Determination of Metal Ion Binding Sites within the Hairpin Ribozyme Domains by NMR†

Samuel E. Butcher, Frédéric H.-T. Allain, and Juli Feigon*

Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095-1569

Received October 8, 1999; Revised Manuscript Received November 23, 1999

ABSTRACT: Cations play an important role in RNA folding and stabilization. The hairpin ribozyme is a small catalytic RNA consisting of two domains, A and B, which interact in the transition state in an ion-dependent fashion. Here we describe the interaction of mono-, di-, and trivalent cations with the domains of the ribozyme, as studied by homo- and heteronuclear NMR spectroscopy. Paramagnetic line broadening, chemical shift mapping, and intermolecular NOEs indicate that the B domain contains four to five metal binding sites, which bind Mn^{2+}, Mg^{2+}, and $Co(NH_3)_6^{3+}$. There is no significant structural change in the B domain upon the addition of $Co(NH_3)_6^{3+}$ or Mg^{2+}. No specific monovalent ion binding sites exist on the B domain, as determined by $^{15}NH_4^+$ binding studies. In contrast to the B domain, there are no observable metal ion interactions within the internal loop of the A domain. Model structure calculations of Mn^{2+} interactions at two sites within the B domain indicate that the binding sites comprise major groove pockets lined with functional groups oriented so that multiple hydrogen bonds can be formed between the RNA and $Mn(H_2O)_6^{2+}$ or $Co(NH_3)_6^{3+}$. Site 1 is very similar in geometry to a site within the P4−P6 domain of the *Tetrahymena* group I intron, while site 2 is unique among known ion binding sites. The site 2 ion interacts with a catalytically essential nucleotide and bridges two phosphates. Due to its location and geometry, this ion may play an important role in the docking of the A and B domains.

Coulombic interactions of cations with nucleic acids have profound effects on molecular and thermodynamic properties (*1*). Metal ions in particular play important roles in biochemistry, and the structure and function of modern day RNAs have been shaped by evolving in the presence of metal ions (*2*). In addition to providing electrostatic stabilization during folding, elegant biochemical experiments have shown that metal ions are directly involved in RNA catalysis for some ribozymes (*3*−*6*).

Despite their obvious importance, metal ions are often difficult to localize in RNA structures. From the small number of RNA crystal structures in which ions have been observed, it is evident that there are numerous ways in which they can interact with RNA (*7*−*14*). The L-shape of $tRNA^{Phe}$ is stabilized by tightly binding four metal ions within loop regions, with many more metal ions bound less tightly (*7*−*9*). Metal ions mediate helical packing of the P4−P6 domain of the *Tetrahymena* group I intron, in a manner akin to that of the hydrophobic core of proteins (*10*). A "metal ion zipper" in 5S rRNA bridges the close approach of phosphates (*13*). Additionally, metal ions have been localized at or near the active sites of some small ribozymes by crystallography (*11*, *14*, *15*). NMR has also emerged as a powerful technique for localizing metal ion binding sites on RNA in solution (*16*−*23*).

The hairpin ribozyme has a unique, two-domain structure which folds in an ion-dependent manner (*24*−*26*). Derived from the negative polarity strand of the tobacco ringspot virus satellite RNA, the hairpin ribozyme functions in processing the RNA during rolling circle replication (*27*). The ribozyme is a reversible, site-specific RNA endonuclease that cleaves RNA substrates to form products with cyclic phosphate and 5′-hydroxyl termini. The hairpin ribozyme is unusually promiscuous in its ionic preference, as it is able to utilize a variety of ions or polyamines for proper folding and catalysis (*28*−*31*). Using kinetically well-behaved molecules, it has been recently demonstrated that the catalytic rate of the hairpin ribozyme approaches similar values with many different ions, including manganese, cobalt hexamine, or magnesium (*32*).

The observation that cobalt hexamine efficiently supports hairpin ribozyme chemistry indicates that direct inner sphere metal ion coordination is not necessary for cleavage, since this ion has a kinetically stable outer hexamine ligand shell (*28*−*30*). These results show that metal ions do not act as Lewis acids in hairpin ribozyme chemistry. Rather, metal ions play an important role in promoting the proper folding required for hairpin ribozyme catalysis. Proper folding of the hairpin ribozyme involves a docking interaction between the A and B domains (Figure 1). Docking of the A and B domains is greatly facilitated by both di- and trivalent metal ions (*25*). Kinetic analysis suggests that approximately one to two metal ions are bound to the docked ribozyme during catalysis (*32*, *33*).

The solution structures of the individual A and B domains of the hairpin ribozyme have been solved by NMR (Figure 1) (*34*, *35*). Exactly how the two domains dock in the transition state is unknown. How do metal ions interact with the hairpin ribozyme to facilitate docking? Do ion binding

† This work was supported by NSF and NIH grants to J.F., a Jonsson Cancer Center postdoctoral fellowship to S.E.B., and Human Frontiers Sciences Program Organization postdoctoral fellowships to F.H.-T.A.

* To whom correspondence should be addressed. Phone: (310) 206-6922. Fax: (310) 825-0982. E-mail: feigon@mbi.ucla.edu.

10.1021/bi9923454 CCC: $19.00
Published on Web 02/12/2000

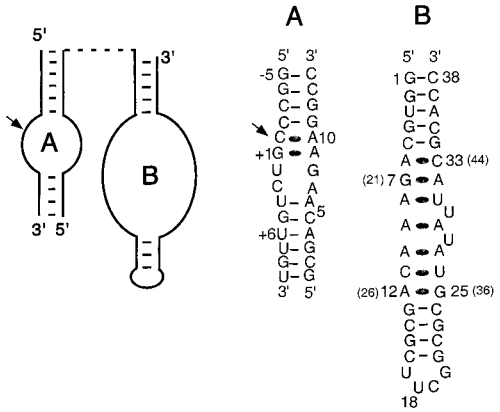

FIGURE 1: Schematic illustration of the hairpin ribozyme secondary structure, along with the sequences of the A and B domains used in this study. The RNA cleavage site in domain A is marked with an arrow. A dashed line between the A and B domains represents a covalent connection in the intact ribozyme secondary structure. The sequences of the individual A and B domains are identical to those used in their NMR structure determinations (*34*, *35*). Non-Watson−Crick base pairings determined by NMR are indicated with ovals. The numbering system for the B domain corresponds to the NMR construct, with the original hairpin ribozyme numbering scheme shown as adjacent numbers in parentheses. The A domain numbering scheme is the same for the NMR construct and the original hairpin ribozyme numbering scheme. The B domain contains a stable helical extension and UUCG tetraloop not normally found in the ribozyme.

sites exist in the ground-state domain structures, or are ions only captured when the A and B domains are in very close proximity? We set out to determine whether metal ion binding sites exist within the individual domains of the hairpin ribozyme, using NMR and the cations $Co(NH_3)_6^{3+}$, Mn^{2+}, Mg^{2+}, and $^{15}NH_4^+$. Binding of a paramagnetic Mn^{2+} ion results in the NMR line broadening of nuclei that are very close to the ion in three-dimensional space, while cobalt hexamine and ammonium binding can be detected by NMR via intermolecular NOEs. Additionally, chemical shift mapping can be used to locate sites of ion interaction. In this report we use the above methods to demonstrate that there are four to five metal binding sites on the loop B domain and none within the A domain. The NMR data are used to precisely localize two of the metal ion binding sites on the loop B domain structure.

MATERIALS AND METHODS

NMR Sample Preparation and Spectroscopy. RNA was transcribed and purified as previously described (*35*). The B domain RNA samples were 0.5−1 mM, in 50 mM NaCl, pH 7.0. The A domain samples were 1.4 mM (0.7 mM in strand), in 150 mM NaCl, pH 7.0. A higher concentration of monovalent salt was required to maintain the two strands of the A domain as a duplex, whereas the B domain maintains its hairpin structure at lower salt concentrations. Samples were in H_2O (90% H_2O/10% D_2O) or D_2O (99.99% D_2O, Isotech). NMR spectra were recorded on Bruker DRX 500 MHz spectrometers. Spectra were processed with Xwinnmr (Bruker instruments) and analyzed with Felix 97 software (Biosym). The B domain spectra in D_2O were recorded at 303 and 310 K, while the A domain spectra were recorded at 298 K. The residual water HDO signal was suppressed by low power presaturation. The 2D NOESY (*36*) spectra of the loop A domain were acquired with 96 scans of 1K points in the acquisition dimension and 512 increments in the indirect dimension. The mixing time was 200 ms. The spectra were apodized with a 90° phase-shifted sine bell and zero filled to 1K points in the indirect dimension. The 2D sensitivity-enhanced ^1H,^{13}C ct-HSQC (*37*) spectra of the B domain were acquired with 32 scans of 1024 points in the acquisition dimension and 220 increments in the indirect dimension. The spectra were processed as described above, except that the indirect dimension was zero filled to 512 increments. The constant time evolution period was set to 30 ms ($2/J_{cc}$) for C5−C6 refocusing. Other spectra that were collected for obtaining the paramagnetic Mn^{2+} line broadening data included 1D and 2D NOESY spectra in H_2O using the 11 echo water suppression pulse sequence (*38*), 2D NOESY and TOCSY (*39*) in D_2O, and 2D ^1H,^{15}N HSQC in H_2O and D_2O (*40*). For the $^{15}NH_4^+$ binding studies, a 1D ^1H ^{15}N-filtered NOESY (τ_m = 80 ms) (*41*) and a 1D ^1H ^{15}N-filtered ROESY (τ_m = 40 ms) (*42*) were acquired with 3072 scans of 8192 points.

Model Structure Calculations. The 30 lowest energy B domain structures obtained from the NMR structure determination (*35*) were used as starting structures for calculating the structures of the RNA−metal complex. Two $Mn(H_2O)_6^{2+}$ ions were placed at a random position relative to the RNA, approximately 30 Å away, by generating an X-PLOR (*43*) structure file containing the $Mn(H_2O)_6^{2+}$ coordinates and the B domain coordinates (*35*). The coordinates for the $Mn(H_2O)_6^{2+}$ ions were obtained from the high-resolution crystal structure (1.5 Å) 1D49 (*44*). Distance restraints of 0−7 Å were assigned to all atoms that were broadened to baseline at 20−40 μM $MnCl_2$. For site 1, these atoms were G1 H1, H2′, N1, H8, C8, and N7; G2 H1, H2′, N1, H8, C8, and N7; and U3 H3, N3, H5, and C5. For site 2, the atoms were A6 H8, C8, N7, H3′, and C3′; and G7 H8, C8, and N7. Repulsive distances were also included to help the ion localization. These distances were restrained to be >7 Å for a subset of atoms which showed no specific line broadening at 60 μM $MnCl_2$. For site 1, these atoms were G1 H1′ and C1′, G2 H1′ and C1′, C38 H1′, H5, and C5, C37 H5 and C5, and A36 H2 and C2. For site 2, the nonbroadened atoms were A6 H1′, C1′, H2, and C2, G21 H1′ and C1′, A8 H8 and C8, and A32 H2 and C2. All calculations were done with X-PLOR version 3.1 (*43*). First, the structures were subjected to an initial simulated annealing protocol of 105 ps at 1000 K, followed by 42 ps of cooling with a temperature step of 7 fs in the absence of van der Waals forces, followed by 50 steps of Powell algorithm energy minimization. Soft NOE potentials with a scale factor of 50 were used at this step. A final round of molecular dynamics and simulated annealing in the presence of van der Waals forces was then performed at 2000 K with cooling to 100 K in 28 ps, using a 0.7 fs time step. Square-well NOE potentials were used with a scale factor of 100, and the protocol was followed by 500 steps of energy minimization using the Powell algorithm.

RESULTS

Interaction of Mn^{2+} with the B Domain. The secondary structure of the hairpin ribozyme is shown, along with the domain constructs used in this study (Figure 1). The loop B domain sequence is identical to the one used in our recent NMR structure determination of the loop B domain (*35*). We chose to probe the loop B structure with $MnCl_2$ to

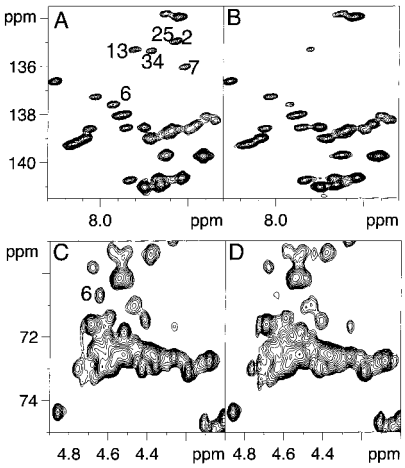

FIGURE 2: 500 MHz ^{1}H,^{13}C ct-HSQC spectra of the B domain, without (A and C) and with (B and D) the addition of 20 μM $MnCl_2$. (A and B) Aromatic region. Aromatic resonances that are broadened as a result of Mn^{2+} binding are labeled with their corresponding numbers in the spectrum without Mn^{2+}. (C and D) C2′−H2′ and C3′−H3′ region. The A6 C3′−H3′ correlation is labeled in (C) and is specifically broadened in (D). Spectra were acquired at 30 °C, with spectral widths of 5000 and 2800 Hz in the proton and carbon dimensions, respectively.

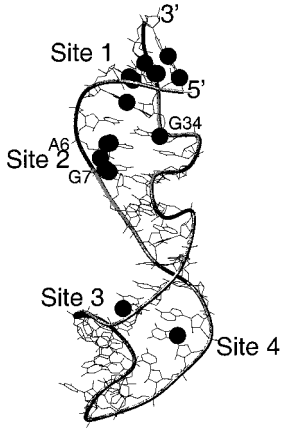

FIGURE 3: Lowest energy NMR structure of the B domain (*35*), with atoms that are specifically broadened as a result of Mn^{2+} binding shown as spheres.

determine if the divalent metal manganese interacts at specific sites within the B domain. Because Mn^{2+} is paramagnetic, it causes line broadening of NMR resonances in close proximity to the ion, with a distance dependence of r^{-6} (*45*). Additionally, it has been observed with other RNAs that only micromolar amounts of manganese are required to observe this effect (*16*, *17*, *23*), which circumvents the potential problem of aggregation at millimolar concentrations of both RNA and metal.

^{1}H,^{13}C ct-HSQC (*37*) spectra were recorded on uniformly ^{13}C,^{15}N-labeled loop B domain RNA with and without Mn^{2+} (Figure 2). At 20 μM Mn^{2+}, a specific subset of resonances within the loop B domain are completely or almost completely broadened to baseline. Most of the aromatic H8 and C8 resonances which are broadened by Mn^{2+} are within the internal loop and belong to the nucleotides A6, G7, G13, and G25 (Figure 2A,B). Additionally, the H8 and C8 resonances of the stem residues G1 and G2 are broadened beyond detection, and the imino protons of these nucleotides are also broadened (data not shown). The Mn^{2+} probing data show that the A6 H3′ proton is specifically broadened (Figure 2C,D). This proton is located in the major groove directly between the A6 and G7 H8 protons, which are also broadened (Figure 2A,B). Since the RNA concentration is millimolar, the Mn^{2+} is in fast exchange and must have an off-rate on the order of 1000 s^{-1}, which is similar to the time scale of Mn^{2+} binding to other RNAs (*16*, *17*, *23*).

The atoms which are observed to be specifically broadened by Mn^{2+} binding to the loop B domain structure are shown as spheres on the loop B domain structure (*35*) (Figure 3). All of the atoms which interact with Mn^{2+} ions are in the major groove. Figure 3 shows that the sites of manganese binding localize to four to five different regions of the B domain RNA. Site 1 is located within the major groove of the terminal helix. Site 2 is located within the internal loop, between nucleotides A6 and G7. G34 is also broadened by manganese. This nucleotide is between site 1 and site 2, so it is not clear as to whether this constitutes an additional site or is the result of ion movement between sites 1 and 2 (Figure 3). The final two sites of manganese interaction, referred to as sites 3 and 4, are located at nucleotides G13 and G25, respectively, at the junction of the internal loop and the second helix. Both of these sites are in the major groove but lie on opposite sides of the helix.

Interaction of Cobalt Hexamine and Magnesium with the B Domain. Since NMR chemical shifts are exquisitely sensitive in reporting ligand binding, ion interactions, and conformational changes, we acquired ^{1}H,^{13}C ct-HSQC spectra of the B domain in the absence and presence of cobalt hexamine and $MgCl_2$. $Co(NH_3)_6^{3+}$, an analogue of hexahydrated magnesium, efficiently stimulates high levels of hairpin ribozyme activity at relatively low concentrations. A ligation rate constant of 4.3 min^{-1} is observed in 0.1 mM cobalt hexamine, whereas 10 mM $MgCl_2$ produces a rate constant of 1.3 min^{-1} (*32*). We find that the B domain aggregates at high concentrations of cobalt hexamine (4 mM) or magnesium (10 mM) under NMR conditions (1 mM RNA; data not shown). We therefore probed the loop B domain in the presence of 1 mM cobalt hexamine or 6 mM magnesium, conditions which are both amenable for NMR spectroscopy and known to promote high levels of ribozyme activity.

The addition of 1 mM $Co(NH_3)_6^{3+}$ or 6 mM Mg^{2+} has little effect on the proton and carbon chemical shifts of the B domain, suggesting that the structure of the domain is the same in the absence and presence of cobalt hexamine (Figure 4A,B). However, small chemical shift changes are observed at positions which are at or near the nucleotides observed to interact with manganese. Resonances in the major groove of the terminal helix are affected, including G1, G2, and G34. The aromatic H8 chemical shifts of these resonances shift by 0.06−0.08 ppm in the proton dimension (Figure 4). These nucleotides are also specifically broadened by manganese and constitute metal binding site 1 (Figures 2 and 3). Within the internal loop, the C8 chemical shifts of A6−A10 shift by 0.4−0.6 ppm in the ^{13}C dimension upon addition of $Co(NH_3)_6^{3+}$ or Mg^{2+}. The A6 and G7 resonances were also specifically broadened by manganese and constitute metal binding site 2 (Figures 2 and 3). Last, similar chemical shift changes are observed for both the proton H8 and carbon C8 resonances of G13 and G25 upon cobalt hexamine binding, and these resonances are also specifically broadened by manganese and constitute metal binding sites 3 and 4.

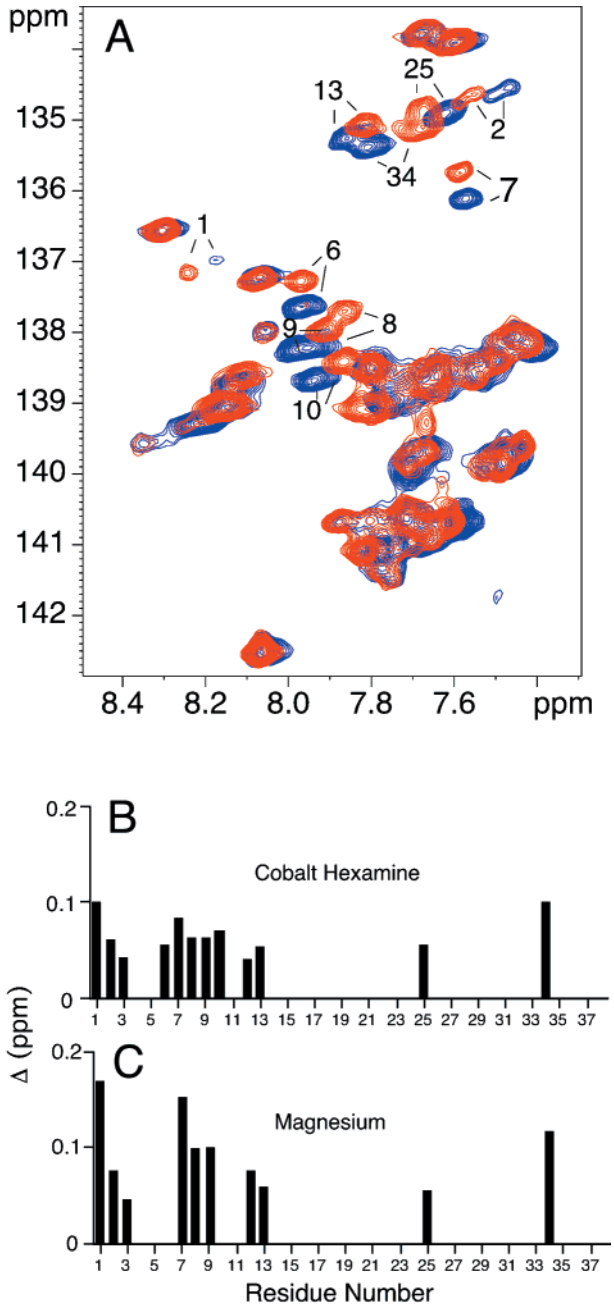

FIGURE 4: (A) $^{1}H,^{13}C$ ct-HSQC spectra of the B domain, with (red) and without (blue) the addition of 1 mM $Co(NH_3)_6^{3+}$. The aromatic region is shown. Resonances which display chemical shift changes upon interaction with cobalt hexamine are labeled. Spectra were acquired as for Figure 2, except that the temperature was 37 °C. (B and C) Observed changes in chemical shift (Δppm) upon the addition of cobalt hexamine or magnesium plotted vs the residue number in the B domain. Changes in chemical shifts were calculated from the equation $\Delta(\text{ppm}) = [(\Delta^{1}\text{H ppm})^2 + (\Delta^{13}\text{C ppm} \times \alpha C)^2)]^{0.5}$ (*56*), where Δppm is the difference in ppm between the chemical shifts of the B domain RNA in the absence and presence of metal ions and αC is a scaling factor to normalize the magnitude of the carbon chemical shift changes relative to the proton scale. The αC correction factor was calculated to be 0.11, given the H8/6 proton chemical shift range of 1 ppm and the C8/6 carbon chemical shift range of 9 ppm. (B) 1 mM cobalt hexamine. (C) 6 mM $MgCl_2$. Chemical shift changes were calculated for all H8/6 and C8/6 resonances in the presence of cobalt hexamine and for all H8/6 and C8/6 resonances in magnesium except 10, 11, and 26, which could not be unambiguously assigned for the magnesium-containing sample.

Therefore, all of the resonances which displayed paramagnetic line broadening in the presence of manganese show chemical shift changes upon the addition of cobalt hexamine, consistent with cobalt hexamine interacting with the same sites as manganese. Cobalt hexamine binding to other RNAs has been observed to induce similar chemical shift changes (*18*, *22*). The chemical shift mapping data indicate that magnesium binds to the same sites as both manganese and cobalt hexamine (Figure 4B).

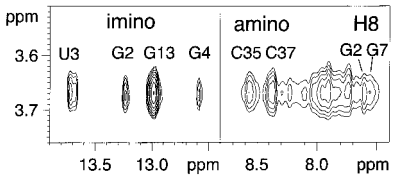

FIGURE 5: Portion of a 500 MHz 2D NOESY spectrum in 90% H_2O/10% D_2O of the B domain in the presence of 1 mM cobalt hexamine, which resonates at 3.65 ppm. The spectral width was 10 000 Hz in both dimensions, the mixing time was 150 ms, and the temperature was 10 °C. NOEs could be unambiguously assigned to the imino protons of nucleotides G2, U3, and G4 (left) and the amino protons of C35 and C37 and the H8 protons of G2 and G7 (right).

We observe that the ribose resonances are unaffected by the addition of cobalt hexamine or magnesium, indicating that the sugar and backbone conformation is unchanged by cobalt hexamine binding (data not shown). Although a specific subset of nucleotides displays chemical shift changes, none of their base pairing partners do (Figures 1 and 4). The observed chemical shift changes are relatively small, with the largest observed Δppm equal to 0.12 ppm (Figure 4B,C). Additionally, the same sequential NOE connectivities are observed in the presence or absence of 1 mM cobalt hexamine (data not shown). We therefore conclude that there are no significant structural changes in the B domain upon the addition of 1 mM cobalt hexamine or 6 mM $MgCl_2$.

Intermolecular NOEs between the cobalt hexamine protons and the RNA are detected in 2D NOESY spectra of the B domain RNA in 1 mM $Co(NH_3)_6^{3+}$ (Figure 5). At site 1, NOEs can be seen from cobalt hexamine to the imino protons of G2, U3, and G4, as well as the amino protons of C35 and C37 and the H8 aromatic proton of G2. At site 2, an NOE can be seen from cobalt hexamine to the H8 aromatic proton of G7. Last, a site 3 NOE is detected at the imino proton of G13. Many more intermolecular NOEs are also visible, but unambiguous assignments for these NOEs could not be made for two reasons. First, line broadening at low temperatures (0−10 °C) in the presence of cobalt hexamine caused severe spectral overlap even in 3D $^{1}H,^{13}C$ HSQC NOESY spectra (data not shown); second, at higher temperatures (20−40 °C), where line widths are narrower, the cobalt hexamine protons exchange too rapidly to be detected. All 18 cobalt hexamine protons resonate at a single frequency (3.65 ppm), indicating that the cobalt hexamine ion tumbles rapidly when bound to the RNA and is in fast exchange between the bound and free forms. Therefore, a single cobalt hexamine molecule can interact with multiple metal binding sites on the RNA during the course of the NMR experiment. The observed fast exchange kinetics of cobalt hexamine binding in solution have thus far been observed for every RNA which contains a cobalt hexamine binding site (*18*, *21*, *22*). The equilibrium dissociation constant for cobalt hexamine binding to a pseudoknot RNA has been estimated at 0.6 ± 0.4 mM, on

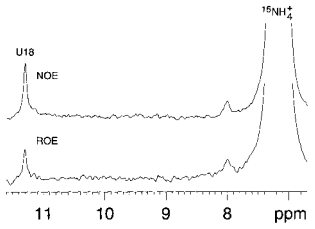

FIGURE 6: 500 MHz 1D ^{15}N-filtered NOE (top) and ROE (bottom) spectra of the B domain RNA (0.5 mM) in the presence of 25 mM $^{15}NH_4Cl$. The NOESY mixing time was 80 ms, while the ROESY mixing time was 40 ms. The temperature was 10 °C and spectral widths were 10 000 Hz. The $^{15}NH_4^+$ resonance and the NOE and ROE to the imino proton of U18 are labeled. The broad exchange peak at 8.0 ppm is not assigned.

the basis of the chemical shift differences between the free and bound states (*22*). The chemical shift differences observed for the B domain with and without cobalt hexamine are nearly identical to these values, and therefore the binding affinities for cobalt hexamine on the B domain appear to be quite similar.

Absence of Detectable Monovalent Ion Binding Sites. Monovalent ion binding sites on nucleic acids have been observed using the ion $^{15}NH_4^+$ (*46, 47*). Since the hairpin ribozyme functions efficiently under high concentrations of ammonium ions (*31, 32*), we wished to ascertain whether the multivalent ion binding sites also bind ammonium ions. 1D ^{15}N-filtered NOE and ROE experiments were performed in the presence of 25 mM $^{15}NH_4^+$ (Figure 6). One intense NOE is observed between the RNA and the $^{15}NH_4^+$ ion, which can be unambiguously assigned to the imino proton of U18 in the UUCG tetraloop. However, the 1D ^{15}N-filtered ROE experiment shows that the majority of this NOE is due to direct chemical exchange between the $^{15}NH_4^+$ ion and the U18 imino proton, which is possible since this proton is directly exposed to solvent (*35, 48*). We therefore find no evidence for specific monovalent ion binding sites within the B domain RNA, since the UUCG tetraloop was engineered for stability and is not part of the natural ribozyme sequence.

Absence of Manganese Interactions within Internal Loop A. The loop A domain sequence used in these studies is identical to that used by Cai and Tinoco in the NMR structure determination of the loop A domain (*34*) (Figure 1). Cai and Tinoco found that the loop A domain structure is the same in the presence or absence of magnesium (*34*). We find that, in contrast to the B domain, nearly all of the resonances in the loop A domain are unaffected by the addition of micromolar amounts of the paramagnetic Mn^{2+} ion. 2D homonuclear NOESY spectra of the loop A domain with and without Mn^{2+} show that specific line broadening is only observed at G1 H8 and G15 H8 (Figure 7). These nucleotides are involved in terminal G·U and G·C base pairs, respectively, and may constitute sites of metal ion localization due to the presence of adjacent 5′-terminal triphosphates. No resonances within the internal loop A are specifically broadened by Mn^{2+}, even at Mn^{2+} concentrations as high as 60 μM (data not shown). The sequential connectivity between the cleavage site nucleotides G+1 and C−1 are visible in both spectra (Figure 7), indicating that no manganese ions interact in the vicinity of the cleavage site.

Model Structure Calculations of the B Domain Metal Binding Sites. Figure 3 allows one to visualize the regions

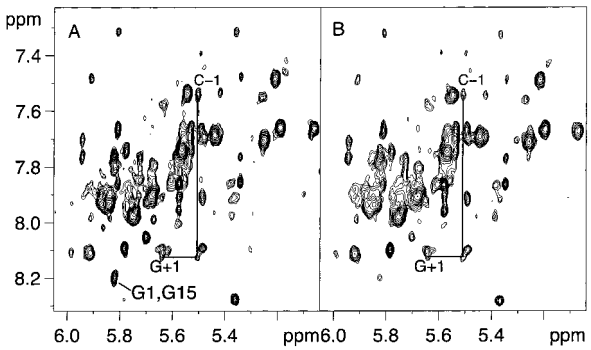

FIGURE 7: 500 MHz 2D NOESY spectra of the A domain, without (A) and with (B) the addition of 20 μM $MnCl_2$. The sequential H8/H6/H2−H1′/H5 region is shown. The G1 and G15 resonances that are broadened as a result of Mn^{2+} binding are labeled in the spectrum without Mn^{2+}. The spectral widths were 5000 Hz in both dimensions, the mixing times were 200 ms, and the temperatures were 20 °C. Lines indicate the sequential connectivity between the nucleotides surrounding the cleavage site, C−1 and G+1.

of manganese localization upon the B domain. However, representing the manganese binding data in this manner does not reveal the precise interaction of the ions at the different sites. Although the paramagnetic effect of manganese binding has not been previously used to calculate structures of manganese bound to RNA, the well-known r^{-6} distance dependence of the paramagnetic effect (*45*) can be used to extract structural information in much the same manner as the NOE. We therefore set out to use the distance dependence of the paramagnetic effect to determine model structures of the manganese ions bound to the B domain RNA.

Two sites within the loop B domain yielded significantly more line broadening data than the others, due to higher proton density and better chemical shift dispersion at those sites. These sites correspond to site 1 at the helical terminus involving G1, G2, and U3 and site 2 within the internal loop at A6 and G7 (Figure 3). These two sites of ion localization are also particularly interesting cases for determining the structure of the ion bound to the RNA. Site 1 has a helical sequence on one strand (5′-GGU-3′) that is identical to the P5 helix of the *Tetrahymena* group I intron P4−6 domain, which binds a hexahydrated divalent ion (*10*). However, the P5 site contains two G·U wobble pairs, whereas site 1 is purely Watson−Crick. Site 2 is especially interesting because it is the only observed site of ion localization involving a catalytically essential nucleotide (G7, corresponding to G21 in the intact hairpin ribozyme).

The 30 lowest energy B domain structures (*35*) were used as starting structures for calculating structure models of the manganese-bound RNA. The manganese ions used in the calculations are modeled as fully hexahydrated ions, for the following reasons. First, the fast off rate of the ions suggests that they are not directly coordinated to the RNA, and similar kinetics of Mn^{2+} binding have been observed at sites that bind a hexahydrated manganese ion (*21*). Second, cobalt hexamine also binds at these sites, as evidenced directly by intermolecular NOEs and chemical shift changes. Last, all divalent ions which influence the hairpin ribozyme structure must be able to do so as hexahydrates, since cobalt hexamine is very proficient at promoting proper folding.

Constraints between the center of the $Mn(H_2O)_6^{2+}$ ions and the RNA were included only for atoms which were

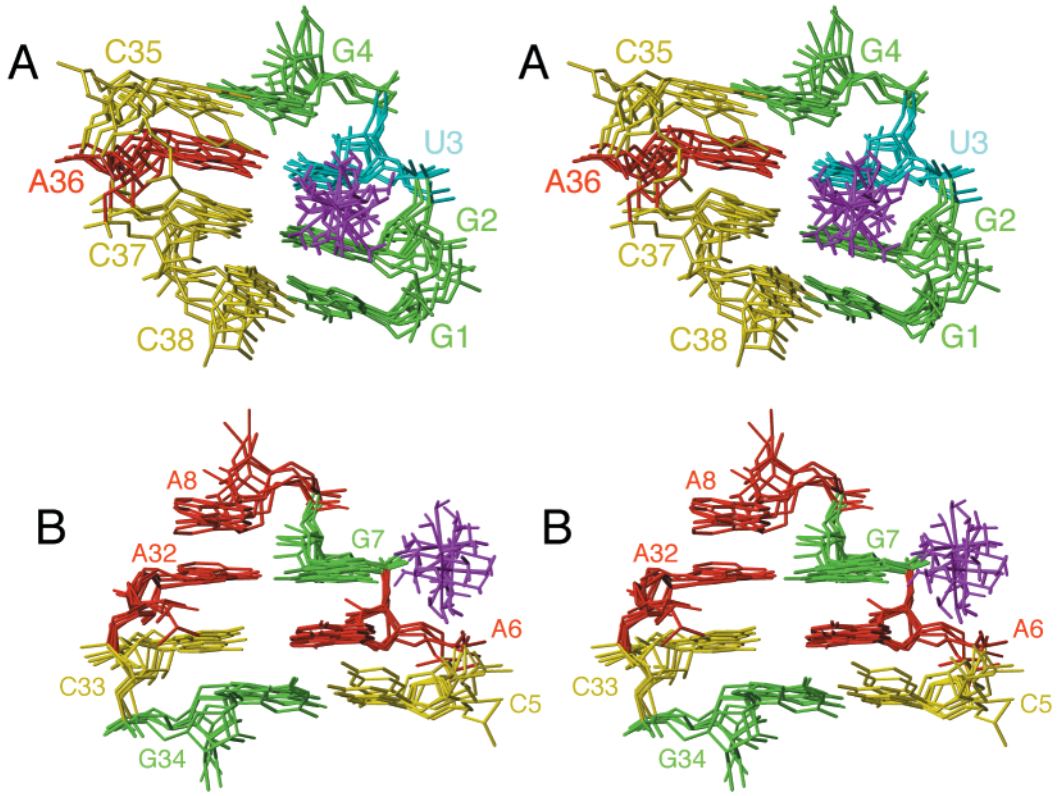

FIGURE 8: The five lowest energy model structures of the B domain with two hexahydrated manganese ions bound. The two sites are shown separately to illustrate the local precision of the RNA and the bound ion. (A) Site 1, superimposed upon G1, G2, U3, G4, C35, A36, C37, and C38. The view is into the major groove of the helical terminus. (B) Site 2, superimposed upon C5, A6, G7, A8, A32, C33, and G34. The view is into the major groove of the internal loop. The $Mn(H_2O)_6^{2+}$ ions were left out of the superimpositions to emphasize the localization of the ion. Note that the five lowest energy structures are shown, but the reported RMSD values in the table are for the 10 lowest energy structures out of the 30 calculated.

Table 1: Structure Determination Statistics for the 10 Lowest Energy Loop B Domain−2[$Mn(H_2O)_6^{2+}$] Complexes

total no. of B domain RNA restraints	1024
intermolecular restraints between RNA and ions	
site 1 to ion 1	18
repulsive	11
site 2 to ion 2	8
repulsive	10
RMSD for all heavy atoms relative to mean structure (Å)	
site 1: G1−G4, C35−C38, ion 1	1.44 ± 0.45
site 2: A6−A8, U31−C33, ion 2	1.35 ± 0.27
NOE violations (Å)	0 > 0.1
angle violations (deg)	0 > 5
mean deviation from covalent geometry	
bond lengths (Å)	0.004
angles (deg)	1.0
impropers (deg)	0.36

specifically broadened to baseline upon addition of 20−40 μM $MnCl_2$. All distances were weakly constrained to be 7 Å or less (Table 1 and Materials and Methods). Additionally, repulsive distances of >7 Å were included for atoms which were not affected by the paramagnetic Mn^{2+} ion (Table 1 and Materials and Methods). Structures of the B domain RNA in complex with two $Mn(H_2O)_6^{2+}$ ions were calculated with X-PLOR (*43*). Briefly, the Mn^{2+} ions started at random coordinates with respect to the RNA structures, approximately 30 Å away from each other. An initial round of simulated annealing in the absence of van der Waals forces allowed the ions to localize to the sites of specific line broadening. A second round of molecular dynamics and energy minimization refinement including van der Waals forces produced structures with low overall energies and RMSD values, good overall geometries, and no distance violations between the RNA and the ions or within the RNA itself (Table 1).

DISCUSSION

Model Structures of the Metal Ion Binding Pockets. Figure 8 shows the lowest overall energy structures of the calculated RNA−2[$Mn(H_2O)_6^{2+}$] complex. In both site 1 and site 2, the position of the ion was precisely localized with respect to the RNA, with a local RMSD for the surrounding RNA residues and the metal ion of about 1.4 Å for both sites (Table 1). The geometry and RMSD values of the calculated RNA−2[$Mn(H_2O)_6^{2+}$] model structures remain essentially unchanged from the starting structures. Therefore, the ion restraints are accommodated without deformation of the RNA structure. In site 1 (Figure 8A), the $Mn(H_2O)_6^{2+}$ ion is localized within the major groove of the helix between G2 and U3. The major groove is the site of highest negative electrostatic potential in A-form RNA. In site 2 (Figure 8B), the second $Mn(H_2O)_6^{2+}$ ion also localizes to the major groove, between purines A6 and G7 within the internal loop.

The orientations of the six water molecules surrounding each ion are less well defined than the position of the ion. However, the five lowest energy structures reveal that the waters are capable of forming hydrogen bonds with acceptor functional groups on the RNA (Figure 9). Four out of the six waters surrounding the ion bound to site 1 at the stem 5′-GGU-3′ sequence are oriented to form five potential

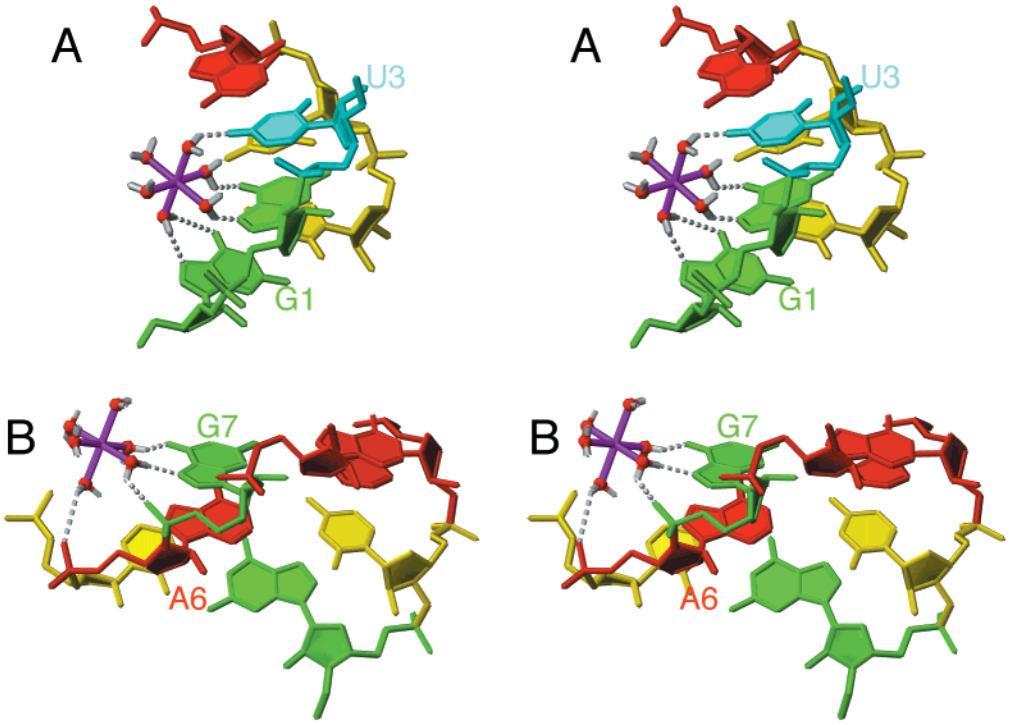

FIGURE 9: Structure models of the metal binding sites. (A) Site 1, rotated to illustrate the five possible hydrogen bonds between the $Mn(H_2O)_6^{2+}$ ion and the RNA. (B) Site 2, rotated to illustrate the four possible hydrogen bonds between the $Mn(H_2O)_6^{2+}$ ion and the RNA.

hydrogen bonds to the guanine N7, guanine O6, and uracil O4 functional groups at this site (Figure 9A). All five of the hydrogen bonds at this site are with base functional groups and none with the backbone. At metal binding site 2 within the internal loop, three out of six coordinated water molecules form four potential hydrogen bonds to functional groups on the RNA (Figure 9B). Two of these hydrogen bonds are to the A6 and G7 phosphate oxygens, while the third and fourth hydrogen bonds are to the G7 N7 and O6 functional groups. Therefore, this G7 metal binding site is quite different from the 5′-GGU-3′ site, in that two backbone contacts are formed in addition to the guanine base contacts.

Both of the calculated metal binding sites hydrogen bond to guanine N7 functional groups. Additionally, sites 3 and 4, which were not calculated, are likely to involve the N7 functional groups of G13 and G25, since the adjacent C8 and H8 atoms are specifically broadened by manganese. However, metal binding to guanine N7 groups cannot be a general effect, because no line broadening is observed at G4, G15, G20, G21, and G23, all of which have exposed N7 atoms. Therefore, there must be other determinants that influence metal ion localization on the RNA. Indeed, the structures of the metal binding pockets show that an array of hydrogen bond acceptors is required to constitute a metal binding site.

Structural Similarities with Other Metal Sites. The hairpin ribozyme B domain contains four to five metal binding sites. In contrast, the A domain only binds Mn^{2+} at the helical termini, where the 5′-terminal guanines and possibly their triphosphates constitute sites of metal interaction. Compared to the A domain, the B domain seems to bind a disproportionately large number of metal ions. In fact, the B domain binds the same average number of metal ions per nucleotide (0.13) in solution as does the recently studied rRNA S8 binding site (*20*), which is a 23-nucleotide RNA that contains

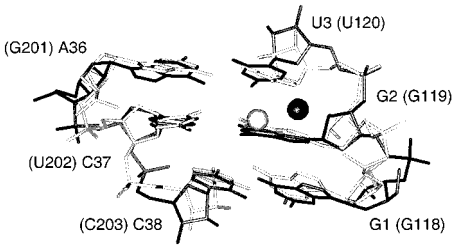

FIGURE 10: Comparison of site 1 from the B domain with the P5 site from the P4−P6 domain of the *Tetrahymena* group I intron (*10*). The two sequences are superimposed over the common heavy atoms of the RNA. The B domain RNA and bound ion are in black, and the P5 RNA and bound ion are in gray. Ions are shown as spheres.

three specific regions of metal ion localization. Therefore, the 28-nucleotide A domain may be unusual in its lack of metal ion binding sites, especially since metal ions are known to be important for the overall folding of the ribozyme. How do the structures of the metal ion binding sites determined for the B domain compare to known sites? Figure 10 compares the X-ray crystal structure of the P5 helical sequence 5′-GGU-3′/5′-GUC-3′ (*10*) with the helical sequence 5′-GGU-3′/5′-ACC-3′ determined in this study. The former sequence contains tandem wobble pairs, while the latter is purely Watson−Crick. Nevertheless, the RNA structures are remarkably similar (1.3 Å RMSD), and both bind a hexahydrated metal ion in the major groove within the 5′-GGU-3′ strand, with the ion located between the G and the U. Both ions interact only with the base functional groups at this site and not with the backbone. Interestingly, the location of the ion in the P5 helix is shifted 2.5 Å toward the center of the helix relative to the position of the ion in the B domain structure (Figure 10). This shift is likely to be due to the fact that the tandem G·U wobble pairs in the P5 helix provide hydrogen bond acceptors in the center of the

helix (e.g., uracil O4 and guanine O6 on the opposite strand), which are not present in Watson−Crick pairs. The electronegative surface area of the major groove containing G•U wobble pairs is therefore larger than one with Watson−Crick functional groups. The major groove of the Watson−Crick helix actually contains partial positive charges in the center of the helix in the form of adenine and cytosine amino groups. Therefore, the ion in the P5 major groove can easily slide over toward the center of the helix, whereas the B domain ion is forced to localize more toward the 5′-GGU-3′ strand.

The metal binding site at G7 lies within a commonly occurring structural motif, termed the "loop E-like" or "S-turn" motif, which has been observed in other RNAs (*49*−*51*). The B domain site 2 ion at G7 is the first experimental evidence of a metal ion binding within the S-turn motif. However, a metal binding site within the S-turn motif of the sarcin−ricin loop domain from 23S rRNA was predicted on the basis of Brownian dynamics simulations of cation diffusion trajectories in the presence of a calculated RNA electrostatic field (*52*). This predicted metal binding site occurs at a location which is equivalent to the site 2 ion in the B domain RNA.

Paramagnetic line broadening was observed at G13 H8 and G25 H8, and these residues also show small chemical shift changes upon addition of cobalt hexamine. However, spectral overlap made it difficult to collect enough restraints to precisely position ions near these nucleotides. Manganese binding has been observed previously within the crystal structure of a hammerhead ribozyme (*12*) at a site that is structurally analogous to site 3 at G13. In the hammerhead ribozyme structure, a manganese was observed to bind at the 5′-AG-3′ step of the sequence 5′-AG-3′/5′-CG-3′, which is the same sequence surrounding G13. In both the hammerhead ribozyme and the B domain structures, the bases form a sheared A•G pair followed by a G•C pair. The hammerhead ribozyme crystal structure shows that a manganese binds to the N7 of the G in the Watson−Crick pair in the hammerhead, which is equivalent to G13 in the B domain. Therefore, a similar type of manganese interaction may be formed at site 3 in the B domain.

Implications for Folding. The metal binding sites at the helical termini of the A domain must have no significance for ribozyme folding, since they are within regions that can be substituted by any combination of Watson−Crick base pairs without affecting the activity of the ribozyme, and triphosphates do not naturally occur at these positions. The observation that the loop resonances are not affected by Mn^{2+} agrees well with chemical probing data, which shows that there is no protection of the phosphodiester backbone of the A domain upon addition of divalent or trivalent metals (*26*). These data are also in agreement with the results of Cai and Tinoco, which report that the loop A structure is the same in the absence or presence of magnesium (*34*).

The chemical shift mapping data and NOE patterns observed for the B domain upon addition of cobalt hexamine or magnesium indicate that the structure of the domain is the same in the absence or presence of di- and trivalent ions. The observed chemical shift changes are similar to those that were observed for the leadzyme structure, which is also the same in the presence or absence of divalent metals (*53*). The largest observed Δppm value obtained from the chemical shift mapping data of the B domain is 0.12 ppm in 6 mM $MgCl_2$. Slightly greater changes were observed for some residues in the leadzyme in 5 mM $MgCl_2$ (*53*). Previously, it had been observed via 1D and 2D ^{1}H NMR spectra that the B domain shows the same number of imino proton resonances and NOE connectivities in 0−25 mM $MgCl_2$ (*35*). Additionally, it was shown that formation of a UV cross-link, which reports the folding of the internal loop, is not influenced by metal ion concentration (*54*). These data indicate that the metal ion binding sites on the B domain are part of the preformed RNA structure.

Is there a possible role for the G7 ion in the folding pathway of the ribozyme? No ions are observed to bind within the isolated internal loop of the A domain, and G7 is the only functionally important nucleotide in the B domain that interacts specifically with a metal ion. The ion in the major groove at G7 may directly participate in docking, by providing the necessary electrostatic shielding required for helical packing. It is therefore intriguing that the ion bound to G7 appears to be ideally situated for such a task. The cation is located between the phosphates of A6 and G7, directly on top of the phosphodiester backbone of the B domain, while half of its surface area (three out of the six coordinated waters) is exposed to the solvent, capable of binding another phosphodiester backbone. This is quite the opposite of the metal ion observed in site 1, which sits deep in the major groove of the helix and only interacts with base functional groups.

The metal binding site within loop B at G7 is unique among ion binding sites identified to date. Biochemical data support the presence of a metal ion bound to this site in the intact ribozyme. The N7 functional group of G7 is accessible to chemical modification in the absence of magnesium (*24*), as would be expected from the structure of the B domain. However, this N7 position is strongly protected upon addition of 12 mM magnesium (*24*). This is the only purine N7 position in the entire ribozyme which displays such a protection pattern (*24*). A metal ion interaction at G7 is also consistent with atomic mutagenesis data which indicate that both the O6 and N7 functional groups are catalytically important (*55*), even though these functional groups do not participate in intradomain interactions (*35*). Additionally, a manganese-induced cleavage within the B domain of the hairpin ribozyme has been observed between G7 and A8 (corresponding to G21 and A22 in the full ribozyme; K. Hampel and J. M. Burke, personal communication). Finally, the observation that cobalt hexamine, magnesium, and manganese interact at the same site is consistent with the wide ionic preference of the ribozyme. While the ground-state ion binding sites determined in this investigation may be involved in the ribozyme folding pathway, it remains to be seen as to whether the same sites exist in the transition state.

ACKNOWLEDGMENT

We thank Dr. K. Hampel and Prof. J. M. Burke for communication of results prior to publication, David Finger for assistance with the labeled NTP preparation, and Dr. Nicholas Hud for assistance with manganese titrations.

REFERENCES

1. Anderson, C. F., and Record, M. T. J. (1995) *Annu. Rev. Phys. Chem. 46*, 657−700.
2. Feig, A. L., and Uhlenbeck, O. C. (1999) in *The RNA World* (Gesteland, R. F., Cech, T. R., and Atkins, J. F., Eds.) p 287, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
3. Piccirilli, J. A., Vyle, J. S., Caruthers, M. H., and Cech, T. R. (1993) *Nature 361*, 85−88.
4. Ruffner, D. E., and Uhlenbeck, O. C. (1990) *Nucleic Acids Res. 18*, 6025−6029.
5. van Tol, H., Buzayan, J. M., Feldstein, P. A., Eckstein, F., and Bruening, G. (1990) *Nucleic Acids Res. 18*, 1971−1975.
6. Slim, G., and Gait, M. J. (1991) *Nucleic Acids Res. 19*, 1183−1188.
7. Jack, A., Ladner, J. E., Rhodes, D., Brown, R. S., and Klug, A. (1977) *J. Mol. Biol. 111*, 315−328.
8. Holbrook, S. R., Sussman, J. L., Warrant, R. W., Church, G. M., and Kim, S. H. (1977) *Nucleic Acids Res. 22*, 3722−3727.
9. Hingerty, B. E., Brown, R. S., and Klug, A. (1982) *Biochim. Biophys. Acta 697*, 78−82.
10. Cate, J. H., Hanna, R. L., and Doudna, J. A. (1997) *Nat. Struct. Biol. 4*, 553−558.
11. Scott, W. G., Finch, J. T., and Klug, A. (1995) *Cell 81*, 991−1002.
12. Pley, H. W., Flaherty, K. M., and McKay, D. B. (1994) *Nature 372*, 68−74.
13. Correll, C. C., Freeborn, B., Moore, P. B., and Steitz, T. A. (1997) *Cell 91*, 705−712.
14. Wedekind, J. E., and McKay, D. B. (1999) *Nat. Struct. Biol. 6*, 261−268.
15. Feig, A. L., Scott, W. G., and Uhlenbeck, O. C. (1998) *Science 279*, 81−84.
16. Ott, G., Arnold, L., and Limmer, S. (1993) *Nucleic Acids Res. 21*, 5859−5864.
17. Allain, F. H.-T., and Varani, G. (1995) *Nucleic Acids Res. 23*, 341−350.
18. Kieft, J. S., and Tinoco, I. (1997) *Structure 5*, 713−721.
19. Gdaniec, T. C., Sierzputowska-Gracz, H., and Theil, C. (1998) *Biochemistry 37*, 1505−1512.
20. Kalurachchi, K., and Nikonowicz, E. P. (1998) *J. Mol. Biol. 280*, 639−654.
21. Colmenarejo, G., and Tinoco, I. J. (1999) *J. Mol. Biol. 2*, 119−135.
22. Gonzalez, R. L., and Tinoco, I. J. (1999) *J. Mol. Biol. 289*, 1267−1282.
23. Hansen, M. R., Simorre, J.-P., Hanson, P., Mokler, V., Bellon, L., Beigelman, L., and Pardi, A. (1999) *RNA 5*, 1099−1104.
24. Butcher, S. E., and Burke, J. M. (1994) *J. Mol. Biol. 244*, 52−63.
25. Walter, N. G., Hampel, K. J., Brown, K. M., and Burke, J. M. (1998) *EMBO J. 17*, 2378−2391.
26. Hampel, K. J., Walter, N. G., and Burke, J. M. (1998) *Biochemistry 37*, 14672−14682.
27. Prody, G. A., Bakos, J. T., Buzayan, J. M., Schneider, I. R., and Bruening, G. (1986) *Science 231*, 1577−1580.
28. Hampel, A., and Cowan, J. A. (1997) *Chem. Biol. 4*, 513−517.
29. Nesbitt, S., Hegg, L. A., and Fedor, M. J. (1997) *Chem. Biol. 4*, 619−630.
30. Young, K. J., Gill, F., and Grasby, J. A. (1997) *Nucleic Acids Res. 25*, 3760−3766.
31. Murray, J. B., Seyhan, A. A., Walter, N. G., Burke, J. M., and Scott, W. G. (1998) *Chem. Biol. 5*, 587−595.
32. Nesbitt, S. M., Erlacher, H. A., and Fedor, M. J. (1999) *J. Mol. Biol. 286*, 1009−1024.
33. Walter, F., Murchie, A. I. H., and Lilley, D. M. J. (1998) *Biochemistry 37*, 17629−17636.
34. Cai, Z., and Tinoco, I., Jr. (1996) *Biochemistry 35*, 6026−6036.
35. Butcher, S. E., Allain, F. H. T., and Feigon, J. (1999) *Nat. Struct. Biol. 6*, 212−216.
36. Macura, S., Huang, Y., Suter, D., and Ernst, R. R. (1980) *J. Magn. Reson. 43*, 259−281.
37. Santoro, J., and King, G. (1992) *J. Magn. Reson. 97*, 202−207.
38. Bax, A., Sklenář, V., Clore, G. M., and Gronenborn, A. M. (1987) *J. Am. Chem. Soc. 109*, 6511−6513.
39. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−360.
40. Palmer, A. G., III, Cavanagh, J., Wright, P. E., and Rance, M. (1991) *J. Magn. Reson. 93*, 151−170.
41. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663−10665.
42. Griesinger, C., and Ernst, R. R. (1987) *J. Magn. Reson. 75*, 261−271.
43. Brünger, A. T. (1992) *X-PLOR (Version 3.1) Manual*, Yale University Press, New Haven and London.
44. Quintana, J. R., Grzeskowiak, K., Yanagi, K., and Dickerson, R. E. (1992) *J. Mol. Biol. 225*, 379−395.
45. Bertini, I., and Luchinat, C. (1986) *NMR of paramagnetic molecules in biological systems*, Benjamin/Cummings, Menlo Park, CA.
46. Hud, N. V., Schultze, P., and Feigon, J. (1998) *J. Am. Chem. Soc. 120*, 6403−6404.
47. Hud, N. V., Sklenář, V., and Feigon, J. (1999) *J. Mol. Biol. 286*, 651−660.
48. Allain, F. H.-T., and Varani, G. (1995) *J. Mol. Biol. 250*, 333−353.
49. Wimberly, B., Varani, G., and Tinoco, I., Jr. (1993) *Biochemistry 32*, 1078−1087.
50. Szewczak, A. A., Moore, P. B., Chan, Y.-L., and Wood, I. G. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 9581−9585.
51. Correll, C. C., Munishkin, A., Chan, Y., Ren, Z., Wool, I. G., and Steitz, T. A. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 13436−13441.
52. Hermann, T., and Westhof, E. (1998) *Structure*, 1303−1314.
53. Legault, P., Hoogstraten, C. G., Metlitzky, E., and Pardi, A. (1998) *J. Mol. Biol. 284*, 325−335.
54. Butcher, S. E., and Burke, J. M. (1994) *Biochemistry 33*, 992−999.
55. Grasby, J. A., Mersmann, K., Singh, M., and Gait, M. J. (1995) *Biochemistry 34*, 4068−4076.
56. Farmer, B. T., Constantine, K., Goldfarb, V., Friedrichs, M. S., Wittekind, M., Yanchunas, J., Robertson, J. G., and Mueller, L. (1996) *Nat. Struct. Biol. 3*, 995−997.

BI9923454