The NMR Structure of an Internal Loop from 23S Ribosomal RNA Differs from Its Structure in Crystals of 50S Ribosomal Subunits[†,‡]

Neelaabh Shankar,[§] Scott D. Kennedy,[§] Gang Chen,[||] Thomas R. Krugh,[||] 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, and Department of Chemistry, University of Rochester, Rochester, New York 14627-0216

Received March 23, 2006; Revised Manuscript Received June 26, 2006

ABSTRACT: Internal loops play an important role in structure and folding of RNA and in recognition of RNA by other molecules such as proteins and ligands. An understanding of internal loops with propensities to form a particular structure will help predict RNA structure, recognition, and function. The structures of internal loops $\begin{matrix}5'^{1009}\text{CUAAG}^{1013}3' \\ 3'_{1168}\text{GAAGC}_{1164}5'\end{matrix}$ and $\begin{matrix}5'^{998}\text{CUAAG}^{1002}3' \\ 3'_{1157}\text{GAAGC}_{1153}5'\end{matrix}$ from helix 40 of the large subunit rRNA in *Deinococcus radiodurans* and *Escherichia coli*, respectively, are phylogenetically conserved, suggesting functional relevance. The energetics and NMR solution structure of the loop were determined in the duplex $\begin{matrix}5'^{1}\text{GGC}\mathbf{UAA}\text{GAC}^{9}3' \\ 3'_{18}\text{CCG}\mathbf{AAG}\text{CUG}_{10}5'\end{matrix}$. The internal loop forms a different structure in solution and in the crystal structures of the ribosomal subunits. In particular, the crystal structures have a bulged out adenine at the equivalent of position A15 and a reverse Hoogsteen UA pair (trans Watson−Crick/Hoogsteen UA) at the equivalent of U4 and A14, whereas the solution structure has a single hydrogen bond UA pair (cis Watson−Crick/sugar edge A15U4) between U4 and A15 and a sheared AA pair (trans Hoogsteen/sugar edge A14A5) between A5 and A14. There is cross-strand stacking between A6 and A14 (A6/A14/A15 stacking pattern) in the NMR structure. All three structures have a sheared GA pair (trans Hoogsteen/sugar edge A6G13) at the equivalent of A6 and G13. The internal loop has contacts with ribosomal protein L20 and other parts of the RNA in the crystal structures. These contacts presumably provide the free energy to rearrange the base pairing in the loop. Evidently, molecular recognition of this internal loop involves induced fit binding, which could confer several advantages. The predicted thermodynamic stability of the loop agrees with the experimental value, even though the thermodynamic model assumes a Watson−Crick UA pair.

Internal loops are important words in the language of RNA. They have implications for folding and stability and often form tertiary and quaternary interactions. An understanding of internal loop structure and energetics in the presence and absence of contacts is important for understanding fundamental aspects of RNA folding and molecular recognition. Predictions of RNA secondary structure often rely on free energy minimization algorithms that assume relatively simple structural models for estimating energetics (*1*−*3*). Crystal structures of ribosomal subunits from *Deinococcus radiodurans* (*4*), *Escherichia coli* (*5*), *Haloarcula marismortui* (*6*), and *Thermus thermophilus* (*7*, *8*) provide insight into whether these structural models reflect the final structure of the RNA. Differences between models and crystal structures may reveal structural rearrangements that are important for function or suggest revisions in structural models.

Here we report the solution structure of a duplex, $\begin{matrix}5'^{1}\text{GGCUAAGAC}^{9}3' \\ 3'_{18}\text{CCGAAGCUG}_{10}5'\end{matrix}$, that contains an internal loop $\begin{matrix}5'^{1009}\text{CUAAG}^{1013}3' \\ 3'_{1168}\text{GAAGC}_{1164}5'\end{matrix}$ derived from domain II of helix 40, *E. coli* helix numbering (*9*), of the *D. radiodurans* large ribosomal subunit (LSU).[1] At an equivalent position, the crystal structure of *E. coli* LSU contains an internal loop with the same sequence, $\begin{matrix}5'^{998}\text{CUAAG}^{1002}3' \\ 3'_{1157}\text{GAAGC}_{1153}5'\end{matrix}$ (*5*). This internal loop is highly conserved in the secondary structures of large ribosomal subunits (Figure 1) and is a part of the binding site of the L20 ribosomal protein (*10*). Other structurally similar internal loops are present in the *D. radiodurans*, *E. coli*, and *H. marismortui* LSUs (Figure 1) (*6*). They adopt similar structure despite different locations and tertiary interactions. In *D. radiodurans*, internal loop residues A1011 and A1012 engage in ribose zipper interactions, and this is also seen for equivalent residues in *H. marismortui* (*11*, *12*). In the *D. radiodurans* crystal, the loop

[†] This work was supported by NIH Grant GM22939 (D.H.T.).

[‡] The coordinates have been deposited as Protein Data Bank entry 2H49.

* 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.

[||] Department of Chemistry.

[⊥] Center for Pediatric Biomedical Research and Department of Pediatrics.

[1] Abbreviations: C_T, total concentration of all strands of oligonucleotides in solution; LSU, large subunit; N, any nucleotide, i.e., A, C, G, or U; Y, any pyrimidine, i.e., U or C; R, any purine, i.e., G or A; T_M, melting temperature in kelvin; T_m, melting temperature in degrees Celsius.

10.1021/bi0605787 CCC: $33.50 © 2006 American Chemical Society
Published on Web 09/09/2006

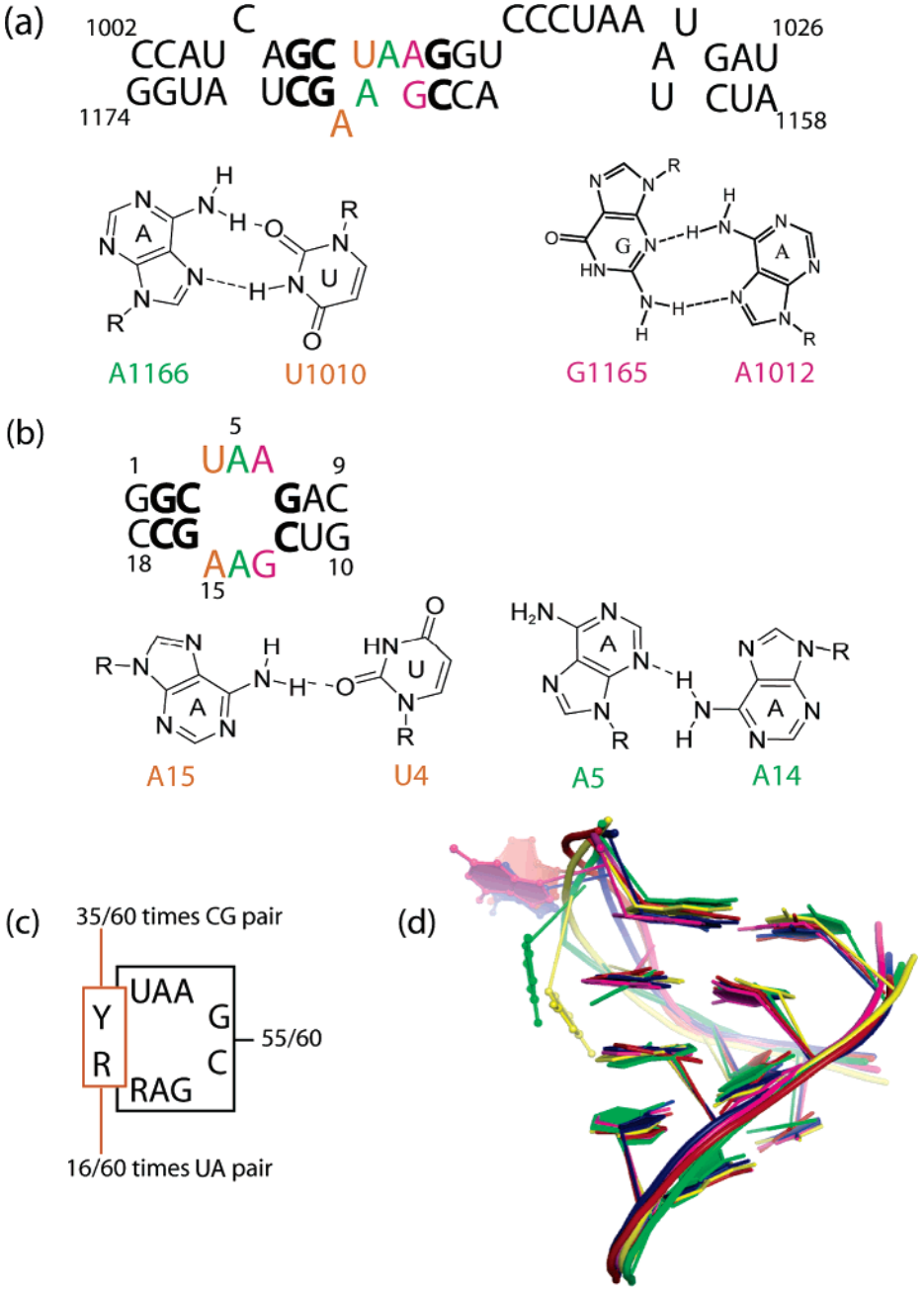

FIGURE 1: (a) Secondary structure of helix 40 in domain II of the large subunit ribosomal RNA of *D. radiodurans*. In the crystal structure, U1010 forms a reverse Hoogsteen pair with A1166 and A1011 is stacked on A1012 and forms a cross-strand stack with A1166. A1012 forms a sheared pair with G1165. A1167 is flipped out of the helix (*4*). (b) Duplex studied by NMR. Boldface and colored nucleotides are identical to those of the natural sequence. Secondary structure prediction of the internal loop predicts it to be a 2 × 2 nucleotide internal loop with U4 forming a Watson−Crick pair with A15. The NMR structure reveals a cis Watson−Crick/sugar edge UA pair with a single hydrogen bond. (c) Conserved features of this internal loop. Of the 60 secondary structures in R. Gutell's secondary structure database [http://www.rna.icmb.utexas.edu/ (*37*)], 55 have the sequence in the box with R being a purine, 35 of 60 times YR is a CG pair, and 16 of 60 times YR is a UA pair. (d) Overlap of all heavy atoms for the five similar internal loops observed in *D. radiodurans* ($^{5'1009}_{3'1168}$CUAAG$^{1013}_{}$GAAGC$_{1164}^{}$5′ colored red and $^{5'1364}_{3'1393}$CUAAG$^{1368}_{}$GUAGC$_{1389}^{}$5′ colored green), *E. coli* ($^{5'998}_{3'1157}$CUAAG$^{1002}_{}$GAAGC$_{1153}^{}$5′ colored blue and $^{5'1351}_{3'1380}$CUAAG$^{1355}_{}$GUAGC$_{1376}^{}$5′ colored yellow), and *H. marismortui* ($^{5'1095}_{3'1261}$UUAAG$^{1099}_{}$AGAGC$_{1257}^{}$5′ colored magenta). The bulged bases are shown in ball-and-stick format. Green and yellow bulged bases are uracils.

has contacts with the hairpin between helices 39 and 40 of the ribosomal RNA and with the L20 protein. L20 is one of nine core proteins (*13*, *14*) that bind to 23S rRNA during an early stage of ribosomal assembly in *E. coli* (*15*). The C-terminal domain of L20 represses the translation of its own mRNA, and it has been suggested that its binding site on mRNA mimics the binding site on the rRNA (*10*, *16*). The residues at the interface of the protein with helix 40 of 23S rRNA are conserved and involved in recognition of rRNA (*17*).

Structure prediction programs such as RNAstructure 4.11 (*1*, *3*) predict a 2 × 2 nucleotide internal loop with a UA closing pair, but in the crystal structures of the large ribosomal subunits of *D. radiodurans* (*4*) and *E. coli* (*5*),

the adenine corresponding to A15 is flipped out from the helix and the adenine corresponding to A14 forms a reverse Hoogsteen pair with U4 (Figure 1a). In contrast, the NMR structure has a single hydrogen-bonded UA pair between U4 and A15 and a sheared AA pair between A5 and A14 (Figure 1b). A sheared GA pair between G13 and A6 in the NMR structure is consistent with the crystal structures. Except for the non-Watson−Crick UA pair, the solution structure agrees with the model used to predict thermodynamic stability. The solution structure suggests that the internal loop undergoes a conformational change when it interacts with ribosomal proteins and rRNA. This internal loop thus exhibits an induced fit (*18*) rather than a lock and key mechanism (*19*) of binding which may be important for function.

MATERIALS AND METHODS

Oligoribonucleotide Synthesis and Purification. Oligonucleotides were synthesized on an Applied Biosystems 392 DNA/RNA synthesizer using standard phosphoramidite chemistry (*20*, *21*). The phosphoramidites and the CPG support were bought from Glen Research. For each 1 μmol synthesis, the CPG support and the base-protecting groups were removed by incubation in 2 mL of an ammonia/ethanol solution (3:1, v/v) at 55 °C for 12 h (*22*). The solid support was removed by filtration, and the filtrate was lyophilized followed by removal of silyl protecting groups on the 2′-hydroxyls by incubation in a 9:1 (v/v) TEA-3HF (triethylamine trihydrofluoride)/DMF (*N*,*N*-dimethylformamide) mixture at 55 °C for ∼2 h. RNA was precipitated with 2-butanol, and the sample was centrifuged at 12 000 rpm. The pellet was washed with 70% ethanol and left to dry at room temperature. The pellet was redissolved in 2 mL of 5 mM ammonium bicarbonate at pH 7, and the sample was loaded onto a Waters Sep-Pak C-18 reverse phase column to remove excess salts. The eluted sample was then applied to a preparative Baker Si500F silica gel plate (20 cm × 20 cm, 500 μm thick). The plates were kept in a sealed TLC chamber with a running solution of 1-propanol, ammonia, and water in the ratio of 50:40:10 (v/v/v). The slowest-running band on the plate was identified with UV light and scraped off. RNase free water was used to dissolve the RNA, and then a Waters Sep-Pak C18 reverse phase column was used to remove salts. The mass of the RNA was verified by ESI-MS with a Hewlett-Packard 1100 LC/MS Chemstation, and the purity was verified by analytical reverse phase HPLC. All the samples were >95% pure.

NMR Sample Preparation. Sample preparation was similar to that of Chen et al. (*23*) with minor modifications. An equal number of moles of each strand was mixed in 300 μL of RNase free water and then dialyzed overnight against 1 L of filtered autoclaved water in a Gibco Life Technologies microdialysis system with a 1000 molecular weight cutoff Spectro-por dialysis membrane and a Rainin Dynamax peristaltic pump. 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 (pH 6.5)]. The sample was dried and reconstituted with 10% D_2O in water to provide a lock signal for the exchangeable proton spectrum. For nonexchangeable proton spectra, D_2O exchange was achieved with three repetitions of lyophilization with 99.96% D_2O as the solvent, and the sample was finally dissolved in 300 μL of 99.96% D_2O from Cambridge Isotope Laboratories. The total duplex concentration was approximately 2 mM. The sample was placed in Shigemi tubes for collection of the spectra. Later, to observe the effect of Mg^{2+} on the structure, 10 mM $MgCl_2$ was added to the NMR sample [80 mM NaCl, 10 mM sodium phosphate, 0.5 mM Na_2EDTA, and 10 mM $MgCl_2$ (pH 6.5)].

NMR Spectroscopy. NMR spectra were collected on Varian Inova 500 and 600 MHz spectrometers. One-dimensional and two-dimensional imino proton spectra at different temperatures were acquired with an S-shaped excitation pulse (*24*) with a spectral width of 12 kHz. SNOESY spectra were recorded in a 90:10 (v/v) H_2O/D_2O mixture at 15 and 25 °C with mixing times of 100 and 150 ms. Each spectrum was acquired with 256 t_1 increments with 2K complex points and 70 scans per FID and a recycle delay of 1.2 s.

NOESY spectra of the sample in D_2O were acquired at 15, 25, and 30 °C with mixing times of 100, 200, 400, and 600 ms. FIDs were acquired with a spectral width of 4200 Hz and 2K complex points. A total of 360 FIDs were obtained with 32 scans per FID and a recycle delay of 2.8 s. TOCSY spectra were collected with mixing times of 8, 18, 30, and 100 ms at 15, 20, and 30 °C. These spectra were acquired with 2K complex points and a spectral width of 4000 Hz. A total of 256−512 FIDs were acquired with 24−64 scans per FID. Known temperature-dependent chemical shifts of H_2O or HDO relative to 3-(trimethylsilyl)tetradeuteriosodium propionate (TSP) were used to reference the proton spectra. A natural abundance $^1H-^{13}C$ HSQC spectrum at 25 °C was acquired on a Varian Inova 600 MHz spectrometer with a spectral width of 6000 Hz for the ^1H dimension and a spectral width of 16 600 Hz for the ^{13}C dimension. The spectrum has 96 increments and was collected with 1K complex points with 672 scans per FID. Natural abundance $^1H-^{13}C$ HMQC spectra were acquired at 25 and 30 °C with a spectral width of 5000 Hz for the ^1H dimension and a spectral width of 14 000 Hz for the ^{13}C dimension. The HMQC spectra consisted of 60 increments each having 1K complex points and 544 scans per FID. T_1 relaxation rates for the base protons were measured by the inversion recovery method. The $^1H-^{31}P$ HETCOR spectrum was acquired with a spectral width of 1400 Hz for ^1H and a spectral width of 2000 Hz for ^{31}P. Two-dimensional spectra were processed using NMRpipe (*25*). Similar one-dimensional imino proton spectra at different temperatures, SNOESY spectra at 15 °C with a mixing time of 150 ms, and NOESY spectra (in D_2O) with mixing times of 100 and 200 ms at 25 and 30 °C were collected after addition of 10 mM $MgCl_2$.

Restraint Generation. NOESY cross-peak volumes were obtained with Sparky (*26*) using box integration. Distance restraints were generated for the duplex from 100 and 200 ms mixing time NOESY spectra at 25 °C and a 150 ms SNOESY spectrum at 15 °C. The Sparky output was converted to Discover input by using a C++ program, which converted the volume for each peak to a distance restraint using $1/r^6$ scaling and a two-spin approximation. Average H5−H6 peak volumes were used as a reference with a distance of 2.45 Å.

NOE-derived distances were assigned error limits of ±30% (100 ms mixing time) and ±40% (200 ms mixing time and SNOESY spectra) to allow for relaxation, spin diffusion, baseline distortions, water exchange, and noise. Restraints

were estimated or discarded in the case of extreme overlaps. No restraints were used from H5′ and H5″ protons. A total of 203 interproton distance restraints were used with 106 internucleotide and 97 intranucleotide restraints (Table S2 of the Supporting Information). NMR data are consistent with the formation of Watson−Crick pairs. Thus, 17 hydrogen bond distance restraints (1.8−2.5 Å between the hydrogen and acceptor) were used for the six Watson−Crick pairs of the stem. Backbone dihedral angles for the Watson−Crick stem residues were loosely restrained: $\alpha = 0 \pm 120°$, $\beta = 180 \pm 30°$, $\gamma = 60 \pm 30°$, $\delta = 85 \pm 30°$, $\epsilon = -140 \pm 40°$, $\zeta = 0 \pm 120°$, and $\chi = -170 \pm 40°$. No hydrogen bonding or dihedral angle restraints were used for the loop nucleotides, except for δ and χ. TOCSY and DQFCOSY spectra were used to identify residues in a *C2′-endo* sugar pucker. A5, terminal residues (C9 and C18), and G13 exhibited H1′−H2′ cross-peaks in DQFCOSY spectra and therefore were restrained to cover both the *C3′*- and *C2′-endo* conformation with δ (122.5 ± 67.5°). The remaining loop residues (U4, A6, A14, and A15) were restrained to be in a *C3′-endo* conformation, δ (85 ± 30°). The glycosidic torsion angles were restrained in the anti conformation, χ (−120 ± 90°), for all loop residues since the intensity of the intranucleotide cross-peak between H1′ and H8/H6 was not comparable to that of the H5−H6 cross-peak. The Supporting Information includes tables of distance, dihedral angle, and endocyclic sugar torsion angle restraints.

Simulated Annealing. The Discover 98 software package running on a Silicon Graphics Octane computer was used to perform NMR-restrained molecular dynamics and energy minimization. A standard A-form starting structure was generated using Biosym Insight II. Simulations were done with the AMBER 95 force field (*27*) using flat-bottom restraint potentials, with force constants of 25 kcal mol^{-1} Å^{-2} for distance restraints and 50 kcal mol^{-1} rad^{-2} for torsion angle restraints and a maximum force of 1000 kcal/mol. Simulations were done in vacuo with the NMR-derived restraints. For van der Waals interactions, group-based summation was used with an 18 Å cutoff. For the electrostatic interactions, the cell-multipole method (*28*) was used with a distance-dependent dielectric constant ($\epsilon = 2r$, where r is the distance in angstroms). There were a total of 14 steps in the simulations (*29*, *30*): (1) van der Waals and electrostatic energy scaled to 1 and 0%, respectively, and dihedral, NOE, and covalent bond energy scaled to 100%; (2) steepest descent minimization started with 500 iterations; (3) restrained molecular dynamics (rMD) at 1000 K for 4 ps with 1 fs time steps; (4) 2 ps rMD at 900 K; (5) 2 ps rMD at 800 K; (6) van der Waals and electrostatic energy increased to 33%; (7) 2 ps rMD at 700 K; (8) van der Waals and electrostatic energy increased to 67%; (9) 2 ps rMD at 600 K; (10) van der Waals and electrostatic energy increased to 100%; (11) 2 ps rMD at 500 K; (12) 2 ps rMD at 400 K; (13) 2 ps rMD at 300 K; and (14) 40 000 iterations of conjugate energy minimization. Steps 3−5 effectively randomize the starting structure. A total of 50 structures were generated and used for analysis.

UV Melting Experiments and Thermodynamics. Oligonucleotides were lyophilized and dissolved in 1.0 M NaCl, 20 mM sodium cacodylate, and 0.5 mM Na_2EDTA (pH 6.5) (melt buffer) or in 100 mM KCl, 10 mM $MgCl_2$, and 20 mM sodium cacodylate (pH 6.5). The concentration of single-stranded RNA was calculated from the absorbance at 260 nm and 80 °C prior to it being dissolved in melt buffer. Extinction coefficients were predicted from those of dinucleotide monophosphates and nucleosides (*31*, *32*). Small mixing errors for noncomplementary duplexes do not affect the thermodynamics appreciably (*33*). Absorbance versus temperature melting curves for the duplex were acquired at 260 nm at heating rates of 1 °C/min with a Beckman Coulter DU640C spectrophotometer having a Peltier temperature controller cooled with flowing water.

Melting curves were fit to a two-state model with Meltwin, assuming linear sloping baselines and temperature-independent $\Delta H°$ and $\Delta S°$ values (*2*, *34*, *35*). The melting temperatures of the duplex in kelvin, T_M, at different concentrations, where C_T is the total strand concentration, were used to calculate the thermodynamic parameters from (*36*)

$$T_M^{-1} = (R/\Delta H°) \ln(C_T/4) + \Delta S°/\Delta H° \qquad (1)$$

The equation $\Delta G°_{37} = \Delta H° - 310.15\Delta S°$ was used to calculate the free energy change at 37 °C (310.15 K). The $\Delta H°$ values were also obtained from averaging fits of individual melting curves and were within 6% of the values derived from the T_M^{-1} versus $\ln(C_T/4)$ plots, suggesting that a two-state model is a good approximation for this duplex.

RESULTS

Sequence and Structure Conservation. A survey of secondary structures of 60 large ribosomal subunits (*37*) showed that 55 of 60 structures had the conserved sequence $\begin{matrix}5'\text{YUAAG}3'\\3'\text{RRAGC}5'\end{matrix}$, where Y represents a pyrimidine, either U or C, and R represents a purine, either G or A (Figure 1c). YR is a CG pair in 35 of 60 cases and a UA in 16 of 60 cases. The crystal structures of the *D. radiodurans* (PDB entry 1NKW), *E. coli* (PDB entry 2AWB), and *H. marismortui* (PDB entry 1JJ2) large subunit rRNAs show that this conserved internal loop adopts a very similar three-dimensional structure (*4*−*6*) (Figure 1d).

In the *D. radiodurans* internal loop, $\begin{matrix}5'^{1009}\text{CUAAG}^{1013}3'\\3'_{1168}\text{GAAGC}_{1164}5'\end{matrix}$, A1167 is flipped out of the helix. U1010 forms a reverse Hoogsteen pair with A1166 (Figure 1a). A1011 is stacked between A1166 and A1012, making a cross-strand as well as an intrastrand stacking pattern. A1012 makes a sheared GA pair with G1165. All the internal loops in Figure 1d have similar three-dimensional structural features, which suggests functional relevance.

The structural similarity between these internal loops from different organisms is striking especially since each is in a different context of tertiary and quaternary contacts (*4*−*6*). The structural similarity motivated the NMR studies presented below. The duplex $\begin{matrix}5'^{1}\text{GGC}\mathbf{UAA}\text{GAC}^{9}3'\\3'_{18}\text{CCG}\mathbf{AAG}\text{CUG}_{10}5'\end{matrix}$ was designed to provide a model system for the $\begin{matrix}5'^{1009}\text{CUAAG}^{1013}3'\\3'_{1168}\text{GAAGC}_{1164}5'\end{matrix}$ internal loop (Figure 1b). NMR studies were conducted in the absence and presence of Mg^{2+}. The observed spectra were very similar and did not indicate any change in structure upon addition of 10 mM $MgCl_2$.

Assignments of Exchangeable Protons. RNAstructure version 4.11 predicts seven Watson−Crick base pairs (two

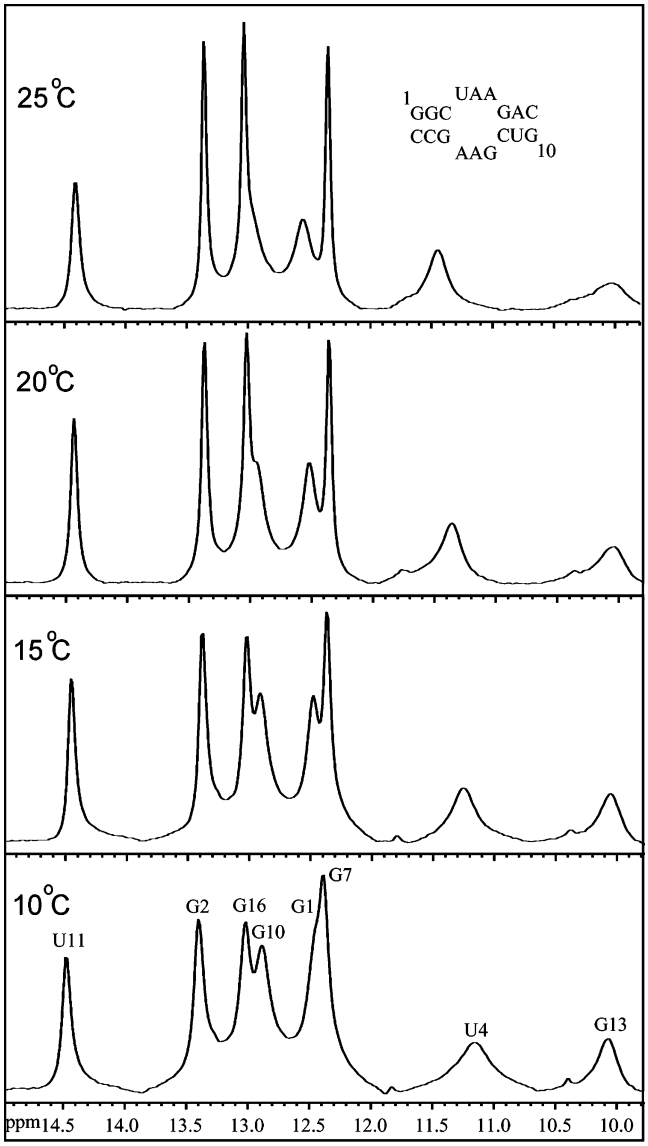

FIGURE 2: One-dimensional imino proton spectra of 5′GGCUAAGAC/3′CCGAAGCUG at temperatures ranging from 10 to 25 °C and pH 6.5. Assignments are given for the bottom spectrum.

UA pairs and five GC pairs) in the stems. Imino proton (10−15 ppm) spectra of the duplex (Figure 2) provide information about base pairing. Five GC pairs were identified (11.5−13.5 ppm), but only the U11H3 peak was found in the usual UA Watson−Crick pair region (13.5−14.5 ppm). Evidently, U4 does not form a Watson−Crick pair. Two broad peaks at 10.1 and 11.2 ppm are from G13H1 and U4H3, respectively. Two weak peaks resonating at 10.4 and 11.8 ppm may be due to imino protons of G13 and U4 in minor conformations. G1H1 and G7H1 are overlapped at 10 °C but are resolved above 10 °C. In contrast, G10H1 and G16H1 are resolved at 10 °C but overlap at higher temperatures. SNOESY spectra were used to assign the imino resonances and to confirm the secondary structure (see the Supporting Information). The cross-peak between G1H1 and G2H1 in the SNOESY spectrum helps assign G1H1 (12.57 ppm). G7H1 (12.35 ppm) was assigned on the basis of its cross-peak to U11H3 (14.46 ppm). G2H1 (13.37 ppm) was assigned on the basis of a cross-peak to G16H1 (13.04 ppm). C3 and C17 amino protons gave cross-peaks with both G2H1 and G16H1. U4H1′ gave a cross-peak to G16H1, and C18H1′

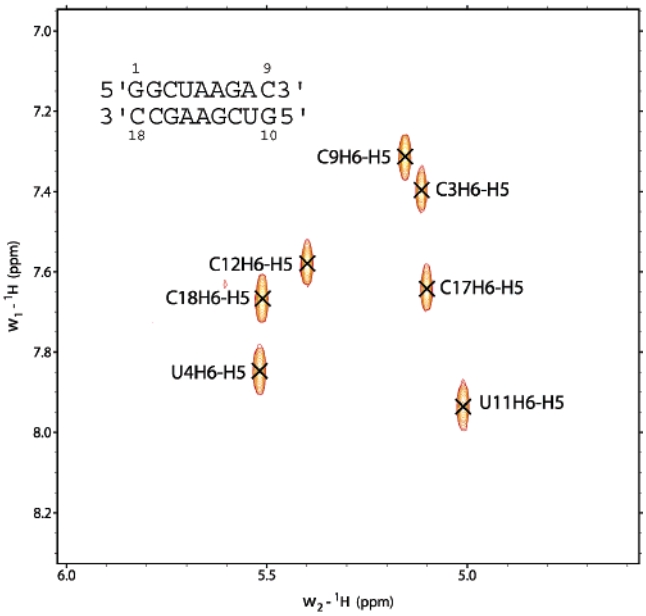

FIGURE 3: TOCSY spectrum of 5′GGCUAAGAC3′ 3′CCGAAGCUG5′ at 25 °C showing H5−H6 cross-peaks for U and C bases.

gave a cross-peak to G2H1 which helped confirm these assignments. The remaining peak in the GC imino region (12.95 ppm) is assigned as G10H1 by exclusion.

Assignments of Nonexchangeable Protons. Resonances for nonexchangeable protons were assigned as described previously (*38*, *39*). The Supporting Information includes details about the chemical shift assignments at 25 °C. A TOCSY spectrum (Figure 3) at 25 °C with a mixing time of 30 ms was used to identify the seven H5−H6 cross-peaks. Lack of extra H5−H6 peaks in the TOCSY spectrum implies one predominant structure. The cross-peaks resulting from H5 and H6 were used to identify the pyrimidine peaks in NOESY spectra, and linked cross-peaks were used to identify consecutive pyrimidines C3, U4, C17, and C18. The H8/H6/H2−H1′/H5 region of the 100 ms NOESY spectrum is shown in Figure 4. Sugar systems associated with the NOESY walk peaks were identified from strong H1′−H2′ (∼2.8−3.0 Å) and H1′−H3′ (∼3.5−4.0 Å) cross-peaks. Assignments follow standard connectivity from G1 to C9. The G7H1′ peak is shifted upfield to 3.82 ppm (25 °C) and was confirmed by natural abundance $^{1}H-^{13}C$ HSQC at 30 °C (Figure S6 of the Supporting Information). The U11−C18 cross-peak also followed a standard NOESY walk, but G10H8 could not be identified because of overlap. Overlap between some H8 peaks caused difficulty in the assignment of sugar protons, and such peaks were not used for modeling. H5′ and H5″ are not assigned stereospecifically, but H5′ protons usually resonate downfield from the H5″ protons due to the negative charge on the phosphate (*40*).

A $^{1}H-^{13}C$ HSQC spectrum (Figure S4 of the Supporting Information) and measurement of base proton relaxation times (data not shown) helped identify and confirm the AH2 resonances. Cross-peaks to A6H1′, G7H1′, and A14H1′ helped assign A6H2. An overlap at 25 °C between the A6H2−A14H1′ and A15H8−A14H1′ cross-peaks was resolved at 30 °C. A8H2 was assigned on the basis of its cross-peaks to G7H1, A8H1′, C9H1′, and C12H1′. Assignment of A14H2 was based on its cross-peaks with A6H1′, A6H2,

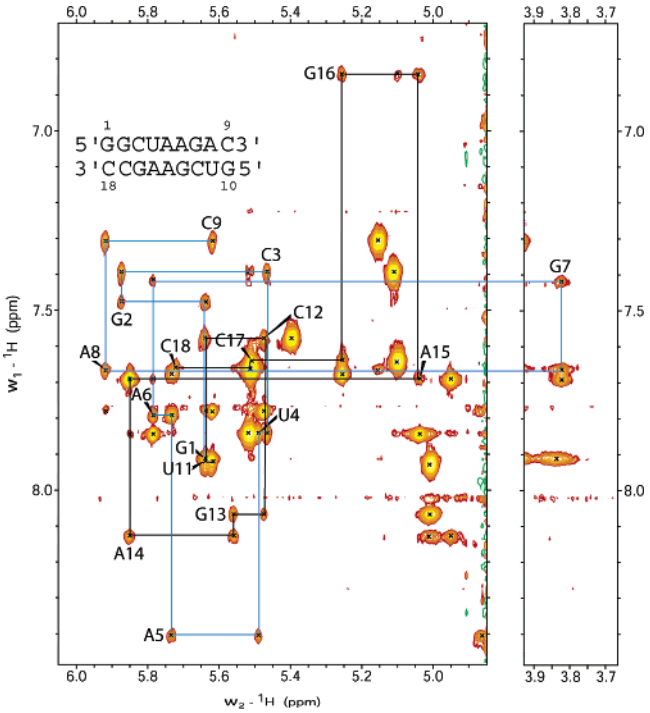

FIGURE 4: NOESY walk region, H8/H6/H2−H1′/H5, of the 100 ms mixing time NOESY spectrum at 25 °C. Blue and black lines trace the walk for the top and bottom strands of the duplex, respectively. The H8/H6−H1′ peaks are labeled.

A14H1′, A14H2′, A15H1′, and A15H8. A15H2 was assigned on the basis of its cross-peaks to A5H1′ and G16H1′. The assignment of A5H2 was based on A5H2−G13H2′ and A5H2−A14H8 cross-peaks.

DQFCOSY, TOCSY and $^{1}H-^{31}P$ HETCOR spectra were used to assign the sugar protons. The H2′ and H3′ peaks were identified from strong cross-peaks to H1′. They were confirmed by walking from strong to medium (*n*) H8/H6−(*n* − 1) H2′/H3′ sequential cross-peaks. Significant overlap in the sugar regions hindered generation of some standard internucleotide restraints for the modeling studies. ^{31}P resonances for all residues except A5 (0.97 ppm) were observed within a 1.5 ppm range (ca. −1.0 to 0.5 ppm) in the HETCOR spectra (see the Supporting Information), which is consistent with A-form geometry (*41*) as the predominant conformation of the duplex.

Spectra Collected upon Addition of 10 mM $MgCl_2$. The addition of 10 mM Mg^{2+} resulted in largely unchanged imino and NOESY spectra (Figures S8 and S9 of the Supporting Information), including frequencies and intensities of cross-peaks. Changes in chemical shifts of imino, H8, and H1′ resonances were mostly ≤0.02 ppm but always within 0.06 ppm (see Table S4 of the Supporting Information). Adenine H2 resonances had cross-peaks similar to those in spectra without $MgCl_2$. In the absence of any relevant changes, all the calculations were performed on spectra obtained without $MgCl_2$.

Structural Determination. Distance and dihedral angle restraints derived from the NMR spectra (see the Supporting Information) were used to model the structure of the $5'^{1}$GGC**UAA**GAC^{9}3′ $3'_{18}$CCG**AAG**CUG$_{10}$5′ duplex. Figure 5 summarizes some of the most important NOEs. A total of 50 structures were generated in the absence of solvent using the simulated

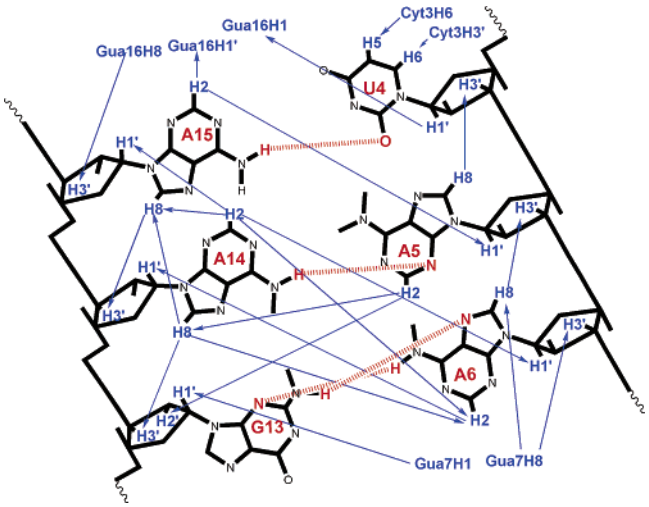

FIGURE 5: Some interstrand and internucleotide NOEs which are important for the structure of the internal loop are shown as blue arrows. Less important hydrogens are omitted. Hydrogen bonds between the bases are colored red.

annealing protocol described in Materials and Methods. The 50 structures converged to satisfy all NMR-derived distance and dihedral angle restraints within 0.1 Å and 2°, respectively. Superposition of the 20 lowest-energy structures reveals that the overall structure and local features are consistently reproduced (Figure 6). The average rmsd values for all atom and heavy atom pairwise superposition of these 20 structures to the average structure are 0.49 and 0.45 Å, respectively. All 20 structures have an A15U4 pair with a single hydrogen bond (cis Watson−Crick/Sugar edge A15U4), a sheared A14A5 pair (trans Hoogsteen/sugar edge A14A5) (*42*) (see Figure 1b), and a sheared A6G13 pair (trans Hoogsteen/sugar edge A6G13) (see Figure 1a). There is cross-strand stacking between A6 and A14 and intrastrand stacking between A14 and A15.

In Solution, A15 (A1167) Is in a Cis Watson−Crick/Sugar Edge Pair with U4 and Not Bulged Out. A15, which was designed to mimic the bulged out base in the crystal structure, is not flipped out in the predominant conformation determined by NMR (Figures 1b, 6, and 7). Cross-peaks in the NOESY spectrum between A15H2 and A5H1′ and between A15H2 and G16H1′ suggest this adenine is inside the helix (Figure 5). The H8 resonance of A15 is no broader than the H8 resonances of other bases and thus does not show evidence of chemical exchange, suggesting that A15 is held in a single conformation. A14H8−A15H8 and A15H8−G16H8 cross-peaks also indicate that the A15 base is located between the A14H8 and G16H8 protons, suggesting stacking of the three purines (Figure 5). The A14H2−A15H1′ NOE is also consistent with A15 stacked on A14. In solution, we find no evidence of A15 being flipped out as in the crystal structure.

The NMR structure has a single hydrogen bond between the amino of A15 and the carbonyl O2 of U4 (cis Watson−Crick/sugar edge A15U4, see Figure 1(b)). The one-dimensional imino region of the ^{1}H spectrum has peaks at 10.2 and 11.6 ppm, either of which could be the U4 imino peak and both of which are outside the usual range for a Watson−Crick AU pair. Typically, 3′-dangling U bases are seen at ∼10.9 ppm (*23*), and most G imino peaks in a GA sheared pair are seen at ∼10 ppm as compared to ∼12 ppm

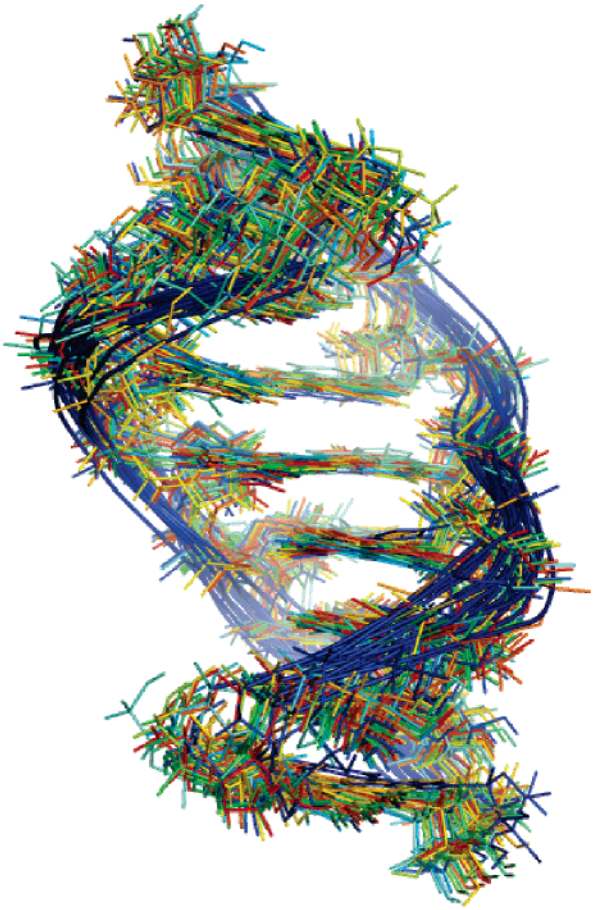

FIGURE 6: Superposition of 20 lowest-energy structures for 5′GGCUAAGAC3′ 3′CCGAAGCUG5′ derived from restrained molecular dynamics. The backbone is distorted between A5 and A6 and between G13 and A14.

in an imino GA pair (*38*, *43*−*46*). The peak at 10.2 ppm becomes broader with an increase in temperature, while the nonterminal Watson−Crick peaks become sharper (Figure 2). This suggests that the peak at 10.2 ppm is the G13 imino proton, which is in chemical exchange with water. By exclusion, the peak at 11.6 ppm is from the U4 imino proton. In contrast, A8−U11 has an imino proton resonance at 14.46 ppm, which is in the expected region for a Watson−Crick pair. Heus et al. (*47*) observed that a UA pair with Watson−Crick geometry but longer than usual hydrogen bonds had the U imino proton resonance shifted upfield to 11.75 ppm. This supports opening of the A15U4 pair in the duplex such that it does not have both the hydrogen bonds of a regular Watson−Crick pair. Several distances are near or outside of the NOE-derived distance limits if the structure is modeled with the A15U4 pair restrained to be a Watson−Crick pair. If the A15U4 pair is a Watson−Crick pair, then the A15H2−A5H1′ distance (3.26 Å) should be shorter (2.88 Å), the A15H2−G16H1′ distance (2.73 Å) should be longer (3.55 Å), the U4H1′−G16H1 distance (2.59 Å) must be longer (3.73 Å, a restraint violation), and the A15H8−A14H3′ distance (3.03 Å) should be shorter (2.17 Å). These restraints, however, are consistent with the proposed model with a single hydrogen bond between A15 and U4.

The unusual UA pair (cis Watson−Crick/sugar edge A15U4) may be stabilized due to stacking of U4 on C3. A 3′-dangling U in a $\frac{CU}{G}$ motif has a $\Delta G^{\circ}_{37,Stack}$ averaging −1.1 kcal/mol (*48*, *49*) and is typically stacked with the C amino H41 closely overlapping the carbonyl O4 atom (∼3.39 Å). In the NMR structure, the U4 carbonyl O4 atom is stacked on C3's amino group in a similar way. Additionally, the U4 carbonyl O4 atom is stacked on the A5 amino H62 atom (∼2.99 Å) which would also favor this unusual UA pair (Figure S7 of the Supporting Information). Quantum mechanical studies of nucleic acid bases have suggested that amino groups are intrinsically nonplanar (*50*, *51*), but empirical potentials such as AMBER consider amino groups to be planar (*27*). The A5 amino group may not be in the planar sp2 geometry but rather in a pyramidal geometry with the amino hydrogens out of plane. This would decrease the distance between the A5 amino hydrogens and U4 carbonyl O4 atom and increase the stability of this interaction.

The crystal structure of a 58-nucleotide RNA and L11 protein from large ribosomal subunit has a similar UA pair between U1065 and A1073 (*52*). This base pair, however, might be stabilized by a tertiary contact in the major groove. The crystal structures of the specificity domain of ribonuclease P (*53*) and of the large ribosomal subunit of *D. radiodurans* (*4*) have similar UA pairs with a single hydrogen bond except that the Hoogsteen edge of A interacts with the sugar edge of U (cis Hoogsteen/sugar edge AU). The hydrogen bonds are between A124H61 and U119O2 and between A911H62 and U890O2. In all cases, the carbonyl O4 atom of U overlaps a C amino group. The positioning of U on C at the end of the helix may be energetically more important than formation of a second hydrogen bond in a UA pair. The face of A that interacts with U may not be a big constraint as long as a hydrogen bond can form between the A amino group and UO2.

Sugar Puckers of Sheared Purine•Purine Pairs. It has been suggested that the sugar pucker of G in a sheared GA pair may depend on the 5′-base (*47*). In the case of a 5′-U, the adjacent G in a sheared GA pair has a *C*2′-*endo* conformation, whereas with a 5′-C, the adjacent G has a *C*3′-*endo* conformation. In the $\frac{5'CUAAG3'}{3'GAAGC5'}$ internal loop described in this work, G13 has a *C*3′-*endo* conformation as predicted. The loop also has a 5′-U adjacent to an AA sheared pair. The A5 sugar has a dynamic conformation indicated by ∼5 Hz coupling between H1′ and H2′. Although both *C*2′-*endo* and *C*3′-*endo* conformations were allowed in the modeling, all modeled structures have a 3′-*endo* sugar conformation. Forcing the A5 sugar into a *C*2′-*endo* conformation did not change the base pairing pattern. The results suggest that rules for the sugar pucker of an A in a sheared AA pair differ from those for a G in a sheared GA pair.

There Is a Sheared A5A14 Pair. In the 20 lowest-energy structures, A5 and A14 form a sheared AA pair (AA trans Hoogsteen/sugar edge A14A5; see Figure 1b) similar to that observed in the structure of the P4−P6 domain of the *T. thermophila* group I intron (*42*). Several observations are consistent with the formation of a sheared AA pair between A5 and A14. Cross-peaks are observed from A5H2 to A14H8 and G13H2′ (Figure 5). The A5H2−A14H8 and −G13H2′ cross-peaks give NMR-derived distances of 4.4 and 4.1 Å, respectively. This indicates that the adenosines are not in a face-to-face geometry, which would result in interstrand A5H2−A14H8 and −G13H2′ distances of at least 7 Å. Also, the A6H2−A14H1′ cross-peak has a NMR-derived distance

Table 1: Thermodynamic Parameters for Duplex Formation by 5′GGCUAAGAC/3′CCGAAGCUG[a]

	$\Delta H°$ (kcal/mol)	$\Delta S°$ (cal K^{-1} mol^{-1})	$\Delta G°_{37}$ (kcal/mol)	T_m[b] (°C)
curve fit parameters	−83.39 ± 10.63 (−85.70 ± 5.59)	−240.04 ± 34.55 (−245.69 ± 17.74)	−8.94 ± 0.24 (−9.50 ± 0.23)	46.2 (48.1)
$1/T_M$ vs $\ln(C_T)$ parameters[c]	−80.83 ± 9.65 (−80.79 ± 7.92)	−231.73 ± 30.54 (−230.27 ± 24.90)	−8.96 ± 0.30 (−9.37 ± 0.25)	46.6 (48.3)
predicted[d]	−56.08	−154.16	−8.23	46.9

[a] Buffer conditions: 1 M NaCl, 20 mM sodium cacodylate, and 0.5 mM Na_2EDTA (pH 6.5). Values in parentheses are for the following buffer conditions: 0.15 M KCl, 20 mM sodium cacodylate, and 10 mM $MgCl_2$ (pH 6.5). [b] Calculated for an oligonucleotide total strand concentration of 10^{-4} M. [c] Plots of T_M^{-1} vs $\ln(C_T/4)$ resulted in an r value of 0.97. [d] Values for $\Delta G°$, $\Delta H°$, and $\Delta S°$ were predicted from updated nearest neighbor parameters that incorporate previously published data (*1*, *2*, *71*, *72*, *96*).

of 2.6 Å, which indicates that the minor groove is narrower than that typically observed for an A-form helix. For equivalent protons in an A-form helix, the distance can range from 3.2 to 4.1 Å (*54*, *55*). Therefore, the observed distances suggest that the minor groove does not widen to accommodate a face-to-face adenine pair.

The Sheared G13A6 Pair Is Consistent with the Crystal Structure. The G13A6 sheared pair is characterized by four pieces of evidence: (1) a G imino peak at 10.1 ppm in the one-dimensional spectrum indicating the absence of an imino proton hydrogen bond, (2) a medium-intensity cross-peak between G7H1 and G13H1′ (2.89 Å ± 40%), (3) a NOE cross-peak between A6H2 and G7H1′ (2.96 Å ± 40%) indicating that the Watson−Crick edge of A6 is in the minor groove, and (4) an upfield shift of the G7H1′ resonance to 3.82 ppm. An upfield shift of H1′ for the nucleotide 3′ of A is expected for a sheared GA pair (*43*, *44*, *47*, *56*−*59*). This upfield chemical shift can be explained by the large ring current effect of adenosine above the H1′ proton of the 3′-residue (*60*, *61*).

Thermodynamics. Thermodynamic parameters for the $\begin{matrix}5'GGCUAAGAC3'\\3'CCGAAGCUG5'\end{matrix}$ duplex were measured in the presence of 1 M NaCl and also in the presence of 0.15 M KCl and 10 mM $MgCl_2$ (Table 1). Results for the two buffers were the same within experimental error (Table 1). The free energy increment at 37 °C for the loop $\begin{matrix}5'CUAAG3'\\3'GAAGC5'\end{matrix}$ was calculated as −1.78 kcal/mol (Table 2) from

$$\Delta G°_{37LOOP}\begin{pmatrix}5'CUAAG3'\\3'GAAGC5'\end{pmatrix} = \Delta G°_{37}\begin{pmatrix}5'GGCUAAGAC3'\\3'CCGAAGCUG5'\end{pmatrix} - \Delta G°_{37}\begin{pmatrix}5'GGCGAC3'\\3'CCGCUG5'\end{pmatrix} + \Delta G°_{37}\begin{pmatrix}5'CG3'\\3'GC5'\end{pmatrix} \quad (2)$$

In this equation, $\Delta G°_{37}\begin{pmatrix}5'GGCUAAGAC3'\\3'CCGAAGCUG5'\end{pmatrix}$ and $\Delta G°_{37}\begin{pmatrix}5'GGCGAC3'\\3'CCGCUG5'\end{pmatrix}$ are the free energy changes for duplex formation with and without the internal loop at 37 °C, respectively. $\Delta G°_{37}\begin{pmatrix}5'GGCUAAGAC3'\\3'CCGAAGCUG5'\end{pmatrix}$ is obtained from T_M^{-1} versus $\ln(C_T/4)$ plots. $\Delta G°_{37}\begin{pmatrix}5'GGCGAC3'\\3'CCGCUG5'\end{pmatrix}$ was calculated from nearest neighbor parameters (*2*).

On the basis of thermodynamics, the internal loop is predicted to be a 2 × 2 nucleotide internal loop closed on one side with a Watson−Crick UA pair. The free energy increment for formation of the $\begin{matrix}5'CUAAG3'\\3'GAAGC5'\end{matrix}$ loop at 37 °C can be predicted by averaging $\Delta G°_{37}\begin{pmatrix}5'UAAA3'\\3'AAAU5'\end{pmatrix}$ and $\Delta G°_{37}\begin{pmatrix}5'CGAG3'\\3'GAGC5'\end{pmatrix}$ (*1*, *62*) and adding the nearest neighbor parameter for $\Delta G°_{37}\begin{pmatrix}5'CU3'\\3'GA5'\end{pmatrix}$. The value of −1.06 kcal/mol is similar to the experimental value of −1.78 kcal/mol (Table 2). The predicted $\Delta G°_{37LOOP}$ for the loop structure in the crystal structures of LSU rRNA is 3.53 kcal/mol if it is considered as a single nucleotide bulge and a 1 × 2 nucleotide internal loop separated by a UA pair or 1.09 kcal/mol if it is treated as a 3 × 3 nucleotide internal loop. These estimated $\Delta G°_{37LOOP}$ values are 5.31 and 2.87 kcal/mol, respectively, less favorable than that measured for the 2 × 2 loop with AA and GA pairs (Table 2).

Table 2: Free Energies of Loop Formation at 37 °C

secondary structure of the loop[a]	source of the structural model	$\Delta G°_{37LOOP}$ (kcal/mol)
C **UAA** G G **AAG** C	NMR[b]	−1.78[c]
CU**AA**G GA**AG**C	predicted	−1.06[d]
C **UAA**G G **A A** G C	crystal structure of *D. radiodurans*	3.53[d] 1.09[e]

[a] Non-Watson−Crick pairs are bold. [b] UA forms a single hydrogen-bonded base pair in the NMR structure. [c] Measured $\Delta G°_{37LOOP}$ as calculated from eq 2. [d] Value for $\Delta G°_{LOOP}$ as predicted from updated nearest neighbor parameters that incorporate previously published data (*1*, *2*, *96*) and assuming a single Watson−Crick UA pair. [e] Value for $\Delta G°_{LOOP}$ predicted assuming a 3 × 3 nucleotide internal loop (*73*) with the non-Watson−Crick UA pair treated as part of the loop.

DISCUSSION

Secondary or three-dimensional structures are known for only a fraction of known RNA sequences. An understanding of the sequence dependence of the energetics and flexibility of various structural elements would facilitate modeling of structures and contribute to determination of structure−function relationships. Internal loops are a common structural element in RNAs. This study shows that the solution structure of a phylogenetically conserved internal loop, 5′CUAAG/3′GAAGC, is close to that predicted by a nearest neighbor energetic model but differs from that found in the crystal structures of ribosomal subunits (*4*−*6*). Evidently, tertiary and quaternary interactions with rRNA and protein are sufficiently strong to remodel the three-dimensional structure of the loop, suggesting induced fit binding (*63*).

The Crystal and Solution Structures Differ. The solution structure of the internal loop differs from that in the crystal structures of large ribosomal subunits from *D. radiodurans* and *E. coli* but is close to the 2 × 2 nucleotide internal loop structure expected from a model used to approximate energetics (Figure 1). For example, in the *D. radiodurans* crystal structure, A1167 is flipped out of the helix. U1010

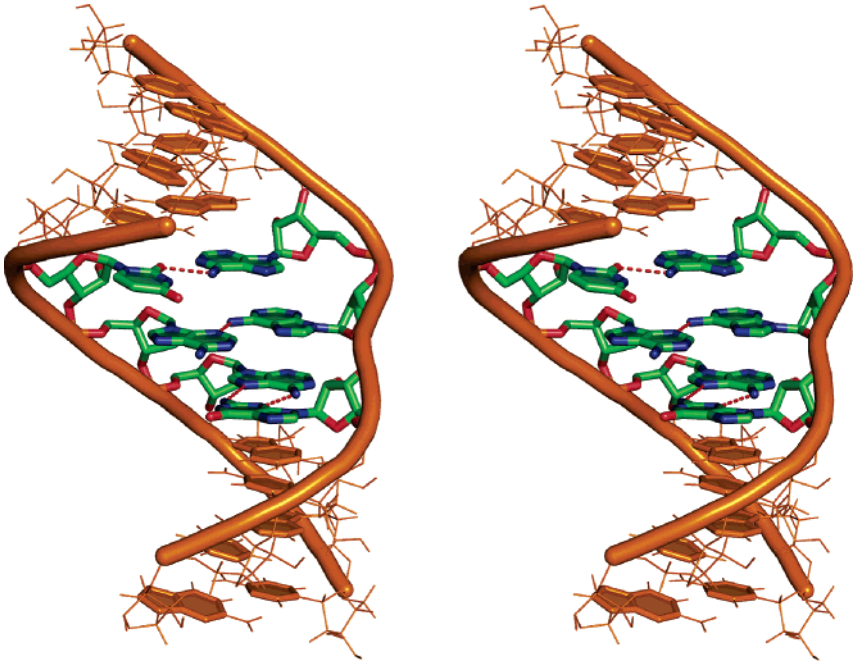

FIGURE 7: Stereoview of the average modeled low-energy structure for $\begin{matrix}5'\text{GGCUAAGAC}3'\\3'\text{CCGAAGCUG}5'\end{matrix}$. The stem regions are depicted as orange lines. The loop region is shown as sticks where red is for oxygen, blue for nitrogen, green for carbon, and an orange cartoon for the backbone. The 5′G1 base is at the top, and the 3′C8 base is at the bottom. Dashed red lines represent hydrogen bonds between bases. Hydrogen atoms are hidden for clarity.

forms a reverse Hoogsteen pair with A1166 (Figure 1a). A1011 forms a cross-strand stack between A1012 and A1166. A1012 forms a sheared GA pair with G1165. The solution structure differs from the crystal structure in that A15, which corresponds to the bulged out A1167 in the crystal structure, forms a single hydrogen-bonded pair with U4, while A5 and A14 form an AA sheared pair adjacent to the GA sheared pair. A6 forms a cross-strand stack with A14, and A14 makes an intrastrand stack on A15 (Figure 8). A6H2−A14H2, A6H2−A14H1′, and A14H2−A6H1′ cross-peaks suggest that A6 is stacking on A14. A14H2−A15H1′, A14H8−A15H8, and A14H2−A15H8 cross-peaks provide evidence for stacking of A14 on A15 (Figure 5).

Magnesium Binding. No Mg^{2+} that might stabilize the internal loop structure is observed in the *D. radiodurans* crystal structure (PDB entry 1NKW), but a Mg^{2+} ion is present between the phosphates of U999 and C1153 (equivalent to U4 and C12, respectively) in the *E. coli* crystal structure (PDB entry 2AWB). A Mg^{2+} ion is also observed at a similar position in the *H. marismortui* crystal structure (PDB entry 1JJ2). The distance between the U999 and C1153 phosphates is ∼6.8 Å in the crystal which is much shorter than the distance of ∼18 Å in the NMR model. Addition of 10 mM $MgCl_2$ to the NMR sample does not induce any significant change in the NMR spectra. Evidently, Mg^{2+} alone cannot induce a conformational switch to flip out A15.

Tertiary and Quaternary Interactions May Stabilize the Bulged A Structure. In *D. radiodurans* $\begin{pmatrix}5'^{1009}\text{CUAAG}^{1013}3'\\3'_{1168}\text{GAAGC}_{1164}5'\end{pmatrix}$, the bulged A1167 is close to the α-carbons of Arg48, Arg51, and Asn52 of ribosomal protein L20 (residues within a 6 Å radius) and to A986, G987, and A1001 of the rRNA (residues within a 5 Å radius). The bulged G1260 in the *H. marismortui* internal loop is close to Lys158 and Gly156 from ribosomal protein L32E and to the A1073, G1074, and A1088 residues of the rRNA (residues within a 5 Å radius). As shown in the bottom panel of Figure 9, the *E. coli* internal loop $\begin{pmatrix}5'^{998}\text{CUAAG}^{1002}3'\\3'_{1157}\text{GAAGC}_{1153}5'\end{pmatrix}$ is close to Arg47, Arg50, Gln51, and Arg54 of ribosomal protein L20 and Lys81 of ribosomal protein L22 (residues within a 5 Å radius). It is also close to A975, A976, and A990 of the rRNA (residues within a 5 Å radius). Thus, different tertiary and quaternary interactions may stabilize the structures of these loops, rather than preorganized interactions within the loops themselves.

Effects of Non-Watson−Crick Pairs on Global Structure. Tandem AA and GA pairs are commonly found in biologically functional RNA (*64*−*66*). One of the features of the isolated loop in solution is that the backbone is slightly narrower at two positions. The C1′−C1′ distance in the typical Watson−Crick pairs in this structure is ∼10.7 Å. The distance between A5C1′ and A14C1′ is 9.4 Å, and the distance between A6C1′ and G13C1′ is 9.5 Å because the backbone is narrower due to formation of sheared pairs. In the crystals, the GA sheared pair has an average C1′−C1′ distance of 9.7 Å. The NMR-derived distance between A6H2 and A14H1′ is 2.6 Å compared to a distance of 3.2−4.1 Å in a typical A-form helix so the minor groove is narrowed. Similar constriction or kinking of the backbone has been observed in previous structures involving purine·purine sheared pairs (*23*, *44*, *47*, *55*). The U4A15 pair in the NMR structure has a C1′−C1′ distance of 10.9 Å, whereas the reverse Hoogsteen pair in the crystals has an average C1′−C1′ distance of 9.6 Å.

The Switch in Conformation May Be Facilitated by A5 Dynamics. NMR spectra of this internal loop indicate structural dynamics at A5. The A5H2 peak is broad (∼21 Hz) and is weak in the HSQC spectrum, suggesting dynamics (Figure S4). This is consistent with the ∼5 Hz coupling

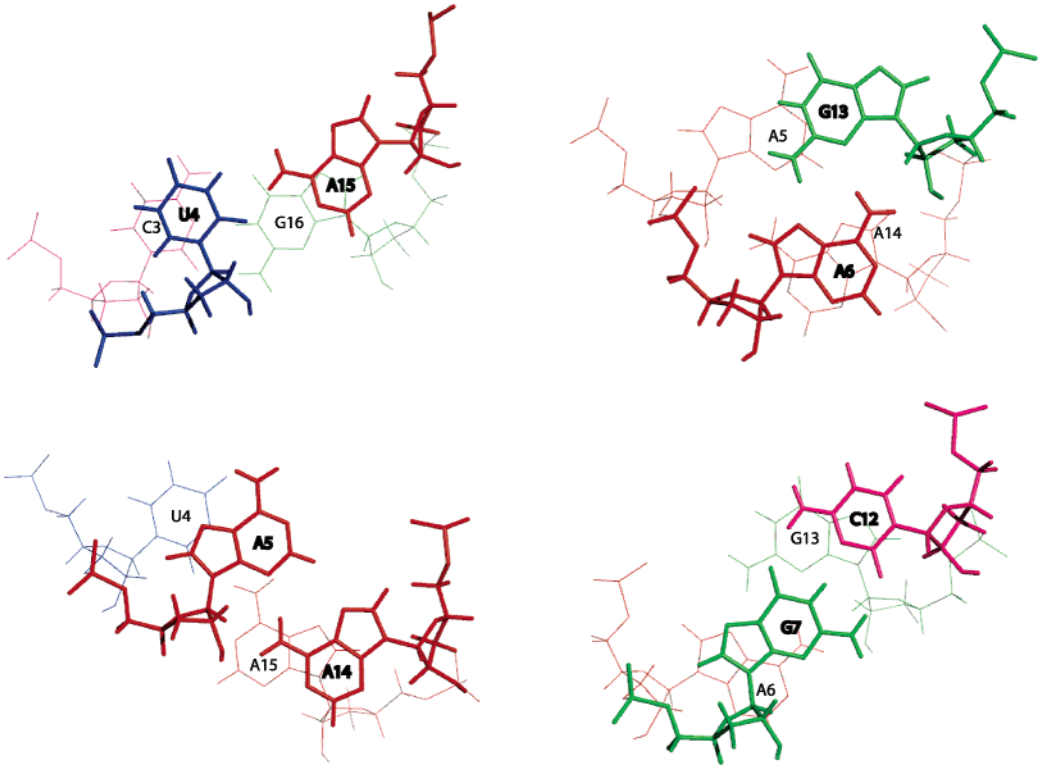

FIGURE 8: Stacking patterns of the residues in the loop generated from the average of the 20 lowest-free energy structures using 3DNA (*97*). Adenines are colored red, cytosines pink, guanines green, and uracils blue. The base pair closer to the viewer is bold.

between A5H1′ and A5H2′ suggesting that the ribose ring is interchanging between *C2′*- and *C3′-endo* conformations. Dynamics at this residue are also consistent with the lack of base stacking interactions between A5 and neighboring residues in the modeled structures (Figures 7 and 8). A5 is held by only a single hydrogen bond in a sheared pair with A14, and A15 is held by only a single hydrogen bond to U4 (Figure 1b). Dynamics associated with A5 may provide a pathway for this internal loop to switch between two very different conformations. Rapid switching between syn and anti guanosines in GG pairs has been observed (*67*) and is thought to be facilitated by dynamics in sugar puckers (*68*). In the crystal structures, the equivalent of A5 forms a cross-strand stack with the equivalent of A14 (Figure 9). This stack allows the Hoogsteen edge of A14 to form two hydrogen bonds with U4 in a reverse Hoogsteen pair, thus compensating for the hydrogen bonds lost between U4 and A15 and between A5 and A14 and helping to stabilize the induced fit conformation.

The Solution Structure and Energetics Are Similar to Those Predicted from Simple Models. While the solution structure of the 5′CUAAG/3′GAAGC loop differs from the crystal structures, it is very similar to that expected from models used to approximate the structures and energetics of 2 × 2 nucleotide internal loops (*1*, *69*). In the energetic model, the free energy increment for a sequence nonsymmetric loop, 5′PXYS/3′QWZT, is the average of that measured for the sequence symmetric loops, 5′PXWQ/3′QWXP and 5′TZYS/3′SYZT, when the non-Watson−Crick pairs, XW and ZY, are sterically matched. Here PQ and TS are Watson−Crick pairs. NMR structures of the internal loops, CGAG (*44*), CAAG (*29*, *70*), and UGAA (*47*), all have sheared purine·purine pairs so that the same is expected for the 5′UAAG/3′AAGC loop. The average of the measured free energy increments at 37 °C for 5′UAAA/3′AAAU (*62*, *71*) and 5′CGAG/3′GAGC (*62*, *72*) is 1.02 kcal/mol, which when added to the −2.08 kcal/mol free energy increment for 5′CU/3′GA predicts a free energy increment of −1.06 kcal/mol for the 5′CUAAG/3′GAAGC sequence. This is within experimental error of the measured value of −1.78 kcal/mol (Table 2). Evidently, there are no unexpected interactions between the sheared AA and GA pairs and the unusual UA pair has a stability similar to that of a Watson−Crick pair in this context, or the two potential effects offset each other.

Potential Functional Advantages of the Difference between Solution and Crystal Structures. The size symmetric loops, $5'^{1009}$CUAAG$^{1013}3'$ $3'_{1168}$GAAGC$_{1164}5'$ and $5'^{1364}$CUAAG$^{1368}3'$ $3'_{1393}$GUAGC$_{1389}5'$, in *D. radiodurans* (*4*) have A1167 and U1392 flipped out, respectively, and similar structures are observed for the equivalent loops in *E. coli* (*5*). This base flip is not expected on the basis of the sequence dependence of internal loop stability (*73*) and is not observed in the solution structure of the $5'^{1009}$CUAAG$^{1013}3'$ $3'_{1168}$GAAGC$_{1164}5'$ internal loop. This internal loop is highly conserved, which suggests a need for it to facilitate a structural switch. The symmetric loop $5'^{616}$UUAAG$^{620}3'$ $3'_{633}$GAGGC$_{629}5'$ in *D. radiodurans* has G631 flipped out. It is surprising that these loops that are expected to be structured are involved in base flipping. Presumably, an unstructured loop would undergo conformational changes more easily. Perhaps the structure of the isolated loop is important for initial folding of the RNA and/or initial recognition by the protein, while the induced structure is necessary for specific binding by protein or RNA. Thus induced fit binding is likely to be

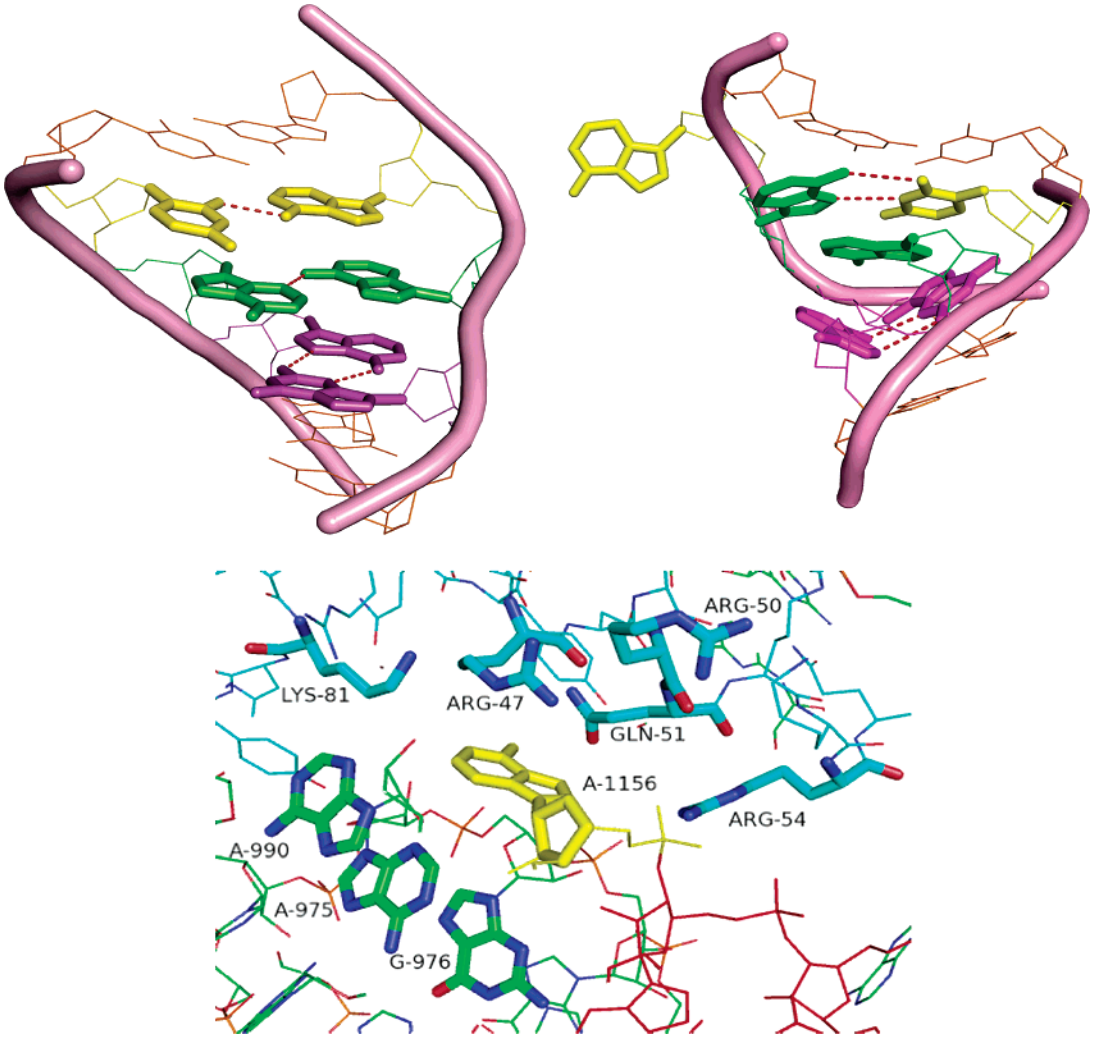

FIGURE 9: Comparison between internal loop structures obtained by NMR (top left panel) and crystallography of the *D. radiodurans* large ribosomal subunit (top right panel) (*4*). Analogous bases are colored the same. The flanking pairs and the backbone of the loop are colored orange. In the NMR structure, A15 and U4 (yellow) are paired in a cis Watson−Crick/sugar edge pair. In the crystal structure, A15 is bulged out and U4 forms a reverse Hoogsteen pair with A14. In the NMR structure, A5 and A14 (green) form a sheared pair but A5 and A14 are stacked in the crystal structure. G13 and A6 (pink) form a sheared pair in both cases. The bottom panel shows the tertiary contacts close to the bulged adenine (A1156) in the *E. coli* crystal structure (residues within a 5 Å radius). The overall folding of the loop in *E. coli* is the same as for *D. radiodurans* (rmsd = 0.67 Å).

important for other internal loops predicted to be preorganized.

Induced fit binding (*63*) is common in RNA−protein complexes, but the reasons for this universality are not known (*74*−*78*). The differences between the NMR and crystal structures of the 5′CUAAG/3′GAAGC loop suggest some possible advantages for this induced fit binding. One potential reason for a bulged nucleotide initially to be enclosed in a structured loop is that it may be protected from chemical cleavage. It is known that backbone cleavage is enhanced at bulged nucleotides due to increased in-line orientation of the 2′-OH with the phosphate (*79*, *80*). A structured loop incorporating a nucleotide to be bulged would also provide increased stability for the initial folding of the RNA. The nucleotide can later be bulged to provide RNA−RNA tertiary interactions or a binding site for protein. The L20 protein is one of the first to bind to the rRNA and is among the six proteins that are required for formation of the first key intermediate in ribosomal assembly (*13*−*15*). The binding of protein could allow temporal ordering of a long-range conformational change required for assembly of a large complex such as the ribosome. Internal loops with flipped bases can lead to considerable misalignment of helix axes (*81*).

The initial structure of the loop may also provide a kinetically important intermediate that allows the protein to rapidly find the location requiring base flipping; i.e., the initial structure may serve as a "bookmark" identifying the site to be opened. Interestingly, Restrictocin recognizes a bulged G-type S-turn in the sarcin−ricin tetraloop of rat 23S rRNA and flips a G on the 5′-side of the cleavage site facilitating in-line orientation of the 2′-OH of G with the P-O5 group of the 3′-A (*82*, *83*).

The nearest neighbor model (*2*, *3*) predicts a free energy increment of 3.53 kcal/mol (Table 2) for the bulged internal loop structure found in crystals of the LSU of *D. radiodurans* (*4*) and *E. coli* (*5*). The predicted free energy is 5.3 kcal/mol less favorable than that measured for the oligonucleotide mimic (Table 2). This predicted free energy difference between the crystal and solution structures suggests that tertiary or quaternary contacts with RNA or protein provide the necessary free energy difference to stabilize the crystal

structure. If this was a case of lock and key binding (*19*), the structures in the crystal and solution would be identical and fewer tertiary and/or quaternary contacts would be needed to achieve the required association constant. The increased number of favorable contacts with the protein and RNA required for induced fit binding may provide increased specificity.

Single-base bulges are commonly occurring secondary structural elements (*84*) that play vital roles in RNA folding and recognition (*85*, *86*). Base flipping is also commonly induced by nucleic acid-modifying enzymes to gain access to the site of modification (*87*−*89*). For example, tRNA pseudouridine synthetase, TruB, flips target base U55 by disrupting a GU base pair (*90*). Modeling studies conducted with 23S rRNA methyltransferase RrmJ suggest that the substrate nucleotide U2552 has to bulge out to be methylated at the 2′-O position (*91*). Initiation factor IF1 binds 16S rRNA and induces flipping of A1492 and A1493 from helix 44 (*92*). Evidently, many proteins are able to provide sufficient free energy to compensate for the cost of flipping a base.

Molecular dynamics and thermodynamics calculations on an interaction between the N-terminal RNP domain of the U1A mutant and stem−loop 2 of U1 snRNA suggest that stacking interactions between a protein residue and RNA bases could contribute as much as ∼4.4 kcal/mol to binding (*93*). This suggests that a single protein contact like stacking could compensate for base flipping and subsequent base pairing rearrangement. Many enzymes insert a protein residue into the cavity created by a flipped out base to stabilize the structure (*90*, *94*). It is also likely that the RNA contacts from other parts of the ribosome play a role in this induced fit recognition.

With a growing body of evidence of base flipping as a functional mechanism used by RNA, prediction of sites and loops where base flipping can occur will become important. Identification of sequences with base flipping propensities may help drug design by identifying sites where small molecules could compete with tertiary and quaternary binding by targeting the stable structure formed in the absence of these interactions (*95*).

NOTE ADDED IN PROOF

A paper published during review of this paper shows that 5′UAA/3′NAG motifs are widespread and involved in long range tertiary interactions (*98*).

ACKNOWLEDGMENT

We thank Dr. Brent Znosko for his help during setup of NMR experiments. We thank Mr. Rahul Tyagi for his help in writing the C++ program used to convert Sparky output to Discover input.

SUPPORTING INFORMATION AVAILABLE

A table of NMR restraints, a table with the summary of NOE distance restraints for each residue, a table listing chemical shift assignments, and ^{31}P HETCOR, SNOESY, HSQC, and DQFCOSY spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. 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*, 7287−92.
2. 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*, 14719−35.
3. 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*, 911−40.
4. 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*, 679−88.
5. Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Cate, J. H. D. (2005) Structures of the Bacterial Ribosome at 3.5 Å Resolution, *Science 310*, 827−34.
6. 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 Å resolution, *Science 289*, 905−20.
7. Clemons, W. M., Jr., May, J. L., Wimberly, B. T., McCutcheon, J. P., Capel, M. S., and Ramakrishnan, V. (1999) Structure of a bacterial 30S ribosomal subunit at 5.5 Å resolution, *Nature 400*, 833−40.
8. Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000) Structure of the 30S ribosomal subunit, *Nature 407*, 327−39.
9. Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H. D., and Noller, H. F. (2001) Crystal Structure of the Ribosome at 5.5 Å Resolution, *Science 292*, 883−96.
10. Guillier, M., Allemand, F., Raibaud, S., Dardel, F., Springer, M., and Chiaruttini, C. (2002) Translational feedback regulation of the gene for L35 in *Escherichia coli* requires binding of ribosomal protein L20 to two sites in its leader mRNA: A possible case of ribosomal RNA-messenger RNA molecular mimicry, *RNA 8*, 878−89.
11. Klosterman, P. S., Tamura, M., Holbrook, S. R., and Brenner, S. E. (2002) SCOR: A Structural Classification of RNA database, *Nucleic Acids Res. 30*, 392−4.
12. Tamura, M., and Holbrook, S. R. (2002) Sequence and structural conservation in RNA ribose zippers, *J. Mol. Biol. 320*, 455−74.
13. Herold, M., and Nierhaus, K. H. (1987) Incorporation of six additional proteins to complete the assembly map of the 50S subunit from *Escherichia coli* ribosomes, *J. Biol. Chem. 262*, 8826−33.
14. Rohl, R., and Nierhaus, K. H. (1982) Assembly map of the large subunit (50S) of *Escherichia coli* ribosomes, *Proc. Natl. Acad. Sci. U.S.A. 79*, 729−33.
15. Nowotny, V., and Nierhaus, K. H. (1980) Protein L20 from the large subunit of *Escherichia coli* ribosomes is an assembly protein, *J. Mol. Biol. 137*, 391−9.
16. Raibaud, S., Vachette, P., Guillier, M., Allemand, F., Chiaruttini, C., and Dardel, F. (2003) How bacterial ribosomal protein L20 assembles with 23S ribosomal RNA and its own messenger RNA, *J. Biol. Chem. 278*, 36522−30.
17. Raibaud, S., Lebars, I., Guillier, M., Chiaruttini, C., Bontems, F., Rak, A., Garber, M., Allemand, F., Springer, M., and Dardel, F. (2002) NMR structure of bacterial ribosomal protein L20: Implications for ribosome assembly and translational control, *J. Mol. Biol. 323*, 143−51.
18. Koshland, D. E. (1958) Application of a Theory of Enzyme Specificity to Protein Synthesis, *Proc. Natl. Acad. Sci. U.S.A. 44*, 98−104.
19. Fischer, E. (1894) Einfluss der Configuration auf die Wirkung der Enzyme, *Ber. Dtsch. Chem. Ges. 27*, 2984−93.
20. 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*, 2677−84.
21. 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*, 7845−54.
22. 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*, 779−82.
23. Chen, G., Znosko, B. M., Kennedy, S. D., Krugh, T. R., and Turner, D. H. (2005) Solution structure of an RNA internal loop with three consecutive sheared GA pairs, *Biochemistry 44*, 2845−56.
24. Smallcombe, S. H. (1993) Solvent suppression with symmetrically-shifted pulses, *J. Am. Chem. Soc. 115*, 4776−85.
25. 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*, 603−14.
26. Goddard, T. D., and Kneller, D. G. (2004) *Sparky: NMR assignment and integration software*, University of California, San Francisco.
27. 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) A 2nd generation force field for the simulation of proteins, nucleic acids, and organic molecules, *J. Am. Chem. Soc. 117*, 5179−97.
28. Ding, H. Q., Karasawa, N., and Goddard, W. A., III (1992) Atomic level simulations on a million particles: The cell multipole method for Coulomb and London nonbond interactions, *J. Chem. Phys. 97*, 4309−15.
29. Znosko, B. M., Burkard, M. E., Schroeder, S. J., Krugh, T. R., and Turner, D. H. (2002) Sheared Aanti·Aanti base pairs in a destabilizing 2 × 2 internal loop: The NMR structure of 5′(rGGCAAGCCU)2, *Biochemistry 41*, 14969−77.
30. Burkard, M. E., and Turner, D. H. (2000) NMR structures of r(GCAGGCGUGC)2 and determinants of stability for single guanosine-guanosine base pairs, *Biochemistry 39*, 11748−62.
31. 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 589−95, CRC Press, Cleveland, OH.
32. 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.
33. Peritz, A. E., Kierzek, R., Sugimoto, N., and Turner, D. H. (1991) Thermodynamic study of internal loops in oligoribonucleotides: Symmetric loops are more stable than asymmetric loops, *Biochemistry 30*, 6428−36.
34. Petersheim, M., and Turner, D. H. (1983) Base-stacking and base-pairing contributions to helix stability: Thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp, *Biochemistry 22*, 256−63.
35. McDowell, J. A., and Turner, D. H. (1996) Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: Solution structure of (rGAGGUCUC)$_2$ by two-dimensional NMR and simulated annealing, *Biochemistry 35*, 14077−89.
36. Borer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974) Stability of ribonucleic acid double-stranded helices, *J. Mol. Biol. 86*, 843−53.
37. 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 Bioinf. 3*, 2.
38. Varani, G., Aboulela, F., and Allain, F. H. T. (1996) NMR investigation of RNA structure, *Prog. Nucl. Magn. Reson. Spectrosc. 29*, 51−127.
39. Varani, G., and Tinoco, I. (1991) RNA structure and NMR spectroscopy, *Q. Rev. Biophys. 24*, 479−532.
40. Varani, G., Cheong, C., and Tinoco, I., Jr. (1991) Structure of an unusually stable RNA hairpin, *Biochemistry 30*, 3280−9.
41. Gorenstein, D. (1984) *^{31}P NMR, Principles and Applications*, Academic Press, New York.
42. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) Crystal structure of a group I ribozyme domain: Principles of RNA packing, *Science 273*, 1678−85.
43. Heus, H. A., and Pardi, A. (1991) Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops, *Science 253*, 191−4.
44. SantaLucia, J., Jr., and Turner, D. H. (1993) Structure of (rGGCGAGCC)$_2$ in solution from NMR and restrained molecular dynamics, *Biochemistry 32*, 12612−23.
45. 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*, 4449−60.
46. Wu, M., and Turner, D. H. (1996) Solution structure of (rGCGGACGC)$_2$ by two-dimensional NMR and the iterative relaxation matrix approach, *Biochemistry 35*, 9677−89.
47. Heus, H. A., Wijmenga, S. S., Hoppe, H., and Hilbers, C. W. (1997) The detailed structure of tandem G·A mismatched base-pair motifs in RNA duplexes is context dependent, *J. Mol. Biol. 271*, 147−58.
48. 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*, 967−82.
49. Freier, S. M., Sugimoto, N., Sinclair, A., Alkema, D., Neilson, T., Kierzek, R., Caruthers, M. H., and Turner, D. H. (1986) Stability of XGCGCp, GCGCYp, and XGCGCYp helixes: An empirical estimate of the energetics of hydrogen bonds in nucleic acids, *Biochemistry 25*, 3214−9.
50. Sponer, J., Burcl, R., and Hobza, P. (1994) Interactions between amino groups in DNA. An ab initio study and a comparison with empirical potentials, *J. Biomol. Struct. Dyn. 11*, 1357−76.
51. Sponer, J., Mokdad, A., Sponer, J. E., Spackova, N., Leszczynski, J., and Leontis, N. B. (2003) Unique tertiary and neighbor interactions determine conservation patterns of cis Watson−Crick A/G base-pairs, *J. Mol. Biol. 330*, 967−78.
52. 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*, 1171−4.
53. Krasilnikov, A. S., Yang, X., Pan, T., and Mondragon, A. (2003) Crystal structure of the specificity domain of ribonuclease P, *Nature 421*, 760−4.
54. Klosterman, P. S., Shah, S. A., and Steitz, T. A. (1999) Crystal structures of two plasmid copy control related RNA duplexes: An 18 base pair duplex at 1.20 Å resolution and a 19 base pair duplex at 1.55 Å resolution, *Biochemistry 38*, 14784−92.
55. Znosko, B. M., Kennedy, S. D., Wille, P. C., Krugh, T. R., and Turner, D. H. (2004) Structural features and thermodynamics of the J4/5 loop from the *Candida albicans* and *Candida dubliniensis* group I introns, *Biochemistry 43*, 15822−37.
56. Flinders, J., and Dieckmann, T. (2001) A pH controlled conformational switch in the cleavage site of the VS ribozyme substrate RNA, *J. Mol. Biol. 308*, 665−79.
57. Hoffmann, B., Mitchell, G. T., Gendron, P., Major, F., Andersen, A. A., Collins, R. A., and Legault, P. (2003) NMR structure of the active conformation of the Varkud satellite ribozyme cleavage site, *Proc. Natl. Acad. Sci. U.S.A. 100*, 7003−8.
58. 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*, 968−80.
59. Michiels, P. J. A., Schouten, C. H. J., Hilbers, C. W., and Heus, H. A. (2000) Structure of the ribozyme substrate hairpin of *Neurospora* VS RNA: A close look at the cleavage site, *RNA 6*, 1821−32.
60. Case, D. A. (1995) Calibration of ring-current effects in proteins and nucleic acids, *J. Biomol. NMR 6*, 341−6.
61. Nowakowski, J., Miller, J. L., Kollman, P. A., and Tinoco, I. (1996) Time evolution of NMR proton chemical shifts of an RNA hairpin during a molecular dynamics simulation, *J. Am. Chem. Soc. 118*, 12812−20.
62. Schroeder, S. J., and Turner, D. H. (2001) Thermodynamic stabilities of internal loops with GU closing pairs in RNA, *Biochemistry 40*, 11509−17.
63. Koshland, D. E., Jr., and Neet, K. E. (1968) The catalytic and regulatory properties of enzymes, *Annu. Rev. Biochem. 37*, 359−410.
64. Leontis, N. B., and Westhof, E. (1998) Conserved geometrical base-pairing patterns in RNA, *Q. Rev. Biophys. 31*, 399−455.
65. Nagaswamy, U., Larios-Sanz, M., Hury, J., Collins, S., Zhang, Z. D., Zhao, Q., and Fox, G. E. (2002) NCIR: A database of non-canonical interactions in known RNA structures, *Nucleic Acids Res. 30*, 395−7.

66. 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.
67. Burkard, M. E., and Turner, D. H. (2000) NMR Structures of r(GCAGGCGUGC)$_2$ and Determinants of Stability for Single Guanosine-Guanosine Base Pairs, *Biochemistry 39*, 11748−62.
68. Mathews, D. H., and Case, D. A. (2006) Nudged Elastic Band Calculation of Minimal Energy Paths for the Conformational Change of a GG Non-canonical Pair, *J. Mol. Biol. 357*, 1683−93.
69. 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*, 12486−97.
70. Znosko, B. M., Burkard, M. E., Krugh, T. R., and Turner, D. H. (2002) Molecular recognition in purine-rich internal loops: Thermodynamic, structural, and dynamic consequences of purine for adenine substitutions in 5′(rGGCAAGCCU)$_2$, *Biochemistry 41*, 14978−87.
71. Wu, M., McDowell, J. A., and Turner, D. H. (1995) A periodic table of symmetric tandem mismatches in RNA, *Biochemistry 34*, 3204−11.
72. 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*, 11349−54.
73. Chen, G., Znosko, B. M., Jiao, X., and Turner, D. H. (2004) Factors affecting thermodynamic stabilities of RNA 3 × 3 internal loops, *Biochemistry 43*, 12865−76.
74. Williamson, J. R. (2000) Induced fit in RNA-protein recognition, *Nat. Struct. Biol. 7*, 834−7.
75. Leulliot, N., and Varani, G. (2001) Current topics in RNA-protein recognition: Control of specificity and biological function through induced fit and conformational capture, *Biochemistry 40*, 7947−56.
76. Lee, B. M., Xu, J., Clarkson, B. K., Martinez-Yamout, M. A., Dyson, H. J., Case, D. A., Gottesfeld, J. M., and Wright, P. E. (2006) Induced Fit and "Lock and Key" Recognition of 5 S RNA by Zinc Fingers of Transcription Factor IIIA, *J. Mol. Biol. 357*, 275−91.
77. Blanchard, S. C., and Puglisi, J. D. (2001) Solution structure of the A loop of 23S ribosomal RNA, *Proc. Natl. Acad. Sci. U.S.A. 98*, 3720−5.
78. Yoshizawa, S., Fourmy, D., and Puglisi, J. D. (1999) Recognition of the codon-anticodon helix by ribosomal RNA, *Science 285*, 1722−5.
79. Nakano, S., Uotani, Y., Uenishi, K., Fujii, M., and Sugimoto, N. (2005) Site-selective RNA cleavage by DNA bearing a base pair-mimic nucleoside, *J. Am. Chem. Soc. 127*, 518−9.
80. Soukup, G. A., and Breaker, R. R. (1999) Relationship between internucleotide linkage geometry and the stability of RNA, *RNA 5*, 1308−25.
81. Luebke, K. J., and Tinoco, I. (1996) Sequence Effects on RNA Bulge-Induced Helix Bending and a Conserved Five-Nucleotide Bulge from the Group I Introns, *Biochemistry 35*, 11677−84.
82. Correll, C. C., Yang, X., Gerczei, T., Beneken, J., and Plantinga, M. J. (2004) RNA recognition and base flipping by the toxin sarcin, *J. Synchrotron Radiat. 11*, 93−6.
83. Yang, X., Gerczei, T., Glover, L. T., and Correll, C. C. (2001) Crystal structures of restrictocin-inhibitor complexes with implications for RNA recognition and base flipping, *Nat. Struct. Biol. 8*, 968−73.
84. Wyatt, J. R., and Tinoco, I., Jr. (1993) RNA structural elements and RNA function, in *The RNA World* (Gesteland, R. F., and Atkins, J. F., Eds.) pp 465−96, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
85. Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Heaphy, S., Karn, J., Lowe, A. D., Singh, M., and Skinner, M. A. (1990) HIV-1 tat protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure, *EMBO J. 9*, 4145−53.
86. Hermann, T., and Patel, D. J. (2000) RNA bulges as architectural and recognition motifs, *Struct. Folding Des. 8*, R47−54.
87. Cheng, X., and Roberts, R. J. (2001) AdoMet-dependent methylation, DNA methyltransferases and base flipping, *Nucleic Acids Res. 29*, 3784−95.
88. O'Gara, M., Horton, J. R., Roberts, R. J., and Cheng, X. (1998) Structures of HhaI methyltransferase complexed with substrates containing mismatches at the target base, *Nat. Struct. Biol. 5*, 872−7.
89. Roberts, R. J., and Cheng, X. (1998) Base flipping, *Annu. Rev. Biochem. 67*, 181−98.
90. Hoang, C., and Ferre-D'Amare, A. R. (2001) Cocrystal structure of a tRNA Psi55 pseudouridine synthase: Nucleotide flipping by an RNA-modifying enzyme, *Cell 107*, 929−39.
91. Hager, J., Staker, B. L., and Jakob, U. (2004) Substrate binding analysis of the 23S rRNA methyltransferase RrmJ, *J. Bacteriol. 186*, 6634−42.
92. Carter, A. P., Clemons, W. M., Jr., Brodersen, D. E., Morgan-Warren, R. J., Hartsch, T., Wimberly, B. T., and Ramakrishnan, V. (2001) Crystal structure of an initiation factor bound to the 30S ribosomal subunit, *Science 291*, 498−501.
93. Blakaj, D. M., McConnell, K. J., Beveridge, D. L., and Baranger, A. M. (2001) Molecular dynamics and thermodynamics of protein-RNA interactions: Mutation of a conserved aromatic residue modifies stacking interactions and structural adaptation in the U1A-stem loop 2 RNA complex, *J. Am. Chem. Soc. 123*, 2548−51.
94. Klimasauskas, S., Kumar, S., Roberts, R. J., and Cheng, X. (1994) HhaI methyltransferase flips its target base out of the DNA helix, *Cell 76*, 357−69.
95. Tok, J. B. H., Bi, L., and Saenz, M. (2005) Specific recognition of naphthyridine-based ligands toward guanine-containing bulges in RNA duplexes and RNA-DNA heteroduplexes, *Bioorg. Med. Chem. Lett. 15*, 827.
96. Schroeder, S. J., Burkard, M. E., and Turner, D. H. (1999) The energetics of small internal loops in RNA, *Biopolymers 52*, 157−67.
97. Lu, X. J., and Olson, W. K. (2003) 3DNA: A software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures, *Nucleic Acids Res. 31*, 5108−21.
98. Lee, J. C., Gutell, R. R., and Russell, R. (2006) The UAA/GAN internal loop motif: A new RNA structural element that forms a cross-strand AAA stack and long-range tertiary interactions, *J. Mol. Biol. 360*, 978−988.

BI0605787