Nucleic Acids Research, 2005, Vol. 33, No. 10 ***3145–3153***
doi:10.1093/nar/gki621

NMR structure and Mg^{2+} binding of an RNA segment that underlies the L7/L12 stalk in the *E.coli* 50S ribosomal subunit

Qin Zhao[1], Uma Nagaswamy[1], Hunjoong Lee[2], Youlin Xia[2], Hung-Chung Huang[1], Xiaolian Gao[1,2] and George E. Fox[1,*]

[1]Department of Biology and Biochemistry and [2]Department of Chemistry, University of Houston, Houston, TX 77204-5001, USA

Received February 11, 2005; Revised and Accepted May 10, 2005

ABSTRACT

Helix 42 of Domain II of *Escherichia coli* 23S ribosomal RNA underlies the L7/L12 stalk in the ribosome and may be significant in positioning this feature relative to the rest of the 50S ribosomal subunit. Unlike the *Haloarcula marismortui* and *Deinococcus radiodurans* examples, the lower portion of helix 42 in *E.coli* contains two consecutive G•A oppositions with both adenines on the same side of the stem. Herein, the structure of an analog of positions 1037–1043 and 1112–1118 in the helix 42 region is reported. NMR spectra and structure calculations support a *cis* Watson–Crick/Watson–Crick (*cis* W.C.) G•A conformation for the tandem $(G•A)_2$ in the analog and a minimally perturbed helical duplex stem. Mg^{2+} titration studies imply that the *cis* W.C. geometry of the tandem $(G•A)_2$ probably allows O6 of G20 and N1 of A4 to coordinate with a Mg^{2+} ion as indicated by the largest chemical shift changes associated with the imino group of G20 and the H8 of G20 and A4. A cross-strand bridging Mg^{2+} coordination has also been found in a different sequence context in the crystal structure of *H.marismortui* 23S rRNA, and therefore it may be a rare but general motif in Mg^{2+} coordination.

INTRODUCTION

Helix 42 in Domain II of the *Escherichia coli* 23S ribosomal RNA (rRNA) (Figure 1a), is the main connecting element between the L7/L12 stalk and the rest of the 50S ribosomal subunit (1). The L7/L12 stalk encompasses the GTPase-associated domain of the 23S rRNA, which has interactions with ribosomal proteins L10, L7/12 and L11 (2–6). The L7/L12 stalk is well known to be dynamic, adopting alternative conformations at different stages in the elongation cycle (7–9). If bending in helix 42 occurred, it would have significant effect on the relative orientation of the L7/L12 stalk and the rest of the 50S subunit. In *E.coli* 23S rRNA, the lower portion of helix 42 contains two tandem G•A juxtapositions (1039/1116 and 1040/1115) in which both adenines are on the same side of the stem. These G•A oppositions are replaced by A–U and G•U in *Haloarcula marismortui*, and U–A and G–C in *Deinococcus radiodurans*. The 50S subunit crystal structures from these other organisms reveal that in both cases, the corresponding bases are paired and the entire stem is a standard A-form helix. In addition, there are no ribosomal proteins in close proximity. Examination of 23S rRNA sequences from other organisms, e.g. *Thermus thermophilus*, *Zea mays* chloroplast and *Bacillus stearothermophilus* revealed that normal Watson–Crick pairs frequently occur in positions corresponding to the tandem $(G•A)_2$ interactions found in *E.coli* (10).

G•A mismatch pairs occur very frequently in RNA and actually are comparable in number in the large ribosomal RNAs to G–U wobble pairs (10–12). Multiple types of G•A mismatch pairs with two hydrogen bonds have been discovered (13–16). These include the *cis* Watson–Crick/Watson–Crick (*cis* W.C.), the *cis* Watson–Crick/Hoogsteen, the *cis* Sugar-edge/Sugar-edge, the *trans* Hoogsteen/Sugar-edge ('sheared' G•A), and the *trans* Sugar-edge/Sugar-edge. Although less common than the sheared G•A, the *cis* W.C. G•A pair plays important structure/function roles in various natural RNAs (15,17–21). For example, the *cis* W.C. G•A pair is essential for the catalysis observed in the autocatalytic genomic hepatitis delta virus (HDV) ribozyme. The deletion of the adenine results in the decreased activity of the genomic ribozyme (15,19). The crystal structure of the P4–P6 domain of the *Tetrahymena* group I intron contains two *cis* W.C. G•A pairs in separate locations, which occur at the interface between helices and other motifs,

*To whom correspondence should be addressed. Tel: +1 713 743 8363; Fax: +1 713 743 8351; Email: fox@uh.edu

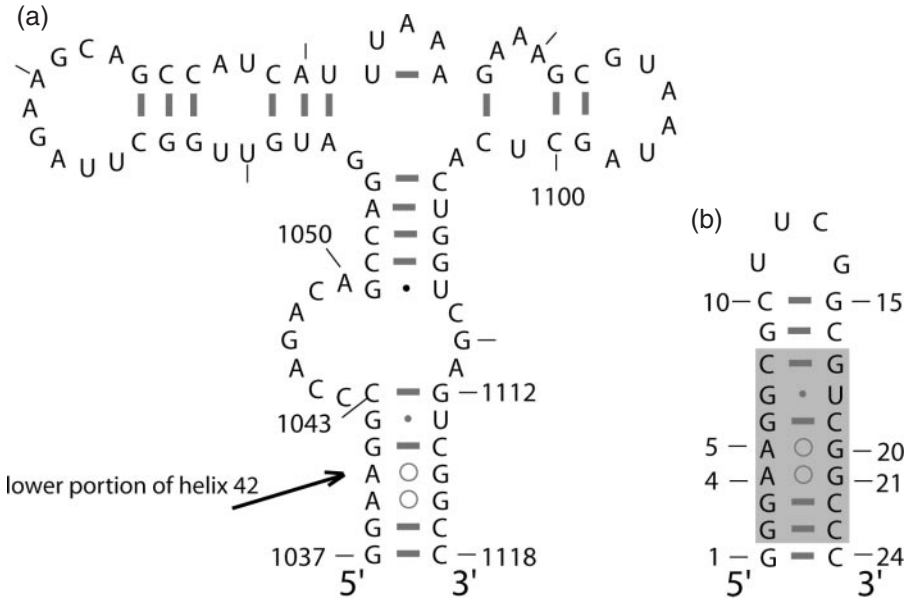

Figure 1. (**a**) The region of *E.coli* 23S rRNA containing helix 42 is illustrated. The lower portion of helix 42 (1037–1043 and 1112–1118) is indicated by an arrow. (**b**) Secondary structure of the 24-mer analog of the lower portion of helix 42. The bases in the shaded region correspond exactly to what occurs in the *E.coli* 23S rRNA.

strongly suggesting a possible functional role of these *cis* W.C. G•A pairs in stabilizing RNA folding (19). Many examples of *cis* W.C. G•A pairs were found in the crystal structure of the 30S ribosomal subunit (16S rRNA) of *T.thermophilus* and the 50S ribosomal subunits (23S rRNA) of the archaeon *H.marismortui* and the bacterium *D.radiodurans* (1,15,17,18). The highly conserved *cis* W.C. G•A base pairs seen in these structures are mostly involved in either tertiary or neighbor contacts in the ribosome (15).

Although the majority of the known G•A mismatch pairs occur as singlets, two or more consecutive GA oppositions such as those seen in helix 42 do occur and multiple examples are found in the ribosomal structures. Many of these are in bulge loops or junction regions rather than being embedded in stem structures. In 18 of the 23 known ribosomal examples (1JJ2.pdb, 1J5E.pdb), both neighboring GA oppositions adopt the sheared base pair conformation and in most cases the adenines are on opposite strands. One of a few examples with the adenines on the same strand involves positions 5′ G816G817 3′/3′ A796A797 5′ in the *H.marismortui* 50S ribosomal subunit. In the remaining five examples, one of the G•A oppositions forms either one or no hydrogen bond while the other is a sheared G•A pair. Tandem sheared $(G•A)_2$ oppositions have also been seen in the group I intron ribozyme (22,23) where they are associated with a change in direction of the backbone that allows creation of a three-helix junction. In contrast, no instances of the 5′ GG 3′/3′ AA 5′ (*cis* W.C.) arrangement were observed in the crystal structures of the 50S (*H.marismortui*, *D.radiodurans*) or 30S (*T.thermophilus*) ribosomal subunits.

Multiple studies of short RNA duplexes containing tandem G•A oppositions have been undertaken previously. In most cases, the adenines were on opposite sides and the resulting pairs were of the sheared type (13,14,24). Although the primary pairing remained the same in all of these examples, the structural details were context dependent being influenced by a change in the 5′ flanking base from U to C (25). When the flanking 5′ base was a G, the structure was dramatically changed with the result that the G•A pairs were now of the *cis* W.C. type resulting in a widening of the backbone (26). The one study to date in which the adenines in the G•A oppositions were on the same strand was a small stem which was in part analogous to the P5abc domain of the group I intron (27) The sequence context in this case differs from that seen in *E.coli* helix 42 in that the 3′ pair is a wobble pair rather than a standard Watson–Crick pair. In this case, the G•A oppositions were of the *cis* W.C. type (27) in the analog but were the more usual sheared configuration in the larger context of the intron. This was not unexpected as the structural context in the intron was dramatically different from that of the analog (23). Instead of a base pair next to the G•A oppositions, the intron helix immediately opens up to form a junction. In total, these results suggested that the G•A oppositions in helix 42 would also be of the *cis* W.C. type, but given the well-established context sensitivity of such interactions this was by no means certain.

Metal ions have long been known to stabilize RNA structure (28) and have been implicated in stabilizing the catalytic core of numerous ribozymes (29). A recent study (30) greatly extended our knowledge in this area by examining the major role played by Mg^{2+} ions in maintaining the structural integrity of the *H.marismotui* 23S rRNA. Four Mg^{2+} ion binding sites were observed in the helix stem of 23S rRNA (1S72.pdb). The metal ion binding was seen to be especially abundant in and around the peptidyl transferase center, where ribosomal proteins are absent, reminding us that metal ions may have been of special significance in a primordial RNA world. In addition to the examples in the ribosome structure, a tandem $(G•A)_2$ in a P5GA hairpin of a group I intron has been observed to form a Mg^{2+} ion binding site (23,27).

In the work described herein, we have used NMR spectroscopy to make a detailed examination of the structure of a

24mer RNA (Figure 1b) that is an analog of positions 1037–1043 and 1112–1118 of helix 42 of *E.coli* 23S rRNA. Traditional NMR strategies were used to obtain indirect evidence of the individual hydrogen bonds, which was confirmed by the direct identification of hydrogen bonds in the tandem $(G \bullet A)_2$ via cross hydrogen bond scalar couplings between the ^{15}N of the acceptor in A and the imino ^{15}N of the donor in G (31,32). Additionally, heteronuclear NMR spectroscopy was used to probe the Mg^{2+} binding sites through ^{1}H and ^{15}N chemical shift changes and the backbone dynamics of the RNA hairpin including the tandem $(G \bullet A)_2$ were characterized by ^{15}N NMR relaxation studies. An analysis of the dynamic characteristics of the tandem $(G \bullet A)_2$ pairs and Mg^{2+} binding provide insights to the possible role of this structurally important stem in the 23S rRNA.

MATERIALS AND METHODS

RNA sample preparation

The 24mer RNA oligonucleotide 5′-GGGAAGGCGCUUCG-GCGUCGGCCC-3′ was synthesized by *in vitro* transcription using T7 RNA polymerase and a synthetic DNA template. Unlabeled RNA samples were prepared from 20 ml transcription reactions using 4 mM 5′-NTPs as described elsewhere (33–36). Aliquots (1 mM) of ^{13}C- and ^{15}N-labeled ATP and GTP (Silantes) were used to uniformly label the RNA samples in 20 ml transcription reactions (37). The RNA samples were purified by electrophoresis on 20% polyacrylamide denaturing gels and eluted with RNA elution buffer (0.5 M NH_4OA_C and 1 mM EDTA, pH 7.0). The samples were desalted with a Sep-pak C-18 cartridge (Waters Inc.) and dialyzed against an NMR buffer containing 10 mM NaH_2PO_4/Na_2HPO_4, 0.1 mM EDTA, pH 6.5 using an Amicon-3 concentrator (Millipore) (38). For exchangeable proton measurements, 8% D_2O was used to provide the lock signal. For non-exchangeable proton measurements, the samples were lyophilized three times and then redissolved in 99.96% D_2O. The unlabeled sample was then dissolved in 0.50 ml of either 92% H_2O: 8% D_2O or 99.96% D_2O (Cambridge Isotope Laboratories, Inc.) to a final concentration of 1.4 mM. The uniformly labeled sample was dissolved in 0.25 ml in a 5-mm Shigemi tube with a final concentration of 1.4 mM.

NMR spectroscopy

NMR spectra were acquired on a Bruker Avance 600 MHz NMR spectrometer. All NMR data were processed and analyzed with the programs XWINNMR (Bruker Instruments) and Sparky (39). One-dimensional (1D) melting profiles and 2D NOESY spectra in 92% H_2O/8% D_2O were collected to study the exchangeable protons in the temperature range 0–25°C using jump-return pulses (40) to suppress the water peak. The samples were analyzed in the 99.96% D_2O to observe non-exchangeable protons. The 2D NOESY (50, 100, 250 and 500 ms mixing times), TOCSY with 80 ms mixing time were acquired in the temperature range 12–35°C. DQF-COSY spectra were recorded at 25 and 35°C using a repetition delay of 4.2 s. A proton-detected heteronuclear ^{31}P–^{1}H COSY spectrum was acquired with sweep width 8389.3 Hz in the ^{1}H dimension and 458.8 Hz in the ^{31}P dimension at 25°C.

Experiments on uniformly labeled RNA samples were performed in 92% H_2O/8% D_2O at both 12°C and 25°C. ^{13}C–^{1}H HSQC spectra were acquired with the ^{13}C carrier at 143.0 p.p.m., the spectral width in the ^{13}C dimension of 3773 Hz and the spectral width in the ^{1}H dimension of 8389 Hz. ^{15}N–^{1}H HSQC spectra were acquired with the ^{15}N carrier at 147.0 p.p.m. The spectral width was set to 912 Hz in the ^{15}N dimension and 12019 Hz in the ^{1}H dimension. The quantitative J_{NN} HNN-COSY experiments were acquired with a total of 128 scans per FID. The spectral width was 6081 Hz in the ^{15}N dimension and 12019 Hz in the ^{1}H dimension. The experiments were performed with a ^{1}H carrier positioned on the H_2O resonance, the ^{15}N carrier at 185.0 p.p.m. and ^{13}C carrier at 153.0 p.p.m. A mixing time of 1 s was applied for HNN-COSY transfer.

The effect of Mg^{2+} on the 24mer RNA fragment was monitored by chemical shift measurements from the ^{1}H–^{15}N HSQC spectra of the uniformly labeled RNA in 92% H_2O/8% D_2O and the 1D NOESY spectra of the non-labeled sample in 99.96% D_2O upon the addition of 0, 5, 10, 15, 20, 25, 30, 35 and 40 mM Mg^{2+}. In order to avoid sample dilution during the Mg^{2+} titration, 5 mM $MgCl_2$ aliquots were dried and then the RNA samples were sequentially combined with each of the nine aliquots to increase the Mg^{2+} concentration from 0 to 40 mM, while keeping the sample volume at 0.5 ml (unlabeled sample) or 0.25 ml (labeled sample). The pH of the sample was 6.5 during the Mg^{2+} titration.

^{15}N T_1 and T_2 and ^{1}H–^{15}N NOE spectra were acquired at 25°C with the pulse sequences as described previously (41,42). ^{15}N T_1 values were measured from the spectra recorded with seven different relaxation delays: 0.05, 0.1, 0.2, 0.3, 0.4, 0.6, 0.8 s. T_2 values were calculated from spectra recorded with seven different delays: 0.016, 0.047, 0.079, 0.110, 0.142, 0.173, 0.205 s. ^{1}H–^{15}N NOE experiments were performed with the repetition delay of 4.0 s. A repetition delay of 2.0 s was employed in the measurement of ^{15}N T_1 and T_2 values. The standard deviation values of the order parameters and phenomenological exchange term were indicated by the sim error (43).

NMR constraints and structure calculations

Structure calculations utilized procedures described previously (38). The cross-peaks were assigned for all the 2D NOESY spectra in D_2O (50 ms, 250 ms and 500 ms mixing times) and H_2O (200 ms mixing time) (Supplementary Material Table S1). Distance restraints were derived from the cross-peak intensities by categorizing the NOEs as strong (1.8–3.2 Å), medium (2.2–4.2 Å) or weak (3.6–5.0 Å). These ranges were based on comparisons with cross-peaks associated with the known C H5–H6 distances.

Torsion angle values were derived from the analysis of ^{31}P–^{1}H COSY, DQF-COSY and NOESY spectra. Residues (U12, C13) with $^{3}J_{H1'-H2'}$ couplings > 3 Hz and $^{3}J_{H3'-H4'}$ coupling < 7–8 Hz were restrained to a δ torsion angle of 140° ± 20° corresponding to C2′-*endo* conformation. Other residues were in the range of a C3′-*endo* conformation corresponding to a δ torsion angle of 80° ± 20°. Chi (−160° ± 20°) is correlated with C3′-*endo* and (−130° ± 15°) with C2′-*endo*. The residue G15 with a downfield phosphorus resonance was

Table 1. Statistics of NMR restraints (47) in the structure calculation and RMSD values among the final converged structures

A. Distance and dihedral restraint statistics	
(1) Distance restraints	
Intra-residue	316
Inter-residue	211
Watson–Crick	23
Mean number per residue	23
(2) Dihedral restraints	
Total NMR derived number	120
Mean number per residue	5
B. Superimposition of structures	
(1) Violations	Number of violations
Bond > 0.10 Å	0
Angle > 5.0°	0
Improper > 15.0°	0
Dihedral > 20.0°	0
Vdw >1.7	0
NOE > 0.3 Å	0
(2) RMSD of residues	(Å)
9 structures (all nt)	0.85 ± 0.50
9 structures (stem)	0.77 ± 0.38
9 structures (loop)	0.29 ± 0.09

restrained to the torsion angles as reported previously (44–46). With the exception of the tandem $(G\bullet A)_2$, the backbone torsion angles of the other stem residues were loosely constrained to the range of A-form parameters ($\alpha = -68° \pm 15°$, $\beta = 178° \pm -15°$, $\gamma = 54° \pm 15°$, $\varepsilon = -153° \pm 15°$ and $\zeta = -71° \pm 15°$) (38).

The restrained molecular dynamics (rMD) simulated annealing protocol in the Crystallography and NMR System (CNS) package was used to determine the structure (47). A total of 550 distance restraints, 120 dihedral restraints (Table 1) and 11 planarity restraints were incorporated into rMD followed by energy minimization with an all-atom force field (47). The initial structures were generated in extended strands in random dihedral angles. After global folding and refinement, the converged structures were accepted on the basis of low total energy, the avoidance of bond violation >0.10 Å, torsion angle violation >5.0°, improper violation >15.0°, dihedral violation >20° (38), vdw violation >1.7 and NOE distance violations >0.3 Å. Structures were viewed with VMD (48) and analyzed with X3DNA (49).

Backbone dynamics analysis

R_1 and R_2 values of ^{15}N-labeled G residues were determined by fitting exponential peak height decay in a series of ^{15}N T_1 and T_2 spectra respectively using the program Sparky (39). The average values of T_1 and T_2 are 0.47 and 0.14 s, respectively. The relaxation parameters of ^{15}N of the G residues were derived. The values of R_1, R_2, NOE, S^2, and R_{ex} are shown in Supplementary Material Table S2. The phenomenological exchange term R_{ex} indicated the contribution of chemical exchange processes to the decay of transverse magnetization through the relaxation measurements. The extent of picosecond-time scale motions were determined by the order parameter S^2. The generalized order parameters S^2 and phenomenological exchange term R_{ex} were calculated with Modelfree 4.0 (43) using the Lipari–Szabo approach (50,51). A grid search of internal motional parameters was performed before optimization of the overall diffusion model. Brent's implementation of the Powell's method was applied to the optimization of the axially symmetric diffusion tensor for all spins (43). A simple spectral density function including S^2 and R_{ex} (52) was selected to fit the experimental NMR transverse relaxation data for the extraction of the motional parameters. The average value of the order parameter S^2, over all residues except those in the tandem $(G\bullet A)_2$ region, is 0.78. The order parameters S^2 of G20 and G21 are 0.78 and 0.91, respectively. In contrast to the neighboring residues, residue G21 exhibits higher R_{ex} contribution to the transverse relaxation rate, arising from conformational processes, of $2.58 \pm 0.643\ s^{-1}$.

RESULTS

Spectral analysis

Assignment of exchangeable protons. 2D NOESY spectra at 5°C in 92% H_2O/8% D_2O are used to assign exchangeable protons. Eleven imino proton resonances are shown in the 1D proton spectrum of the 24mer RNA sample. The starting point for imino proton assignments was established by identification of the G7•U18 wobble pair, which provides a very strong cross-peak between the G7 imino and U18 imino protons in the 2D NOESY spectra. The sequential walk was evident based on the characteristic downfield imino resonance of U18 and the upfield shifted G7 imino. G6 was identified by the cross-peak between G6 and G7 imino protons. G17 was determined by the cross-peak between G17 and U18. The identification of the imino protons of G9, G2 and G3 was facilitated by the observation of cross-peaks between G(H1)–C(NH2) and C(NH2)–C(H5). This is the characteristic pattern associated with Watson–Crick G–C base pairs.

Among the remaining imino peaks, two resonated in the region between 11.0 and 12.0 p.p.m. This is the typical chemical shift range for the guanine imino protons that are found in *cis* W.C. G•A mismatches (14). The imino peak at 11.2 p.p.m. had a cross-peak to G3 imino proton and therefore was assigned to G21. The other non-canonical base-pairing pattern was assigned to residue G20. The intense cross-peaks between A5 H2 and the G20 amino group are assigned unambiguously. Cross-peaks between A4 H2 and the G21 amino group were also observed. This indicates that both G•A base pairs are of the *cis* W.C. G•A type.

Assignment of non-exchangeable protons. The 2D-NOESY and DQF-COSY spectral analysis confirmed formation of a hairpin by the 24mer RNA. Supplementary Material Figure S1 shows the expanded base (H8/H6/H2) to sugar (H1′/H5) region of the NOESY spectrum recorded with 500 ms mixing time at 25°C. All the signature spectral features of the well-characterized UUCG tetraloop (44–46) were observed in the 2D spectra. The sequential connectivities are absent at the loop residues U11, U12, G14 and G15. In particular, G14 adopts a *syn* conformation with the very strong H1′–H2′ and intense intranucleotide H8–H1′ cross-peaks even in the 50 ms 2D NOESY spectra. The H1′ of the adjacent G15 is unusually upfield shifted (44–46). The sequential NOEs of residues 2 through 10 and 15 through 23 are consistent with the duplex stem formation, except that the internucleotide cross-peaks of A4H1′/A5H8 and G21H1′/C22H6 are absent. This indicates

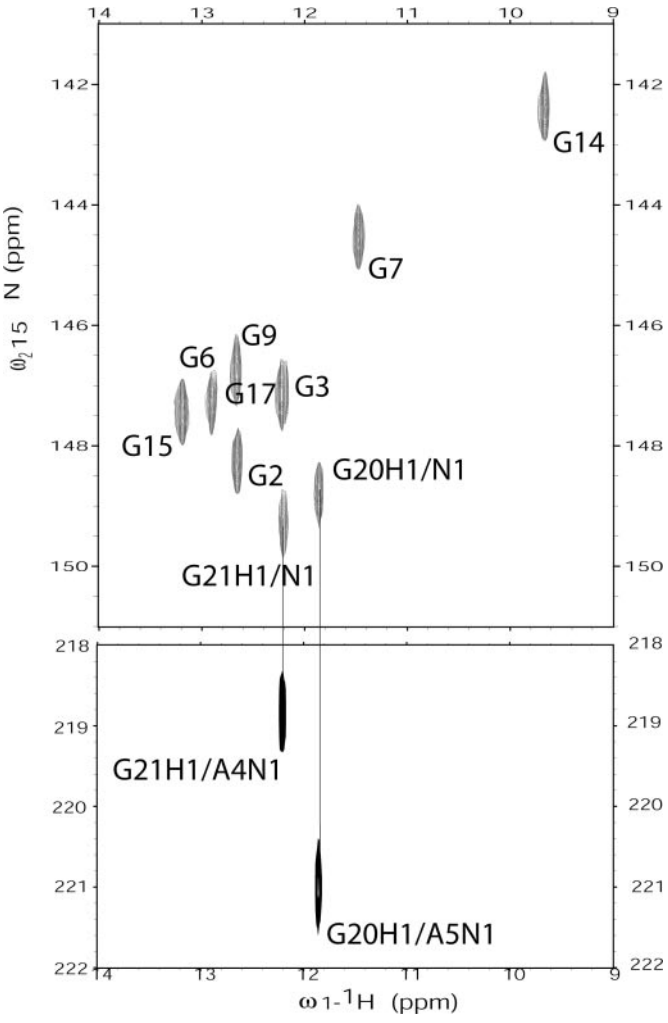

Figure 2. Direct observation of a $^{2h}J_{NN}$ coupling of a G•A base pair in HNN-COSY spectrum at 12°C. The cross-correlations between H1 of G and N1 of A from G20•A5 and G21•A4 are shown with solid lines.

local structural alteration at the tandem $(G•A)_2$ site (27). However, the presence of coupling cross-peaks interstrand G(N1)–G(H1)–A(N1) correlations in HNN-COSY spectrum (Figure 2) unambiguously reveals the form of *cis* W.C. G•A pairs.

Assignments of ^{31}P resonances. Nineteen chemical shifts of ^{31}P were assigned with chemical shift dispersing over a range of 2.57 p.p.m. [from 1.76 (G15) to −0.81 p.p.m. (C13)]. Most large deviations of the ^{31}P chemical shifts are from the loop residues or those at the junction of the loop to stem; the ^{31}P resonances of G3–A5 and of C19–G21 appeared at a narrow chemical shift range (0.8 p.p.m.), indicating the tandem $(G•A)_2$ adopts a regular helical backbone conformation.

Backbone dynamics. Backbone dynamics measurements and modeling results (Supplementary Material Figure S3 and Table S2) did not reveal any clear abnormalities in relaxation, the order parameters or the phenomenological exchange term. This indicates that the tandem $(G•A)_2$ does not induce additional amplitude motions at a picosecond to nanosecond time scale. The small variations observed in S^2 for the 5′-terminal residues G2 and G3 do not provide sufficient information for isolation of local dynamics.

Specific spectral features of the tandem $(G•A)_2$

The tandem $(G•A)_2$ exhibits well-defined spectral patterns. The imino protons of G20 and G21 were clearly observed and they can be linked to A(N1) through the measurement of J_{NN} in the heteronuclear correlation spectrum. The amino protons of both G residues are in close contact with the corresponding H2 of the A residues. These spectral observations provide evidence for a well-defined *cis* W.C. type of tandem $(G•A)_2$ with the two G and A residues in opposite strands (Figure 3a). Stacking between the two G•A pairs was observed from the adjacent imino protons of G residues to the adjacent H8 protons of G and A residues, establishing that the *cis* W.C. type of tandem $(G•A)_2$ pairs are well-accommodated in the hairpin stem helix.

Overall structure

An overlay of the nine alternative structures generated is shown in Figure 3b and the structural calculation statistics are provided in Table 1. The global structure of the converged hairpin stem generally resembles the standard A-form geometry. The UUCG loop is very similar to that described previously (44–46), with the U11•G14 pair stabilized by base–base and base–sugar hydrogen bonds and the *syn* conformation of G14 as well as C2′-*endo* sugar puckering of U12 and C13. The unusually upfield shifted H1′ resonance of residue G15 is attributed to the ring current effects of the *syn* base of residue G14 (44). The overall average helical twist of the stem region is 29.9° which is close to the 32.7° (24) associated with the standard A-form (Supplementary Material Table S3). The stem region contains a G•U wobble pair, which exhibits unwinding at the step 5′ to G but overwinding at the step 3′ to G (Supplementary Material Table S3, 20.5° versus 42.2°). The tandem $(G•A)_2$ region deviates from the average helical twist accounting in part for the reduced overall average. Specifically, the helical twists of A4G21–A5G20, A5G20–G6C19 are 25.3° and 28.1°, respectively, suggesting an underwinding between these *cis* W.C. G•A mismatches compared to the A-form helix (24). This underwinding of the helix at the mismatches is associated with the intrastrand stacking of the *cis* W.C. G•A mismatches. In contrast, G3C22–A4G21 is clearly overwound with a twist 38.8°. The twist angle of G6C19–G7U18 is 20.5°, reveals considerable underwinding, which could reflect local distortions from the neighboring G•U wobble pair.

Mg^{2+} binding in the tandem $(G•A)_2$

Chemical shifts are sensitive to metal ion binding and its induced structural changes. The potential binding of Mg^{2+} to the 24mer RNA hairpin was investigated by monitoring the ^1H and ^{15}N chemical shifts of the G and A residues. The induced changes in the chemical shifts of the imino-N1, imino-H of guanines and H8/H6 protons of all the residues were monitored and plotted as a function of increasing concentration of Mg^{2+} up to 40 mM (Supplementary Material Figure S2). The observed patterns of chemical shift variation

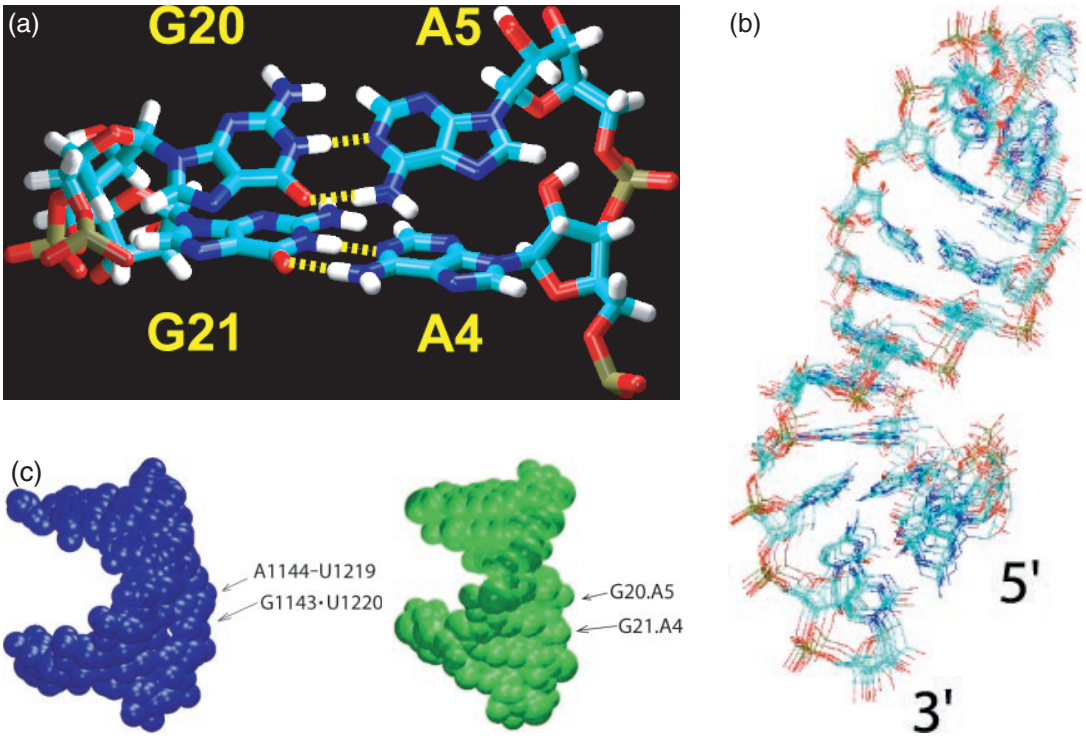

Figure 3. Solution structures of the RNA analog determined by NMR, distance geometry and restrained molecular dynamics. (**a**) Expanded view of the tandem $(G\bullet A)_2$ region containing A4, A5, G20 and G21. (**b**) Superposition of nine converged structures of the RNA molecule on the average structure. A4, A5, G20 and G21 orientation faces the major groove. (**c**) The equivalent fragment of the conserved *E.coli* 23S rRNA region of the solution structure (green), the *H.marismortui* crystal structure solved at 2.4 Å resolution (blue). The residues of tandem $(G\bullet A)_2$ and corresponding base pairs in *H.marismortui* are indicated by the arrows.

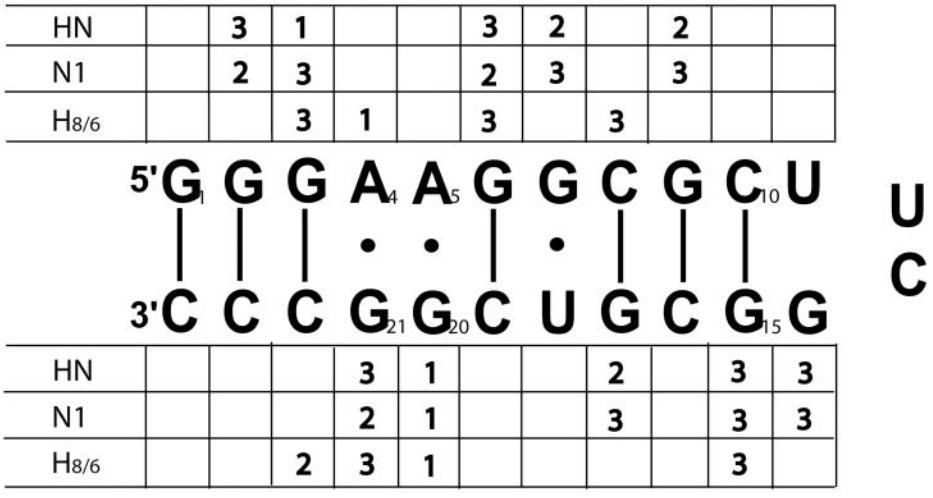

Figure 4. Mg^{2+} induced changes in the chemical shift of adenine N1, guanine H1 and H8/H6 of A, C, G residues. The symbols are 1 for the most affected, 2 for the moderately affected and 3 for the least affected.

for the imino nitrogens and protons are summarized in Figure 4. The majority of the net chemical shift changes, though subtle, are shown in clearly defined trends (Supplementary Material Figure S2). While the resonances of G20 are associated with the largest (Figure 4, marked with '1') upfield shift with increasing Mg^{2+} concentration, other residues around the tandem $(GA)_2$, i.e. the chemical shifts of G3NH, G20NH, G20N1, G20H8 and A4H8, were also affected by the addition of Mg^{2+} (Figure 4). These resonance perturbations relate to major groove functional groups of residues central to the hairpin stem, suggesting the localization of the Mg^{2+} ion in the major groove and at the vicinity of the tandem $(G\bullet A)_2$ pair (27). Mg^{2+} binding is commonly found near the phosphate groups of ribosomal RNA (30), and when it occurs the interaction tends to be non-specific and mediates RNA–protein complex formation. Our results do not exclude this binding mode of Mg^{2+} in the 24mer RNA.

We further examined possible modes of Mg^{2+} binding by docking the metal ion at the *cis* W.C. type of tandem $(G\bullet A)_2$ site (Figures 5a and 6b). The alignment of the G and A bases

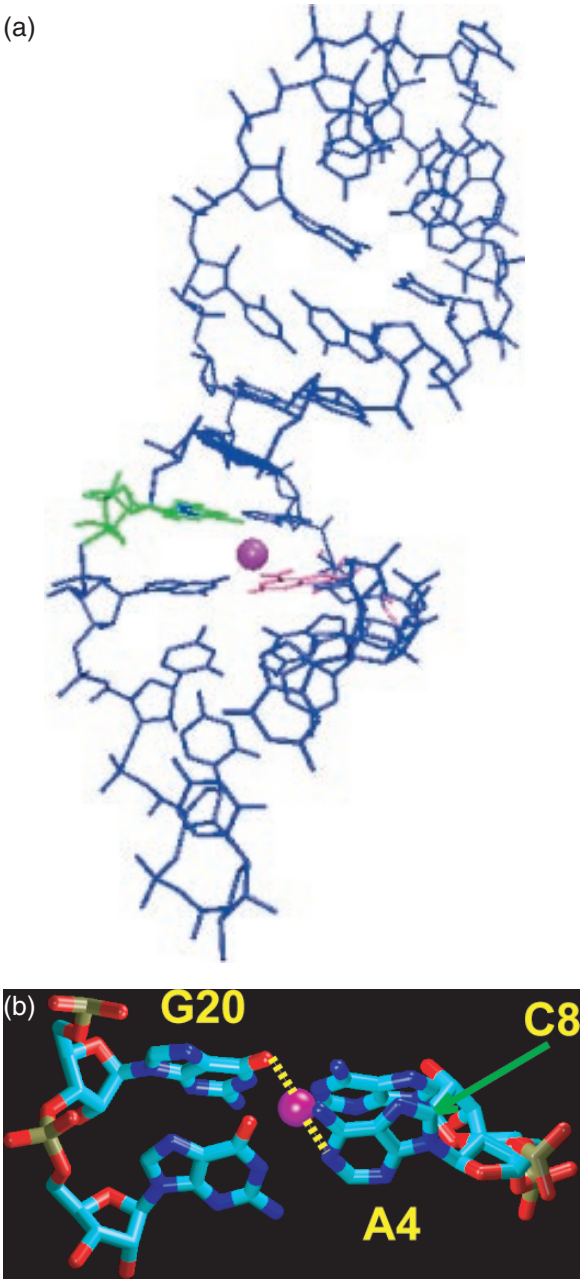

Figure 5. (**a**) Tandem $(G \bullet A)_2$ of the RNA hairpin functions as the pocket for Mg^{2+} (purple). Residues A4 and G20 are shown in pink and green, respectively. (**b**) G20O6 and A4N1 provide the coordination site with Mg^{2+} (purple) via gold bonds. C8 is shown by a green arrow.

with their Watson–Crick edge forming hydrogen bonds presents juxtaposed electron donors: O6 from G20 and N1 from A4. The distance (2.3 Å) between these ligands and the Mg^{2+} suggest that they are interacting with the inner-sphere (30). The close contacts of these atoms with the metal ion would be expected to introduce electron inductive and conjugation effects in the nucleobase ring system, giving rise to the observed chemical shift changes, while the resonances in other regions of the 24mer RNA hairpin are essentially insensitive to the addition of Mg^{2+}. The divalent metal ion binding bridging the cross-strand G and A residues would require two consecutive base pairs. The geometry of the tandem $(G \bullet A)_2$

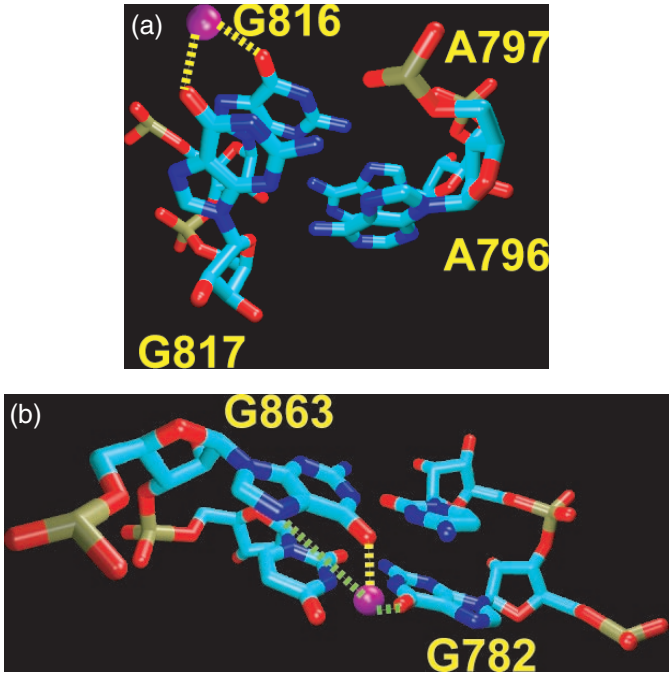

Figure 6. (**a**) G816O6 and G817O6 of *H.marismortui* 23S rRNA provide the inner-sphere coordination sites (gold bonds) with Mg^{2+} (purple). (**b**) G863O6, G863N7 and G782O6 from the consecutive pairs of *H.marismortui* 23S rRNA coordinate with Mg^{2+} (purple). The inner sphere ligation is shown by gold, while the outer sphere ligation is indicated by green.

region determined here is consistent with the proposed Mg^{2+} binding site. The interphosphate distance of the tandem $(G \bullet A)_2$ region (between 18 and 20 Å) is more open than the normal distance of an A-form helix. This wider major groove allows the tandem $(G \bullet A)_2$ region to present a larger surface for Mg^{2+} binding and potentially for auxiliary water to participate as an outer-sphere ligand in stabilizing the metal ion coordination. The major groove of the tandem G•A sequences is shown to be electron negative (30), offering additional attractive force for the binding of positively charged ions. The tandem $(G \bullet A)_2$ region described in this work provides an additional example of Mg^{2+} binding in an A-form helix.

Comparison of the NMR structure with 23S rRNA crystal structures

From a comparative biology perspective, the *E.coli* version of helix 42 is expected to closely resemble the equivalent regions in other 23S rRNAs despite the presence of the tandem $(G \bullet A)_2$. In order to examine this, the NMR structure determined here and the corresponding regions in the crystal structures of the *H.marismortui* (1JJ2.pdb) and *D.radiodurans* (1NKW.pdb) 50S ribosomal subunits were superimposed. The NMR structure has an RMSD value of 2.64 Å with *H.marismortui* and 2.80 Å with *D.radiodurans,* whereas the two crystal structures differ with an RMSD value of 1.67 Å. Figure 3c compares the overall geometry of the NMR structure with the equivalent region in *H.marismortui*. The typical structural features of RNA, a very deep major groove and a rather shallow minor groove are clearly observed in the corresponding regions in both structures. The NMR structure

has clearly adapted what is essentially an A-form geometry despite the presence of the tandem $(G \bullet A)_2$ as confirmed by examination of the helical parameters (Supplementary Material Tables S3 and S4). The geometry of the G•A mismatches determined by NMR is also consistent with the primary sequence exchanges between GA and W.C. oppositions that are seen in the primary sequences of the *E.coli* and *H.marismortui* 23S RNAs.

DISCUSSION

It has been unequivocally shown that the two G•A oppositions in the analog of the lower portion of helix 42 in *E.coli* 23S rRNA form pairs of the *cis* W.C. type in solution. In view of the comparative sequence data and the number of canonical Watson–Crick base-pairs flanking the two G•A oppositions, it is virtually certain that the structure seen here will be essentially the same as that is found in the 23S rRNA. Thus, the lower portion of helix 42 in *E.coli* 23S rRNA represents the first known occurrence of tandem $(G \bullet A)_2$ of the *cis* W.C. type embedded in an otherwise normal stem structure in a naturally occurring functional RNA.

The *cis* W.C. G•A mismatch is frequently found at the ends of canonical helices where it provides an interface between the helix and other structural motifs (15). However, in the present case it is located within a canonical helix accompanied by an equally unusual Mg^{2+} binding site. The *cis* W.C. geometry of the tandem $(G \bullet A)_2$ allows N1 of A4 to coordinate with the Mg^{2+} while the Mg^{2+} serves as a bridge between G20 and A4 via O6 of G20 on the other side. This Mg^{2+} binding duplex motif allows the Mg^{2+} to locate in the major groove of the helix.

The vast majority of Mg^{2+} binding sites in the best-studied case, *H.marismortui* 23S rRNA, occur at helical bulges and interhelical junctions. For example, there is one Mg^{2+}-binding site at an interhelical junction in *H.marismortui* 23S rRNA that consists of non-canonical tandem $(G \bullet A)_2$ with G residues on same side. The two O6 atoms serve as inner-sphere coordination sites for the Mg^{2+} (Figure 6a). On the opposite side, neither A796 nor A797 is involved because the sheared conformation of the tandem $(G \bullet A)_2$ shields them from the Mg^{2+} site. Thus, despite the sequence similarity the Mg^{2+} coordination in this case is quite different than seen in the structure studied here.

Only 4 of 116 Mg^{2+} ions were found to bind to residues in helical stem regions in the large ribosomal subunit of *H.marismortui* (30). In just one case (Figure 6b), the Mg^{2+} has a bridge role similar to that seen in the helix 33 analog. In this instance, G863–C783 and G782•U864 are located in the major groove of the A-form helix with the Mg^{2+} bridging the cross-stand residues G863 and G782. The inner-sphere ligand for Mg^{2+} is from O6 atom of G863 and O4 atom of U864 as well (figure not shown). O6 atom of G782 provides outer-sphere ligand for Mg^{2+} while N7 of G863 is simultaneously coordinated to the Mg^{2+} as another outer-sphere ligand. The result is a structure with isosteric similarity to the Mg^{2+} binding site proposed here for the *E.coli* helix 42 analog. This suggests that the cross-strand bridge may be a general but rare motif in Mg^{2+} coordination.

The rare occurrence of the 5′ GG/3′ AA (*cis* W.C.) tandem interaction and its accompanying Mg^{2+} coordination suggests that the structure obtained here may be uniquely suitable for inclusion in helix 42 of 23S rRNA. The obvious question is 'what functional significance if any this unusual stem structure has?'. The structural comparisons (Figure 3c) show that the inclusion of the two imino pairs minimally disrupts the overall structure. The backbone dynamics reveal that residue G21 (*E.coli* position 1116) has far more potential for slow conformational change than the surrounding residues. Thus, one possibility is that the non-canonical pairs may facilitate flexibility in this stem. Consistent with this, the equivalent positions in *H.marismortui* (but not *D.radiodurans*) are relatively weak pairs. Metal ion binding sites are frequently found to be important in stabilizing RNA structures required for specific protein recognition in the large ribosomal subunit (30). However, if the *H.marismortui* 50S subunit crystal structure is representative, the lower portion of helix 42 may not be in close proximity to any proteins. It is not unreasonable to speculate that as a result of one or more of the conformational changes known to occur in the L7/L12 stalk, this situation changes. Thus, the importance of the tandem $(G \bullet A)_2$ may be that it provides flexibility which is needed to accommodate the functionally important conformational changes that occur in the L7/L12 stalk. It must, nevertheless, be emphasized that this analog structure was investigated in the absence of the ribosomal proteins and therefore may lack subtle structural rearrangements upon binding of nearby ribosomal proteins.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

We thank the W.M. Keck Foundation which partially funds the 600 MHz NMR spectrometer at the University of Houston. The Institute of Molecular Design and the W.M. Keck Center provided computer resource support. The work was supported by grants from the Robert A. Welch Foundation (E-1451) and the National Aeronautics and Space Administration (NAG5–12366) to G.E.F, and the Robert A. Welch Foundation (E-1270) and NIH/GM (RO149957) to X.G. Funding to pay the Open Access publication charges for this article was provided by indirect cost funds.

Conflict of interest statement. None declared.

REFERENCES

1. 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**, 905–920.
2. Ostergaard,P., Phan,H., Johansen,L.B., Egebjerg,J., Ostergaard,L., Porse,B.T. and Garrett,R.A. (1998) Assembly of proteins and 5S rRNA to transcripts of the major structural domains of 23 S rRNA. *J. Mol. Biol.*, **284**, 227–240.
3. Klein,D.J., Schmeing,T.M., Moore,P.B. and Steitz,T.A. (2001) The kink-turn: a new RNA secondary structure motif. *EMBO J.*, **20**, 4214–4221.
4. Baranov,P.V., Gurvich,O.L., Bogdanov,A.A., Brimacombe,R. and Dontsova,O.A. (1998) New features of 23S ribosomal RNA folding: the long helix 41–42 makes a ''U-turn'' inside the ribosome. *RNA*, **4**, 658–668.

5. Mohr,D., Wintermeyer,W. and Rodnina,M.V. (2002) GTPase activation of elongation factors Tu and G on the ribosome. *Biochemistry*, **41**, 12520–12528.
6. 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.
7. Gonzalo,P. and Reboud,J.P. (2003) The puzzling lateral flexible stalk of the ribosome. *Biol. Cell*, **95**, 179–193.
8. Zhao,Q., Ofverstedt,L.G., Skoglund,U. and Isaksson,L.A. (2004) Morphological variation of individual *Escherichia coli* 50S ribosomal subunits in situ, as revealed by cryo-electron tomography. *Exp. Cell Res.*, **300**, 190–201.
9. Christodoulou,J., Larsson,G., Fucini,P., Connell,S.R., Pertinhez,T.A., Hanson,C.L., Redfield,C., Nierhaus,K.H., Robinson,C.V., Schleucher,J. *et al.* (2004) Heteronuclear NMR investigations of dynamic regions of intact *Escherichia coli* ribosomes. *Proc. Natl Acad. Sci. USA*, **101**, 10949–10954.
10. Noller,H.F. (1984) Structure of ribosomal RNA. *Annu. Rev. Biochem.*, **53**, 119–162.
11. Cech,T.R. and Bass,B.L. (1986) Biological catalysis by RNA. *Annu. Rev. Biochem.*, **55**, 599–629.
12. Traub,W. and Sussman,J.L. (1982) Adenine–guanine base pairing ribosomal RNA. *Nucleic Acids Res.*, **10**, 2701–2708.
13. Jang,S.B., Baeyens,K., Jeong,M.S., SantaLucia,J.Jr, Turner,D. and Holbrook,S.R. (2004) Structures of two RNA octamers containing tandem G.A base pairs. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 829–835.
14. 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.
15. 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**, 967–978.
16. Leontis,N.B., Stombaugh,J. and Westhof,E. (2002) The non-Watson–Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res.*, **30**, 3497–3531.
17. Wimberly,B.T., Brodersen,D.E., Clemons,W.M.Jrz, Morgan-Warren,R.J., Carter,A.P., Vonrhein,C., Hartsch,T. and Ramakrishnan,V. (2000) Structure of the 30S ribosomal subunit. *Nature*, **407**, 327–339.
18. 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.
19. Wadkins,T.S., Shih,I., Perrotta,A.T. and Been,M.D. (2001) A pH-sensitive RNA tertiary interaction affects self-cleavage activity of the HDV ribozymes in the absence of added divalent metal ion. *J. Mol. Biol.*, **305**, 1045–1055.
20. Siwkowski,A., DeYoung,M.B., Anderson,P. and Hampel,A. (1997) Mutagenesis and modeling of the hairpin ribozyme family. *Methods Mol. Biol.*, **74**, 357–363.
21. Siwkowski,A., Shippy,R. and Hampel,A. (1997) Analysis of hairpin ribozyme base mutations in loops 2 and 4 and their effects on cis-cleavage *in vitro*. *Biochemistry*, **36**, 3930–3940.
22. Cate,J.H., Gooding,A.R., Podell,E., Zhou,K., Golden,B.L., Szewczak,A.A., Kundrot,C.E., Cech,T.R. and Doudna,J.A. (1996) RNA tertiary structure mediation by adenosine platforms. *Science*, **273**, 1696–1699.
23. 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**, 1678–1685.
24. SantaLucia,J.Jr and Turner,D.H. (1993) Structure of (rGGCGAGCC)2 in solution from NMR and restrained molecular dynamics. *Biochemistry*, **32**, 12612–12623.
25. 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**, 147–158.
26. Wu,M. and Turner,D.H. (1996) Solution structure of (rGCGGACGC)2 by two-dimensional NMR and the iterative relaxation matrix approach. *Biochemistry*, **35**, 9677–9689.
27. Rudisser,S. and Tinoco,I.Jr (2000) Solution structure of Cobalt(III)hexammine complexed to the GAAA tetraloop, and metal-ion binding to G.A mismatches. *J. Mol. Biol.*, **295**, 1211–1223.
28. Draper,D.E. (2004) A guide to ions and RNA structure. *RNA*, **10**, 335–343.
29. Lilley,D.M. (2003) The origins of RNA catalysis in ribozymes. *Trends Biochem. Sci.*, **28**, 495–501.
30. Klein,D.J., Moore,P.B. and Steitz,T.A. (2004) The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA*, **10**, 1366–1379.
31. Hennig,M. and Williamson,J.R. (2000) Detection of N-H. . .N hydrogen bonding in RNA via scalar couplings in the absence of observable imino proton resonances. *Nucleic Acids Res.*, **28**, 1585–1593.
32. Wohnert,J., Dingley,A.J., Stoldt,M., Gorlach,M., Grzesiek,S. and Brown,L.R. (1999) Direct identification of NH. . .N hydrogen bonds in non-canonical base pairs of RNA by NMR spectroscopy. *Nucleic Acids Res.*, **27**, 3104–3110.
33. Liu,S.W. and Milman,G. (1983) Purification and characterization of *Escherichia coli* guanine-xanthine phosphoribosyltransferase produced by a high efficiency expression plasmid utilizing a lambda PL promoter and CI857 temperature-sensitive repressor. *J. Biol. Chem.*, **258**, 7469–7475.
34. Rasmussen,U.B., Mygind,B. and Nygaard,P. (1986) Purification and some properties of uracil phosphoribosyltransferase from *Escherichia coli* K12. *Biochim. Biophys. Acta*, **881**, 268–275.
35. Andersen,P.S., Smith,J.M. and Mygind,B. (1992) Characterization of the upp gene encoding uracil phosphoribosyltransferase of *Escherichia coli* K12. *Eur. J. Biochem.*, **204**, 51–56.
36. Milligan,J.F. and Uhlenbeck,O.C. (1989) Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.*, **180**, 51–62.
37. Cromsigt,J., Schleucher,J., Gustafsson,T., Kihlberg,J. and Wijmenga,S. (2002) Preparation of partially $^{2}H/^{13}C$-labelled RNA for NMR studies. Stereo-specific deuteration of the H5" in nucleotides. *Nucleic Acids Res.*, **30**, 1639–1645.
38. Nagaswamy,U., Gao,X., Martinis,S.A. and Fox,G.E. (2001) NMR structure of a ribosomal RNA hairpin containing a conserved CUCAA pentaloop. *Nucleic Acids Res.*, **29**, 5129–5139.
39. Goddard,T.D. and Kneller,D.G. (2001) *SPARKY 3*. University of California, San Francisco.
40. Figueroa,N., Keith,G., Leroy,J.L., Plateau,P., Roy,S. and Gueron,M. (1983) NMR study of slowly exchanging imino protons in yeast $tRNA^{asp}$. *Proc. Natl Acad. Sci. USA*, **80**, 4330–4333.
41. Farrow,N.A., Muhandiram,R., Singer,A.U., Pascal,S.M., Kay,C.M., Gish,G., Shoelson,S.E., Pawson,T., Forman-Kay,J.D. and Kay,L.E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry*, **33**, 5984–6003.
42. Clore,G.M., Driscoll,P.C., Wingfield,P.T. and Gronenborn,A.M. (1990) Analysis of the backbone dynamics of interleukin-1 beta using two-dimensional inverse detected heteronuclear ^{15}N-^{1}H NMR spectroscopy. *Biochemistry*, **29**, 7387–7401.
43. Palmer,A.G. (1998) *Modelfree 4.0*. Columbia University, New York.
44. Varani,G., Cheong,C. and Tinoco,I.Jr (1991) Structure of an unusually stable RNA hairpin. *Biochemistry*, **30**, 3280–3289.
45. Allain,F.H. and Varani,G. (1995) Structure of the P1 helix from group I self-splicing introns. *J. Mol. Biol.*, **250**, 333–353.
46. Colmenarejo,G. and Tinoco,I.Jr (1999) Structure and thermodynamics of metal binding in the P5 helix of a group I intron ribozyme. *J. Mol. Biol.*, **290**, 119–135.
47. 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. *et al.* (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D. Biol. Crystallogr.*, **54**, 905–921.
48. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38, 27–38.
49. 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**, 5108–5121.
50. Shoup,D., Lipari,G. and Szabo,A. (1981) Diffusion-controlled bimolecular reaction rates. The effect of rotational diffusion and orientation constraints. *Biophys. J.*, **36**, 697–714.
51. Lipari,G. and Szabo,A. (1981) Nuclear magnetic resonance relaxation in nucleic acid fragments: models for internal motion. *Biochemistry*, **20**, 6250–6256.
52. Powers,R., Clore,G.M., Stahl,S.J., Wingfield,P.T. and Gronenborn,A. (1992) Analysis of the backbone dynamics of the ribonuclease H domain of the human immunodeficiency virus reverse transcriptase using ^{15}N relaxation measurements. *Biochemistry*, **31**, 9150–9157.