recognizable pair in the calculated structures (Figures 8 and 10). The width and chemical shift of the G6 imino resonance in the 1D spectrum (Figure 3) suggest that the proton is not involved in strong hydrogen bonding. With one exception, the NOESY walk extends throughout the molecule, suggesting mostly A-form helix. It is discontinuous in the internal loop between U5 and G6, suggesting deviation from the A-form. At 35 °C, the G6H8−U5H1′ peak was only seen at mixing times longer than 200 ms. In a HETCOR spectrum (data not shown), the $^{31}$P chemical shift for G6 (−0.060 ppm) is outside the ∼1 ppm range for the A-form helix. This is also consistent with the lack of A-form geometry for the G6 residue (*70*).

The amplitude of G6H8 is low compared to most other H8 peaks in 1D $D_2O$ and $^{13}$C HSQC spectra. This suggests line broadening due to conformational sampling. Because of this line broadening and partial spectral overlap with A17H8, it is difficult to measure the NOE volumes for G6H8. All of the peaks associated with G6H8, including G6H8-G6H2′ and G6H8-G6H3′, are very weak compared to similar peaks in the spectrum. The G6 H1′−H2′ coupling is larger than other loop residues, indicating a much greater C2′-*endo* population, perhaps as much as 50% (Supporting Information). This also indicates dynamics at G6. Base dynamics is further supported by very weak NOE cross-peaks between G6H1′ and G6H8 and between G6H1′ and G7H8. In the calculated models, the G6 base is flipped out of the helix. The base position, however, could place the Watson−Crick side of the G facing the minor or major groove. The position of the G6 base cannot be conclusively modeled because of the lack of direct cross-peaks to the imino proton. (See the blue colored G6 base in Figure 8.) The lack of any cross-peaks to the G6 imino proton could be because no proton is close enough to show an NOE, or it could be because the imino proton is exchanging rapidly with water. If G6 were exclusively in the major groove, then a stronger than observed A17H2−G6H8 peak would be expected. Complete major groove positioning of G6 would also increase the distance between G6H1′ and G7H8, resulting in a minor distance violation of ∼0.1 Å.

The symmetric tandem mismatch motif $^{5'\mathbf{GA}3'}_{3'\mathbf{AG}5'}$ in RNA has sheared GA pairs (Figure 1) when closed by CG pairs, $^{5'\mathrm{C}\mathbf{GA}\mathrm{G}3'}_{3'\mathrm{G}\mathbf{AG}\mathrm{C}5'}$ (*71*), and imino GA pairs (Figure 1) when closed by GC pairs, $^{5'\mathrm{G}\mathbf{GA}\mathrm{C}3'}_{3'\mathrm{C}\mathbf{AG}\mathrm{G}5'}$ (*72*). In contrast, the $^{5'\mathbf{AG}3'}_{3'\mathbf{GA}5'}$ motif cannot form sheared AG pairs because of geometric constraints (*73*). Sheared AG pairs have a short P−P interstrand distance of 12.5 Å 5′ to the adenine, which is not compatible with the P−P distance of 17.5 Å required for a Watson−Crick base pair in an A-form helix. For the internal loop in this study, $^{5'\mathrm{G}\mathbf{UG}\mathrm{G}3'}_{3'\mathrm{C}\mathbf{UA}\mathrm{C}5'}$, the same geometric constraint on the AG pair should apply, i.e., the AG mismatch with a CG pair 5′ to the adenine cannot form a sheared conformation. The lack of a strong cross-strand NOE from A17H2 to G7H1′ (*71*) and the denaturation study with G−A and I−A mismatches (Figure 9) supports this conclusion.

There is no severe geometric constraint, however, to prevent the G6A17 mismatch from forming an imino hydrogen bonded conformation (*22*). An imino bonded AG pair has an imino proton chemical shift at ∼12 ppm (*22*, *72*), but the G6 imino proton does not resonate close to that region. Also, a hydrogen bond between G6 and A17 would slow the imino−water exchange rate in favor of a cross-peak between A17H2 and the G6 imino proton (*72*). Lack of this or any other cross-peak to the G6 imino proton suggests that G6 and A17 do not form an imino pair.

An imino G6A17 pair would result in an outward stretched backbone with a C1′−C1′ distance of ∼13 Å between the AG, compared with ∼10.9 Å for a Watson−Crick pair. The ∼13 Å also contrasts with dimensions for a UU pair with one or two hydrogen bonds between the bases and C1′−C1′ distances of ∼10.5 Å and ∼9 Å, respectively (*24*, *25*, *35*). Perhaps, forming hydrogen bonds between the two U bases is unfavorable because the cross-strand distance would lead to unfavorable steric interactions between the G6 and A17 bases. The C1′−C1′ distance between G6 and A17 is ∼12.5 Å in the average structure calculated from the 10 accepted structures. This suggests that the grooves are wide, exposing the internal loop to solvent.

Because neither the sheared nor imino-hydrogen bonded conformations for the AG mismatch has a geometry compat-

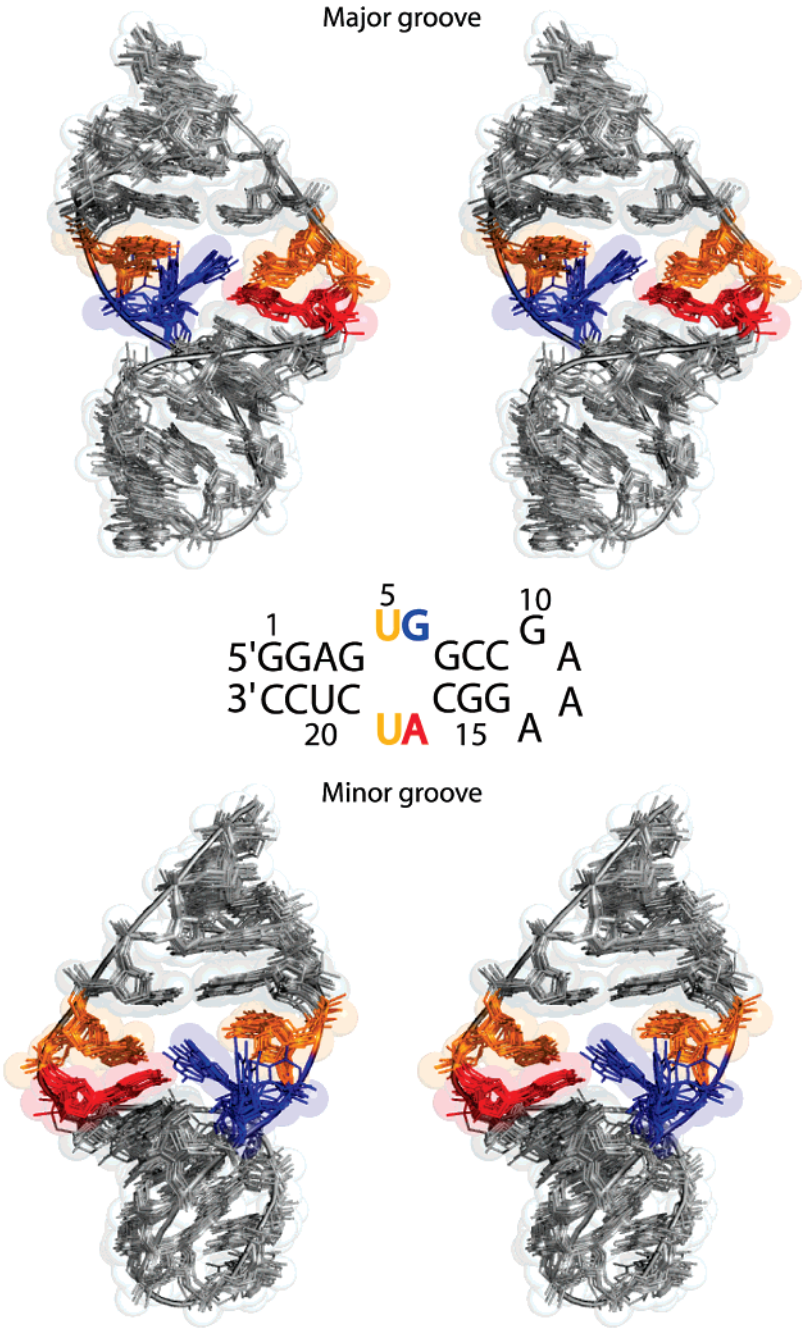



FIGURE 8: Stereoview of the major groove and the minor groove of the ten accepted structures calculated with RDCs. All structures are superimposed by the heavy atoms. The backbone and the spheres trace residues of the calculated average structure. The residues are color coded as shown in the hairpin sequence.

ible with both the adjacent Watson−Crick CG base pair and an adjacent hydrogen-bonded UU pair, the GA and UU are unlikely to form hydrogen bonded conformations simultaneously.

*No Base Pairing between U5 and U18 in the Internal Loop.* A strong NOE is seen between U iminos involved in a UU pair (*18*, *23*, *35*). No imino to imino cross-peaks are observed, however, between U5 and U18, suggesting that they do not form hydrogen bonds. The imino (UH3) might be exchanging rapidly with water, thus explaining the lack of cross-peaks, but rapid exchange also excludes the possibility of tight hydrogen bonding.

The C1′−C1′ distance between U5 and U18 in the average model is 13.0 Å, which is wider than the ∼10.9 Å distance for a Watson−Crick pair. Evidently, functional groups of U5 and U18 are too far away to form hydrogen bonds involving imino protons, which is consistent with the broad imino peaks in the exchangeable proton spectrum (Figure 3). U18 also has a much broader H5−H6 cross-peak compared to other H5−H6 cross-peaks, which suggests conformational sampling for the pyrimidine ring. A cross-peak is observed between A17H2 and C19H1′ (distance restraint 2.8−5.6 Å), when the normal distance between AH2 and ($n$ + 2) H1′ is >7.5 Å. This suggests that the U18 base is turned away from the U5 base such that the distance between A17 and C19 is reduced. These observations are consistent with the lack of hydrogen bonding and stacking of U18 in the modeled structure. The NMR data suggest that U18 primarily spends time outside the hydrogen-bonding distance of U5 but do not rule out the possibility of U18

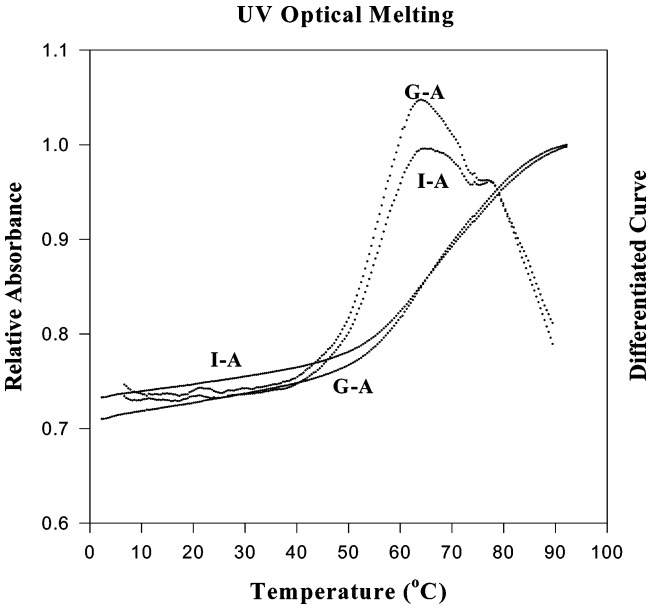



FIGURE 9: UV melting curves for the hairpin (G−A) and the hairpin with G6 to inosine (I−A) substitution in NMR buffer, 80 mM NaCl, 0.5 mM EDTA, and 10 mM phosphate at pH 7. The two melting curves are similar with the first transition at 64 °C and the second transition at 75 °C.

momentarily forming hydrogen bonds with U5. U5H5−G4H8 and U5H6−G4H8 cross-peaks suggest that U5 maintains the A-form stacking pattern (Figure 5).

*RNA Secondary Structure Prediction.* For sequence asymmetric 2 × 2 nucleotide internal loops, models for estimating stabilities have incorporated averages of relevant sequence symmetric internal loops (*9*, *19*). The implicit assumption was that stacking between a closing base pair and a mismatch would be similar to that in sequence symmetric 2 × 2 loops. The NMR structure of $^{5'G\mathbf{UG}G3'}_{3'C\mathbf{UA}C5'}$ indicates that this is not a reasonable assumption. NMR spectra of 2 × 2 internal loops with tandem UU (*17*, *18*, *23*) or tandem AG pairs (*22*) reveal hydrogen bonding within the UU and AG pairs. This pattern contrasts with the lack of hydrogen bonds within the UU and AG pairs in the $^{5'G\mathbf{UG}G3'}_{3'C\mathbf{UA}C5'}$ loop. This suggests that models for estimating stabilities of sequence asymmetric 2 × 2 internal loops should be reconsidered.

Table 1 lists measured values of free energy increments at 37 °C for sequence asymmetric 2 × 2 internal loops closed on both sides by a GC pair. The values are based on the experiments of Xia et al. (*21*) and Burkard et al. (*74*) and calculated from equations equivalent to the following:

$$\Delta G^\circ_{37}\left(^{5'G\mathbf{UG}G3'}_{3'C\mathbf{UA}C5'}\right) = \Delta G^\circ_{37}\left(^{5'GAG\mathbf{UG}GAG3'}_{3'CUC\mathbf{UA}CUC5'}\right) - \Delta G^\circ_{37}\left(^{5'GAGGAG3'}_{3'CUCCUC5'}\right) + \Delta G^\circ_{37}\left(^{5'GG3'}_{3'CC5'}\right)$$

Here, $\Delta G^\circ_{37}\left(^{5'GAG\mathbf{UG}GAG3'}_{3'CUC\mathbf{UA}CUC5'}\right)$ and $\Delta G^\circ_{37}\left(^{5'GAGGAG3'}_{3'CUCCUC5'}\right)$ are measured free energy changes for duplex formation, and $\Delta G^\circ_{37}\left(^{5'GG3'}_{3'CC5'}\right)$ is the free energy increment for the $^{5'GG3'}_{3'CC5'}$ nearest neighbor interaction as determined by Xia et al. (*75*). The values in Table 1 suggest four categories and approximations for sequence asymmetric internal loops, as discussed below.

Category 1 (red in Table 1) contains loops unlikely to have a mismatch with two hydrogen bonds between bases, as observed for $^{5'G\mathbf{UG}G3'}_{3'C\mathbf{UA}C5'}$. This category includes loops with (a) a UU pair adjacent to a purine−purine pair as in $^{5'G\mathbf{UG}G3'}_{3'C\mathbf{UA}C5'}$, (b) a GA or AG pair adjacent to a pyrimidine−pyrimidine pair, or (c) adjacent AC, UC, CU, CC, CA, or AA pairs. Seventeen of the 18 such pairs in Table 1 have free energy increments at 37 °C between 0.5 and 2.2 kcal/mol, with an average of 1.1 ± 0.5 kcal/mol. The outlier is the $^{5'G\mathbf{CG}G3'}_{3'C\mathbf{CA}C5'}$ loop, which has a free energy increment of −0.6 kcal/mol. Ten sequence symmetric loops closed by two GC pairs and with the mismatches listed in criterion (c) have free energy increments between 0.9 and 2.0 kcal/mol, with an average of 1.3 ± 0.3 kcal/mol (*9*). This suggests that a value of 1.2 kcal/mol is a reasonable approximation for 2 × 2 loops of category 1 closed by GC pairs. The value of 1.2 kcal/mol is close to the average of 1.0 kcal/mol for 2 × 2 loops with UU pairs adjacent to GA, AG, or GG pairs, as listed in Table 1 (*21*, *74*). Unfortunately, these loops were not included in fitting the parameters currently included in the RNAstructure program for predicting secondary structure (*19*).

Category 2 (blue in Table 1) contains loops with two identically sized pairs likely to have two hydrogen bonds between the bases. These loops have any combination of GA and AG pairs or have two adjacent UU pairs. The four such loops in Table 1 are approximated well by the average of the relevant sequence symmetric loops. This comparison suggests that in these cases the structures of the sequence asymmetric loops may be similar to those of the sequence symmetric loops.

Category 3 (bolded black in Table 1) contains loops that are likely to have one pair with two hydrogen bonds between the bases. These loops have a GA or AG pair adjacent to an AC, CA, or AA pair, or have a UU pair adjacent to an AC, CA, or pyrimidine−pyrimidine pair. The assumption is that such combinations will not require large destabilizing contortions of the backbone. An NMR structure of the loop, $^{5'U\mathbf{AA}G3'}_{3'A\mathbf{AG}C5'}$, is consistent with this model (*50*). The 16 such loops in Table 1 have free energy increments at 37 °C between −1.0 and 0.7 kcal/mol with an average of 0.0 ± 0.5 kcal/mol.

Category 4 (underlined black in Table 1) contains loops with a single GG pair not adjacent to a UU pair. The 10 such loops in Table 1 have free energy increments at 37 °C between −0.6 and 0.2 kcal/mol with an average of −0.2 ± 0.3 kcal/mol.

Many of the sequence combinations listed in Table 1 have not been measured. Values predicted from the models described above are listed in a separate row in Table 1. Values for several loops listed in Table 1 have been measured by Bourdelat-Parks and Wartell (*20*). They report values for loops closed by two GC pairs and by one AU and one GC pair. Surprisingly, they see little dependence of stability on closing base pair, whereas experiments on a wide variety of other internal loops indicate that each substitution of a closing GC pair with an AU pair is expected to destabilize an internal loop by about 0.7 kcal/mol (*19*, *76*, *77*). One possible reason for this difference between measurement and expectation is that Bourdelat-Parks and Wartell included a GU pair adjacent to one of the two closing GC pairs but no GU pair in the duplexes with loops closed on one side with an AU pair. This may suggest a non-nearest neighbor effect that destabilizes the loops closed by two GC pairs. In general, the approximations in Table 1 adjusted by 0.7 kcal/mol for a single AU closing pair make reasonable predictions for the

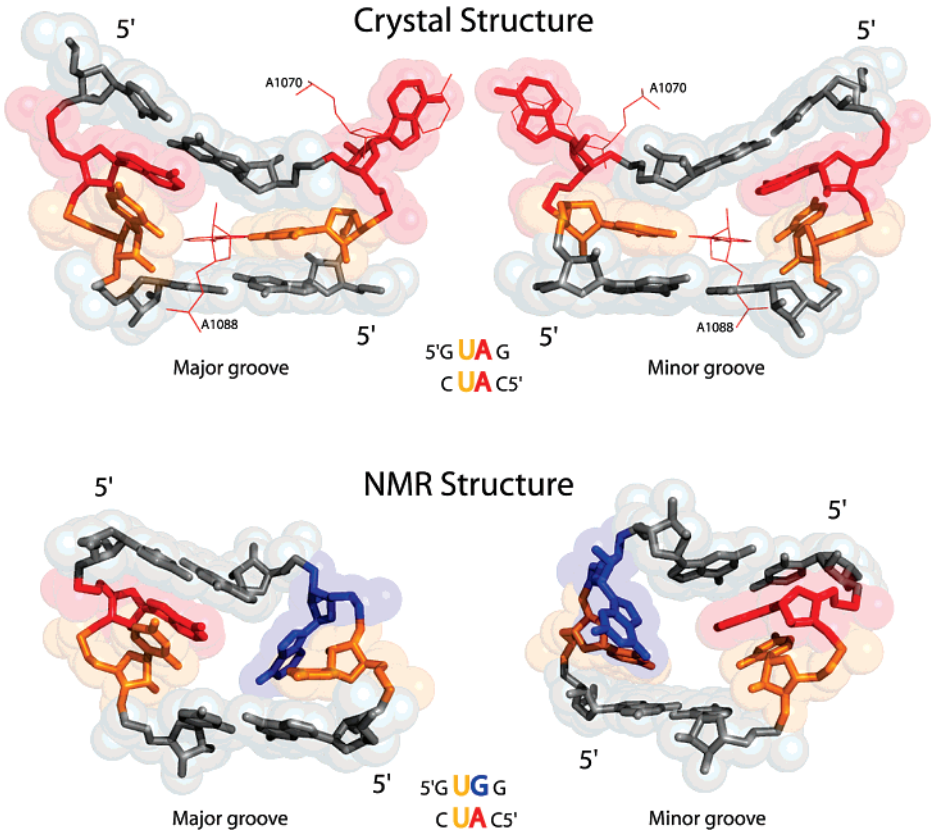



FIGURE 10: (Top) Major and minor groove view of the crystal structure of a ribosomal protein complex from *E. coli* ( $^{5'1059}$GUAG$^{1062}$3′, $^{3'}{}_{1079}$CUAC$_{1076}$5′, PDB entry 1QA6) (*36*). Residue A1070 (shown in red wireframe) from hairpin A stacks on the flipped out A1061. Residue A1088 (shown in red wireframe) from helix C forms a reverse Hoogsteen pair with U1060. (Bottom) Major and minor groove view of the average NMR structure ( $^{5'4}$ GUGG$^{7}$3′, $^{3'}{}_{19}$CUAC$_{16}$5′). The residues are color coded as in the internal loop sequence shown below.

Category 1 loops measured by Bourdelat-Parks and Wartell (*20*), with an average absolute difference of 0.5 kcal/mol between predicted and measured values. For the Category 3 loops, the average absolute differences between predicted and measured values are 0.8 and 1.6 kcal/mol, respectively, for loops closed with one and two GC pairs. A non-nearest neighbor effect from a GU pair could account for much of the difference observed for loops closed by two GC pairs. Evidently, there is still much to be discovered about the sequence dependence of duplexes containing internal loops.

The approximations described above for 2 × 2 internal loops were tested in the RNAstructure 4.5 program, where loops that have been optically melted are assigned the experimentally measured stability (*17*, *18*, *21*). The predicted lowest free energy structures were compared against a database of known RNA secondary structures (*9*). On average, the program predicted 71% of known canonical base pairs for domains of 700 nucleotides or less. This is lower than the 72.8% reported using the previous approximations for 2 × 2 loops (*19*). The difference probably reflects the fact that the parameters for multibranch loops were only optimized for the previous approximations.

## DISCUSSION

Reliable methods for the prediction of secondary and 3D structures of RNA would facilitate many applications of sequence information, including discovery of structure−function relationships and design of therapeutics. In principle, various free energy minimization methods can predict structure. Their success, however, is limited by current knowledge of the interactions determining RNA structure, especially in loops. Stabilities and structures of 2 × 2 nucleotide internal loops are very sequence dependent (*17*, *21*) and therefore provide good benchmarks for testing methods to predict local stability and structure.

*Structure of the 2 × 2 Internal Loop,* $^{5'}$GUGG3′ / 3′CUAC5′. The $^{5'}$GUGG3′ / 3′CUAC5′ internal loop studied here shows no hydrogen bonding within the middle two pairs (Figures 8 and 10), which is consistent with crystal structures of ribosomal components having small sequence variations of this loop, and bound to L11 protein (*36*, *37*, *43*). This is surprising because internal loops with tandem GA or UU pairs exhibit hydrogen bonding within the internal loop (*18*, *22*, *23*, *71*, *72*). Evidently, the 2 × 2 loop is partially preorganized in its functional structure. This contrasts with results for the structure of the $^{5'}$UAAG3′ / 3′AAGC5′ internal loop, which is part of the L20 protein binding site where the NMR structure of the isolated loop is very different from the structure in the ribosome (*50*).

The A of the $^{5'}$GUGG3′ / 3′CUAC5′ loop is stacked on the neighboring CG Watson−Crick pair. Because the $\Delta G°_{37}$ for a $^{5'}_{G}$CA stack is −1.7 kcal/mol (*78*), this is consistent with the observation that bases with a $\Delta G°_{37}$ for stacking more favorable than −0.7 kcal/mol are usually stacked on their adjacent base pair (*79*). G6 (Figure 2d) is very dynamic and is flipped out in

Table 1: Experimental and Predicted Free Energy Increments at 37 °C (kcal/mol) for Internal Loops with Sequence $^{5'G\mathbf{XY}G3'}_{3'C\mathbf{WZ}C5'}$ in 1M NaCl[a]

| X/W | Y/Z | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | G | U | G | A | U | C | C | C | A | |
| | G | A | U | G | C | C | U | C | A | A | |
| G | -1.7 | | 0.5 | | | | | 0.7 | 0.1 | -0.3 | Exp. |
| A | -1.7 | -1.5 | 1.2 | -0.2 | 0 | 1.2 | 1.2 | 1.2 | 0 | 0 | Pred. |
| A | -0.5 | -0.7 | 1.4 | -0.3 | 0.7 | | 0.5 | 0.7 | 0.5 | 0.4 | Exp. |
| G | -1.0 | -1.0 | 1.2 | -0.2 | 0 | 1.2 | 1.2 | 1.2 | 0 | 0 | Pred. |
| U | 0.9 | 0.9 | -0.6 | 1.4 | -0.2 | 0.2 | | -0.1 | 0.0 | 1.5 | Exp. |
| U | 1.2 | 1.2 | -0.5 | 1.2 | 0 | 0 | 0 | 0 | 0 | 1.2 | Pred. |
| G | -0.3 | | 0.7 | | | 0.2 | -0.6 | | -0.3 | -0.3 | Exp. |
| G | -0.2 | -0.2 | 1.2 | | -0.2 | -0.2 | -0.2 | -0.2 | -0.2 | -0.2 | Pred. |
| C | -0.4 | -0.7 | | -0.4 | | | | | | | Exp. |
| A | 0 | 0 | 0 | -0.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | Pred. |
| C | | | | -0.5 | | | | | | | Exp. |
| U | 1.2 | 1.2 | 0 | -0.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | Pred. |
| U | | | | 0.2 | | 1.7 | | | | | Exp. |
| C | 1.2 | 1.2 | 0 | -0.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | Pred. |
| C | | -0.6 | | | | | | | | | Exp. |
| C | 1.2 | 1.2 | 0 | -0.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | Pred. |
| A | 0.7 | -1.0 | 0.3 | | | | | 1.9 | 1.1 | | Exp. |
| C | 0 | 0 | 0 | -0.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | Pred. |
| A | -0.2 | 0.0 | 1.4 | -0.1 | | | | 2.2 | 0.6 | 1.3 | Exp. |
| A | 0 | 0 | 1.2 | -0.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | Pred. |

[a] Data from Xia et al., 1997 (*21*) are adjusted for Watson−Crick nearest neighbor parameters in Xia et al., 1998 (*75*). Values for loops with GG pairs are from Burkard et al., 2001 (*74*). Values in red are for category 1, loops unlikely to have two hydrogen-bonds between bases within a mismatch. They have (a) adjacent UU and RR pairs, (b) one GA or AG pair and one YY pair, or (c) adjacent AC, UC, CU, CC, CA, or AA pairs. Omitting −0.6 of $^{5'\mathbf{CG}3'}_{3'\mathbf{CA}5'}$, the average of 17 non-sequence symmetric loops is 1.1 ± 0.5 kcal/mol (range = 0.5 to 2.2 kcal/mol). The average for 10 sequence symmetric loops not likely to base pair with 2 base−base H-bonds is 1.3 ± 0.3 kcal/mol (range = 0.9 to 2.0) (*9*). The overall average for 27 loops = 1.2 ± 0.4 kcal/mol. The values in blue are for category 2, any combination of adjacent GA and AG pairs or adjacent UU pairs. The predicted $\Delta G^{\circ}_{\mathrm{LOOP}}$ is the average of relevant sequence symmetric loops as listed by Mathews et al. (*9*). Bold values not underlined are for category 3, loops without a GG pair and likely to have one mismatch with 2 base−base hydrogen bonds. They have a GA or AG pair adjacent to a CA, AC, or AA pair, or a UU pair adjacent to a YY or CA or AC pair. The average of 16 loops is 0.0 ± 0.5 kcal/mol (range = −1.0 to 0.7). The values underlined are for category 4, loops with a single GG pair not adjacent to a UU pair. The average of 10 cases is −0.2 ± 0.3 kcal/mol (range = −0.6 to 0.2).

most of the structures generated. The experimental restraints, however, do not conclusively position the G6 base in the minor or major groove.

*Internal Loop Is Partially Preorganized for Tertiary Interactions.* The internal loop, $^{5'4\ \mathbf{GUG}7'}_{3'_{19}\mathbf{CUAC}_{16}5'}$, has G6 at the position corresponding to 1061 in the crystals (Figure 2). The G6 base is bulged out similar to A1061, U1061, or G1165 in the crystal structures (*36*, *37*, *41*−*43*, *80*). Thus, G6 is somewhat preorganized (Figures 8 and 10).

In the crystal structures, A1088 is in a *syn* conformation and intercalates into the distorted major groove of the internal loop of helix A to form a Hoogsteen pair with U1060 (Figure 10). There is a reversal of chain direction at U1060, which is facilitated by bulging out of the base 1061. This requires A1077 and U1078 to be unpaired and turned out of the helix in order to open up the major groove. U1078 then forms a stack with A1077 and G1062 across the width of the helix, sharply changing the direction of the helix axis. The NMR structure of the internal loop $^{5'\mathbf{GUG}3'}_{3'\mathbf{CUAC}5'}$ has similar characteristics. The UU mismatch does not pair, A17 protrudes into the major groove, and U18 is partially stacked on A17 and turned out of the helix, resulting in an opened major groove. Therefore, an adenine (A1088) can easily intercalate into the distorted region to pair with U5. There is no reversal of the chain direction at U5, however. Evidently, the *S. acidocaldarius* internal loop is partially preorganized. This is somewhat surprising in that the free energy increment for this loop is unfavorable by 0.9 kcal/mol (Table 1) (*21*). This also contrasts with the *E. coli* loop, $^{5'^{1059}\mathbf{GUUG}^{1062}3'}_{3'_{1079}\mathbf{CUAC}_{1076}5'}$, where the isolated loop has a Watson−Crick AU pair and single hydrogen bond UU pair (*35*) that are broken in the complex with L11 (*37*).

*Lack of Hydrogen Bonding within the Internal Loop May Be Important for Tertiary Contacts.* In the crystal structures, a highly conserved A1070 base from hairpin A stacks on the flipped out U1061 base (equivalent to G6) (Figures 2a, 8, and 10). This is called a high five motif (*36*, *37*). The stacking of U1061 with A1070 is important for the global folding of the 58-nt fragment. A mutation U1061 → A in the internal loop region of helix A to give $^{5'\mathbf{GUAG}3'}_{3'\mathbf{CUAC}5'}$ (Figure

2a) stabilizes the structure of the mutant *E. coli* rRNA domain (*36*). One reason U1061A and U1061G mutants are more stable (*37*, *81*, *82*) is because stacking a purine ring on A1070 is energetically more favorable than U1061 stacking on A1070. This stacking interaction probably forces the bulged base into the minor groove, whereas the NMR data on the isolated loop allows it to move between the minor and major grooves.

The NMR structure of $^{5'G\mathbf{UG}G3'}_{3'C\mathbf{UA}C5'}$ suggests that a second reason that the U1061 → G mutation stabilizes tertiary structure is the partial preorganization of the loop, which contrasts with the necessity to break the U1061−A1077 base pair in *E. coli* (*82*). Interestingly, the U1061 → purine mutation is naturally present in some thermophilic Archaea and presumably provides higher stability (*82*) for those organisms in high-temperature environments.

*Implications for Tertiary Structure Prediction.* The structures of $^{5'G\mathbf{UN}G3'}_{3'C\mathbf{YA}C5'}$ loops from LSU rRNA show remarkable sequence and structure conservation (Figure 2) (*36*, *37*, *41*, *43*). The NMR structure of the isolated internal loop, $^{5'G\mathbf{UG}G3'}_{3'C\mathbf{UA}C5'}$, shows that there is no base pairing within the internal loop and that the structure is similar to the corresponding loop in crystals. Evidently, $^{5'G\mathbf{UN}G3'}_{3'C\mathbf{YA}C5'}$ internal loops are preorganized to form tertiary interactions by stacking of a tertiary purine on the N nucleotide and hydrogen bonding of a tertiary A with the U. This may facilitate the prediction of tertiary structures and contacts based only on sequence information. The variability in required sequence remains to be determined. The $\Delta G°_{LOOP}$ for $^{5'G\mathbf{AU}G3'}_{3'C\mathbf{GU}C5'}$ is similar to that of $^{5'G\mathbf{UG}G3'}_{3'C\mathbf{UA}C5'}$ (Table 1), and its 1D imino proton spectrum provides no evidence of hydrogen bonding in the mismatches. Thus, it has the potential to also be unstructured.

The fundamental reasons for pre-organization and lack of hydrogen bonds in the $^{5'G\mathbf{UG}G3'}_{3'C\mathbf{UA}C5'}$ loop are not understood, which presents a challenge to computational chemists. Presumably, there is a subtle balance among hydrogen bonding, stacking, solvation, and backbone distortion that determines the structure of an isolated internal loop. Success in parsing these interactions would allow prediction of motifs likely to be preorganized for tertiary interactions. Understanding such interactions may also facilitate the prediction of dynamics, which could allow for the prediction of loops that can reorganize for binding such as that observed for $^{5'U\mathbf{AA}G3'}_{3'A\mathbf{AG}C5'}$ (*50*).

## SUPPORTING INFORMATION AVAILABLE

NMR chemical shift assignments and comparison to Nuchemics predictions, NOE restraints, RDC restraints, dihedral restraints, RMSD values for residues in 10 accepted structures, 1D imino spectra at different temperatures and TOCSY and HSQC spectra. This material is available free of charge via the Internet at http://pubs.acs.org

## REFERENCES


1. Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena, *Cell 31*, 147−157.
2. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme, *Cell 35*, 849−857.
3. Hannon, G. J. (2002) RNA interference, *Nature 418*, 244−251.
4. Zamore, P. D., Tuschl, T., Sharp, P. A., and Bartel, D. P. (2000) RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals, *Cell 101*, 25−33.
5. Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R., and Tuschl, T. (2002) Single-stranded antisense siRNAs guide target RNA cleavage in RNAi, *Cell 110*, 563−574.
6. Blount, K. F., and Breaker, R. R. (2006) Riboswitches as antibacterial drug targets, *Nat. Biotechnol. 24*, 1558−1564.
7. Tinoco, I. Jr., Uhlenbeck, O. C., and Levine, M. D. (1971) Estimation of secondary structure in ribonucleic acids, *Nature 230*, 362−367.
8. Zuker, M., and Stiegler, P. (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information, *Nucleic Acids Res. 9*, 133−148.
9. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure, *J. Mol. Biol. 288*, 911−940.
10. Mathews, D. H., and Turner, D. H. (2006) Prediction of RNA secondary structure by free energy minimization, *Curr. Opin. Struct. Biol. 16*, 270−278.
11. Rivas, E., and Eddy, S. R. (1999) A dynamic programming algorithm for RNA structure prediction including pseudoknots, *J. Mol. Biol. 285*, 2053−2068.
12. Ding, Y., Chan, C. Y., and Lawrence, C. E. (2005) RNA secondary structure prediction by centroids in a Boltzmann weighted ensemble, *RNA 11*, 1157−1166.
13. Hofacker, I. L., Fekete, M., and Stadler, P. F. (2002) Secondary structure prediction for aligned RNA sequences, *J. Mol. Biol. 319*, 1059−1066.
14. Dirks, R. M., and Pierce, N. A. (2004) An algorithm for computing nucleic acid base-pairing probabilities including pseudoknots, *J. Comput. Chem. 25*, 1295−1304.
15. Ren, J., Rastegari, B., Condon, A., and Hoos, H. H. (2005) HotKnots: heuristic prediction of RNA secondary structures including pseudoknots, *RNA 11*, 1494−1504.
16. Xia, T., Mathews, D. H., and Turner, D. H. (1999) Thermodynamics of RNA Secondary Structure Formation, in *Prebiotic Chemistry, Molecular Fossils, Nucleotides, and RNA* (Soll, D., Moore, P. B., and Nishimura, S., Eds.) pp 21−47, Elsevier Science Ltd., Oxford, England.
17. Wu, M., McDowell, J. A., and Turner, D. H. (1995) A periodic table of symmetric tandem mismatches in RNA, *Biochemistry 34*, 3204−3211.
18. SantaLucia, J. Jr., Kierzek, R., and Turner, D. H. (1991) Stabilities of consecutive A·C, C·C, G·G, U·C, and U·U mismatches in RNA internal loops: Evidence for stable hydrogen bonded U·U and $C·C^+$ pairs, *Biochemistry 30*, 8242−8251.
19. Mathews, D. H., Disney, M. D., Childs, J. L., Schroeder, S. J., Zuker, M., and Turner, D. H. (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure, *Proc. Natl. Acad. Sci. U.S.A. 101*, 7287−7292.
20. Bourdelat-Parks, B. N., and Wartell, R. M. (2005) Thermodynamics of RNA duplexes with tandem mismatches containing a uracil-uracil pair flanked by C·G/G·C or G·C/A·U closing base pairs, *Biochemistry 44*, 16710−16717.
21. Xia, T., McDowell, J. A., and Turner, D. H. (1997) Thermodynamics of nonsymmetric tandem mismatches adjacent to G·C base pairs in RNA, *Biochemistry 36*, 12486−12497.
22. Wu, M., SantaLucia, J. Jr., and Turner, D. H. (1997) Solution structure of $(rGGCAGGCC)_2$ by two-dimensional NMR and the iterative relaxation matrix approach, *Biochemistry 36*, 4449−4460.
23. Nikonowicz, E. P., and Pardi, A. (1992) Three-dimensional heteronuclear NMR studies of RNA, *Nature 355*, 184−186.
24. Baeyens, K. J., De Bondt, H. L., and Holbrook, S. R. (1995) Structure of an RNA double helix including uracil-uracil base pairs in an internal loop, *Nat. Struct. Biol. 2*, 56−62.
25. Lietzke, S. E., Barnes, C. L., Berglund, J. A., and Kundrot, C. E. (1996) The structure of an RNA dodecamer shows how tandem U-U base pairs increase the range of stable RNA structures and the diversity of recognition sites, *Structure 4*, 917−930.
26. Walter, A. E., Wu, M., and Turner, D. H. (1994) The stability and structure of tandem GA mismatches in RNA depend on closing base pairs, *Biochemistry 33*, 11349−11354.

27. Schmidt, F. J., Thompson, J., Lee, K., Dijk, J., and Cundliffe, E. (1981) The binding site for ribosomal protein L11 within 23 S ribosomal RNA of *Escherichia coli*, *J. Biol. Chem. 256*, 12301−12305.
28. Ryan, P. C., Lu, M., and Draper, D. E. (1991) Recognition of the highly conserved GTPase center of 23S ribosomal RNA by ribosomal protein L11 and the antibiotic thiostrepton, *J. Mol. Biol. 221*, 1257−1268.
29. Cundliffe, E., and Thompson, J. (1981) Concerning the mode of action of micrococcin upon bacterial protein synthesis, *Eur. J. Biochem. 118*, 47−52.
30. Lu, M., and Draper, D. E. (1994) Bases defining an ammonium and magnesium ion-dependent tertiary structure within the large subunit ribosomal RNA, *J. Mol. Biol. 244*, 572−585.
31. Laing, L. G., Gluick, T. C., and Draper, D. E. (1994) Stabilization of RNA structure by Mg ions. Specific and non-specific effects, *J. Mol. Biol. 237*, 577−587.
32. Xing, Y., and Draper, D. E. (1996) Cooperative interactions of RNA and thiostrepton antibiotic with two domains of ribosomal protein L11, *Biochemistry 35*, 1581−1588.
33. Xing, Y., and Draper, D. E. (1995) Stabilization of a ribosomal RNA tertiary structure by ribosomal protein L11, *J. Mol. Biol. 249*, 319−331.
34. Stark, M. J., Cundliffe, E., Dijk, J., and Stoffler, G. (1980) Functional homology between *E. coli* ribosomal protein L11 and *B. megaterium* protein BM-L11, *Mol. Gen. Genet. 180*, 11−15.
35. Wang, Y. X., Huang, S., and Draper, D. E. (1996) Structure of a U·U pair within a conserved ribosomal RNA hairpin, *Nucleic Acids Res. 24*, 2666−2672.
36. Conn, G. L., Draper, D. E., Lattman, E. E., and Gittis, A. G. (1999) Crystal structure of a conserved ribosomal protein-RNA complex, *Science 284*, 1171−1174.
37. Wimberly, B. T., Guymon, R., McCutcheon, J. P., White, S. W., and Ramakrishnan, V. (1999) A detailed view of a ribosomal active site: the structure of the L11-RNA complex, *Cell 97*, 491−502.
38. Xing, Y., Guha Thakurta, D., and Draper, D. E. (1997) The RNA binding domain of ribosomal protein L11 is structurally similar to homeodomains, *Nat. Struct. Biol. 4*, 24−27.
39. Hinck, A. P., Markus, M. A., Huang, S., Grzesiek, S., Kustonovich, I., Draper, D. E., and Torchia, D. A. (1997) The RNA binding domain of ribosomal protein L11: three-dimensional structure of the RNA-bound form of the protein and its interaction with 23S rRNA, *J. Mol. Biol. 274*, 101−113.
40. Markus, M. A., Hinck, A. P., Huang, S., Draper, D. E., and Torchia, D. A. (1997) High resolution solution structure of ribosomal protein L11-C76, a helical protein with a flexible loop that becomes structured upon binding to RNA, *Nat. Struct. Biol. 4*, 70−77.
41. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 angstrom resolution, *Science 289*, 905−920.
42. Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F., and Yonath, A. (2001) High resolution structure of the large ribosomal subunit from a mesophilic Eubacterium, *Cell 107*, 679−688.
43. Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Cate, J. H. (2005) Structures of the bacterial ribosome at 3.5 Å resolution, *Science 310*, 827−834.
44. Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y., Feng, B., Lin, N., Madabusi, L. V., Muller, K. M., Pande, N., Shang, Z., Yu, N., and Gutell, R. R. (2002) The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs, *BMC Bioinformatics 3*, 2.
45. Heus, H. A., and Pardi, A. (1991) Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops, *Science 253*, 191−194.
46. Jucker, F. M., Heus, H. A., Yip, P. F., Moors, E. H. M., and Pardi, A. (1996) A network of heterogeneous hydrogen bonds in GNRA tetraloops, *J. Mol. Biol. 264*, 968−980.
47. Wincott, F., DiRenzo, A., Shaffer, C., Grimm, S., Tracz, D., Workman, C., Sweedler, D., Gonzalez, C., Scaringe, S., and Usman, N. (1995) Synthesis, deprotection, analysis and purification of RNA and ribozymes, *Nucleic Acids Res. 23*, 2677−2684.
48. Usman, N., Ogilvie, K. K., Jiang, M. Y., and Cedergren, R. J. (1987) Automated chemical synthesis of long oligoribonucleotides using 2′-O-silylated ribonucleoside 3′-O-phosphoramidites on a controlled-pore glass support: synthesis of a 43-nucleotide sequence similar to the 3′-half molecule of an *Escherichia coli* formylmethionine tRNA, *J. Am. Chem. Soc. 109*, 7845−7854.
49. Stawinski, J., Stromberg, R., Thelin, M., and Westman, E. (1988) Evaluation of the use of the tert-butyldimethylsilyl group for 2′-protection in RNA: Synthesis via the H-phosphonate approach, *Nucleosides Nucleotides 7*, 779−782.
50. Shankar, N., Kennedy, S. D., Chen, G., Krugh, T. R., and Turner, D. H. (2006) The NMR structure of an internal loop from 23S ribosomal RNA differs from its structure in crystals of 50S ribosomal subunits, *Biochemistry 45*, 11776−11789.
51. Ruckert, M., and Otting, G. (2000) Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments, *J. Am. Chem. Soc. 122*, 7793−7797.
52. Johnson, B. A., and Blevins, R. A. (1994) NMR View - A computer-program for the visualization and analysis of NMR data, *J. Biomol. NMR 4*, 603−614.
53. Goddard, T. D., and Kneller, D. G. (2004) *Sparky*, NMR Assignment and Integration Software. http://www.cgl.ucsf.edu/home/sparky
54. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr., Sect. D 54*, 905−921.
55. Warren, J. J., and Moore, P. B. (2001) A maximum likelihood method for determining D(a)(PQ) and R for sets of dipolar coupling data, *J. Magn. Reson. 149*, 271−275.
56. Lukavsky, P. J., Kim, I., Otto, G. A., and Puglisi, J. D. (2003) Structure of HCV IRES domain II determined by NMR, *Nat. Struct. Biol. 10*, 1033−1038.
57. Zweckstetter, M., Hummer, G., and Bax, A. (2004) Prediction of charge-induced molecular alignment of biomolecules dissolved in dilute liquid-crystalline phases, *Biophys. J. 86*, 3444−3460.
58. Zweckstetter, M., and Bax, A. (2000) Prediction of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR, *J. Am. Chem. Soc. 122*, 3791−3792.
59. DeLano, W. L. (2002) The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos CA.
60. Borer, P. N. (1975) Optical Properties of Nucleic Acids, Absorption and Circular Dichroism Spectra, in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D. Ed.) 3rd ed., pp 589−595, CRC Press, Cleveland, OH.
61. Richards, E. G. (1975) Use of Tables in Calculation of Absorption, Optical Rotatory Dispersion and Circular Dichroism of Polyribonucleotides, in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D. Ed.) 3rd ed., pp 596−603, CRC Press, Cleveland, OH.
62. Gutell, R. R., Schnare, M. N., and Gray, M. W. (1992) A compilation of large subunit (23S- and 23S-like) ribosomal RNA structures, *Nucleic Acids Res. 20*, 2095−2109.
63. Gutell, R. R., Gray, M. W., and Schnare, M. N. (1993) A compilation of large subunit (23S- and 23S-like) ribosomal RNA structures: 1993, *Nucleic Acids Res. 21*, 3055−3074.
64. Smallcombe, S. H. (1993) Solvent suppression with symmetrically-shifted pulses, *J. Am. Chem. Soc. 115*, 4776−4785.
65. Varani, G., Aboulela, F., and Allain, F. H. T. (1996) NMR investigation of RNA structure, *Prog. Nucl. Magn. Reson. Spectrosc. 29*, 51−127.
66. Varani, G., and Tinoco, I. (1991) RNA structure and NMR spectroscopy, *Q. Rev. Biophys. 24*, 479−532.
67. Cromsigt, J. A., Hilbers, C. W., and Wijmenga, S. S. (2001) Prediction of proton chemical shifts in RNA. Their use in structure refinement and validation, *J. Biomol. NMR 21*, 11−29.
68. SantaLucia, J., Kierzek, R., and Turner, D. H. (1991) Functional-group substitutions as probes of hydrogen-bonding between GA mismatches in RNA internal loops, *J. Am. Chem. Soc. 113*, 4313−4322.
69. SantaLucia, J., Kierzek, R., and Turner, D. H. (1992) Context dependence of hydrogen-bond free-energy revealed by substitutions in an RNA hairpin, *Science 256*, 217−219.
70. Gorenstein, D. (1984) *31P NMR, Principles and Applications*, Academic Press, New York.
71. SantaLucia, J., Jr., and Turner, D. H. (1993) Structure of $(rGGCGAGCC)_2$ in solution from NMR and restrained molecular dynamics, *Biochemistry 32*, 12612−12623.

72. Wu, M., and Turner, D. H. (1996) Solution structure of (rGCG-GACGC)$_2$ by two-dimensional NMR and the iterative relaxation matrix approach, *Biochemistry 35*, 9677−9689.
73. Gautheret, D. F., Konings, D., and Gutell, R. R. (1994) A major family of motifs involving G·A mismatches in ribosomal RNA, *J. Mol. Biol. 242*, 1−8.
74. Burkard, M. E., Xia, T., and Turner, D. H. (2001) Thermodynamics of RNA internal loops with a guanosine-guanosine pair adjacent to another noncanonical pair, *Biochemistry 40*, 2478−2483.
75. Xia, T., SantaLucia, J., Jr., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs, *Biochemistry 37*, 14719−14735.
76. Schroeder, S. J., and Turner, D. H. (2000) Factors affecting the thermodynamic stability of small asymmetric internal loops in RNA, *Biochemistry 39*, 9257−9274.
77. Schroeder, S. J., Burkard, M. E., and Turner, D. H. (1999) The energetics of small internal loops in RNA, *Biopolymers 52*, 157−167.
78. Freier, S. M., Burger, B. J., Alkema, D., Neilson, T., and Turner, D. H. (1983) Effects of 3′ dangling end stacking on the stability of GGCC and CCGG double helixes, *Biochemistry 22*, 6198−6206.
79. Burkard, M. E., Kierzek, R., and Turner, D. H. (1999) Thermodynamics of unpaired terminal nucleotides on short RNA helixes correlates with stacking at helix termini in larger RNAs, *J. Mol. Biol. 290*, 967−982.
80. Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H., and Noller, H. F. (2001) Crystal structure of the ribosome at 5.5 A resolution, *Science 292*, 883−896.
81. Lu, M., and Draper, D. E. (1995) On the role of rRNA tertiary structure in recognition of ribosomal protein L11 and thiostrepton, *Nucleic Acids Res. 23*, 3426−3433.
82. Lu, M., and Draper, D. E. (1994) Bases defining an ammonium and magnesium ion-dependent tertiary structure within the large subunit ribosomal RNA, *J. Mol. Biol. 244*, 572−585.

BI700802S