STRUCTURAL INVESTIGATION OF THE EFFECT OF MAGNESIUM-ION BINDING ON STEM-LOOP V OF THE *NEUROSPORA* VS RIBOZYME

by

DEAN O. CAMPBELL

(Under the direction of Pascale Legault)

ABSTRACT

An essential step in the substrate recognition of the *Neurospora* VS ribozyme is the formation of a loop-loop interaction between the terminal loops of stem-loop I and stem-loop V. Site-specific substitution and mutagenesis have suggested that this interaction is facilitated by the presence of a U-turn in the loop of stem-loop V. We have determined the structure of stem-loop V by nuclear magnetic resonance (NMR) spectroscopy and show that it adopts a U-turn conformation, a common motif found in RNA. Structural comparisons indicate that the U-turn of stem-loop V (SL5) fulfills some but not all of the structural characteristics found in canonical U-turn motifs. The formation of the stem-loop I / stem-loop V interaction is magnesium-ion dependent. Chemical shift mapping indicated that the addition of magnesium ions to SL5 resulted in a conformational change in loop. We have determined the structure of SL5 in the presence of magnesium ions and show that the binding of magnesium ions to the loop causes a conformational change. This conformational change results in the U-turn motif of SL5 having more characteristics of a canonical U-turn motif in the presence of magnesium ions. We used paramagnetic line broadening effect of manganese ions to localize four divalent metal ion binding sites in the loop of stem-loop V. Three of these divalent metal ion binding sites are

specific to the U-turn motifs and have been found in structures of other U-turn motifs. We also show that the U-turn conformation in stem-loop V allows the bases involved in the stem loop I / stem-loop V interaction (G697, A698, and C699) to be accessible, by exposing their Watson-Crick faces to the solvent. This study brings some light on the mechanism of substrate recognition in the VS ribozyme and expands our understanding of the role of U-turn motifs in RNA structure and function.

INDEX WORDS: U-turn motif, *Neurospora* VS ribozyme, Magnesium-ion binding, Chemical-shift mapping, RNA structure.

STRUCTURAL INVESTIGATION OF THE EFFECT OF MAGNESIUM-ION BINDING ON STEM-LOOP V OF THE *NEUROSPORA* VS RIBOZYME

by

DEAN O. CAMPBELL

B.Sc., The University of the West Indies, Jamaica, 1998

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2004

STRUCTURAL INVESTIGATION OF THE EFFECT OF MAGNESIUM-ION BINDING ON STEM-LOOP V OF THE *NEUROSPORA* VS RIBOZYME

by

DEAN O. CAMPBELL

Major Professor: Pascale Legault

Committee: Sidney Kushner
Kojo Mensa-Wilmot
Michael Pierce
Michael Terns

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2004

DEDICATION

This is dedicated to Beverly Dale

ACKNOWLEDGEMENTS

I would like to thank Drs. Bao Nguyen and Paola Di Lello for their many helpful tips and instruction. And also for making my stay in the “downstairs crew” a memorable one. Also to the members of the Leagult/Omichinksi labs, past and present for all the fun and excitement that allowed the years to go by so quickly.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS.....v

CHAPTER

1 LITERATURE REVIEW.....1

The *Neurospora* VS Ribozyme..... 1

U-Turn Motifs in the VS Ribozyme 6

Metal-Ions Binding in the *Neurospora* VS Ribozyme 8

Study of RNA by NMR Spectroscopy..... 10

Purpose of Study..... 11

References..... 20

2 RESONANCE ASSIGNMENT AND STRUCTURAL RESTRAINTS

OF SL5..... 28

Introduction..... 28

PART A: Resonance assignment of the ^{13}C, ^{15}N, and ^{1}H

chemical shifts of SL5 29

A.1 Atomic composition of stem-loop V (SL5) 29

A.2 Assignment procedure for SL5 29

PART B: Obtaining the structure calculation restraints for the structure calculation of SL5 45

B.1. Restraints for structure calculation 45

B.2 NOE distance restraints 45

B.3. Torsion angle restraints 50

B.4. Hydrogen bonds 54

References 94

3 NMR STRUCTURE OF STEM-LOOP V OF THE VS RIBOZYME AND MAGNESIUM-ION BINDING FROM CHEMICAL-SHIFT MAPPING 98

Abstract 99

Introduction 99

Materials and Methods 102

Results 108

Discussion 118

Acknowledgement 123

Figures and Tables 124

References 137

Supplementary Material 147

4 NMR STRUCTURE OF THE VS RIBOZYME STEM-LOOP V IN THE PRESENCE OF MAGNESIUM IONS AND LOCALIZATION OF DIVALENT METAL ION BINDING SITES 150

Abstract 151

Introduction 151

Materials and Methods 155

Results 161

Discussion 172

Acknowledgement 177

Figures and Tables 178

References 189

Supplementary Material 197

5. CONCLUSIONS 202

Summary of study 202

Impact of research and future directions 203

References 207

CHAPTER 1

LITERATURE REVIEW

The *Neurospora* VS Ribozyme

Ribozymes were discovered in the 1980s independently in the labs of Drs. Sydney Altman and Thomas R. Cech (1,2). Ribozymes are *ribo*nucleic acid en*zymes*, and naturally-occurring ribozymes are capable of a variety of chemical reactions, most notably phosphoryl transfer reactions and peptide bond formation (3). Synthetic ribozymes have been developed that are capable of a wide range of chemical activity, such as aminoacylation, alkylation, and carbon-carbon bond formation (4,5). There are many types of naturally-occurring ribozymes including the self-splicing group I and group II introns (2,6), ribonuclease P (1), the hepatitis delta virus (7), hairpin (8), hammerhead (9), and the *Neurospora* VS ribozymes (10). The ribosome (11,12) is a ribonucleoprotein complex which has been postulated to have RNA-catalyzed chemistry. The hammerhead, hairpin, HDV and *Neurospora* VS ribozymes belong to the group of small self-cleaving nucleolytic ribozymes (13). The small nucleolytic ribozymes produce 2', 3'-cyclic phosphate and 5'-hydroxyl termini after cleaving their substrate (14,15).

The *Neurospora* VS RNA was isolated from the Varkud Satellite (VS) plasmid of the *Neurospora* mitochondria in the laboratory Dr. Richard A. Collins at the University of Toronto (10). In the presence of divalent metal ions, the VS RNA is capable of site-specific *cis*- and

trans-cleavage and ligation reactions (in *vitro* and in *vivo*) (16,17). The physiological role of the VS RNA is that it is a step in the replication cycle of the double-stranded VS plasmid DNA. The double-stranded plasmid DNA is of monomeric length and circular and replicates by a rolling circle mechanism (Figure 1.1) (18). The VS RNA is transcribed from the VS DNA as a multimeric transcript and processes itself into monomeric units by site-specific self-cleavage. The VS RNA then self-ligates the ends of the linear monomers, circularizing them. The circular VS RNA is then reverse-transcribed and a second strand is synthesized to reproduce the VS DNA plasmid. The monomeric VS RNA is 881 nucleotides (nt) in length. However, the minimal contiguous nucleotide sequence required to retain full *cis*-cleavage activity is 154 nt in length with only a single nucleotide upstream of the cleavage site (16). This minimal sequence was deduced from deletion analysis and site-directed mutagenesis and the identity of the single upstream nucleotide is not critical. Even smaller self-cleaving versions of the VS ribozyme have been designed from deletions, substitutions and circular permutations (19).

The secondary structure of the VS ribozyme consists of six helices or stem-loops (stem-loop I-stem-loop VI) (Figure 1.2) (20). The secondary structure of the VS ribozyme was deduced from the correlation between chemical modification structure probing, site-directed mutagenesis, and the structure prediction program MFOLD (20). The secondary structure was found to contain a 5' stem-loop structure representing the substrate (stem-loop I) with an internal loop containing the cleavage site. The catalytic domain of the ribozyme is made up of five helices (stem-loops II-VI) in a H-shaped configuration with two helical junctions. By separating the substrate domain from the catalytic domain the VS ribozyme is transformed into a *trans*-cleaving molecule (21). The minimal contiguous substrate sequence required for cleavage is 20 nt, with only one nucleotide upstream of the cleavage site. The optimal activity of the ribozyme is dependent on

metal ions with divalent ions, such as magnesium, likely being the most preferred *in vivo*. The reaction rate is not affected by pH in the range 5.5 to 8.9, which suggests that a hydroxide intermediate is not involved in the cleavage chemistry or that the chemistry of the reaction is not the rate-limiting step (22).

The rate of the *cis*-cleavage reaction is ~0.1 min^{-1} for the 154 nt sequence, however the smaller constructs have reaction rates up to 30 times faster (19). Recently, variants of the VS ribozyme have been synthesized that have *cis*-cleavage rates of up to 600 min^{-1} (23). The *trans* cleavage reaction has a reaction rate of ~0.7–1.0 min^{-1} with a K_M of ~0.13-1.0 μM (21,24). Extension of the 5' and 3' ends of the 154 nt contiguous sequence in a complementary fashion to form another helical domain enables the ligation reaction (25).

The secondary structure of the VS ribozyme was determined in the absence of divalent metal ions. The addition of divalent metal ions produced a folded three-dimensional structure. The best-characterized tertiary interaction in the VS ribozyme was the formation of a loop-loop interaction between stem-loops I and V (Figure 1.3), which was characterized by site-directed mutagenesis and chemical modification (26). The loop-loop interaction is formed from the Watson-Crick base pairing between nucleotides 630-632 in stem-loop I, and 697-699 in stem-loop V (Figure 1.3). By performing substitutions that disrupted the formation of the base pairs involved in the interaction, activity was diminished. Compensatory mutations that restored the Watson-Crick base pairing between the loops restored the activity, although not to wild-type levels. This loop-loop interaction is important for the cleavage and ligation activities of the VS ribozyme (26,27).

The loop-loop interaction between stem-loops I and V causes a conformational change in the substrate (stem-loop I) (27). This was shown from a combination of *in vitro* selection (with

the formation of the loop-loop interaction as the determinant), site-directed mutagenesis and chemical structure probing. There is a shift in the 3' side of stem-loop I, in which nucleotides 623-625 in stem-loop I change their base-paring partners from nucleotides 634-636 to nucleotides 635-637 (Figure 1.3). This shift is accompanied by an extrusion of the base of residue C634 from the base-paired stem. Formation of the loop-loop interaction between the substrate and an isolated stem-loop V is sufficient to shift the wild-type substrate (28). The conformational change also affects the cleavage site as residue C637, unpaired in the unshifted conformation, becomes base paired to G623 in the shifted conformation. This was proposed to lead to the formation of a metal-ion binding site in the cleavage site internal loop of the shifted conformation (27). Both the loop-loop interaction and the conformational change in stem-loop I are magnesium-ion dependent (27,28).

There have been three NMR structures published of the substrate domain (29-31). Two of the studies were on the substrate in the unshifted conformation (29,31). Both studies found that the internal loop is comprised of two sheared G·A base pairs (G620-A639 and G638-A621), and a protonated A^{+}·C pair (A622-C637). The third structure is one in which the cleavage site is in the shifted conformation (30). This structure shows that the proposed metal-ion binding site does indeed forms in the shifted conformation exclusively. It also suggested that the cleavage site adopts a new RNA fold composed exclusively of sheared G·A base pairs (one between G620 and A639 and a second shared sheared base pair between G638 and A621 and A622) that is homologous to a long-range interaction motif found in rRNA (30).

Attempts have been made to determine the overall fold of the VS ribozyme, as well as, to identify the active site. Hydroxyl radical footprinting analysis showed that the ribozyme domain forms a solvent-inaccessible core, with several nucleotides important for the structure and/or

activity being protected (32). These nucleotides are found near the two three-helix junctions, in the loop-loop interaction, and at other sites. Site-specific mutations and chemical modifications were used to identify the importance of the III-IV-V helical junction in the formation of the loop-loop interaction (33). The junction contains a putative U-turn motif that if disrupted reduces the stability of the loop-loop interaction. The U-turn could be replaced by a hairpin, which suggests that the role of the junction is purely structural. By replacing the substrate with a mutation that is constitutively in the shifted conformation, it was shown by deletion analysis that the upper subdomain of the ribozyme (stem-loops III/IV/V) is not necessary for cleavage, and hence does not contain the active site (34). This meant that the active site is contained in the lower subdomain of the ribozyme comprising of stem-loops II/III/VI. The folding of the 2-3-6 helical junction (in isolation) was investigated using comparative gel electrophoresis and fluorescence resonance energy transfer (FRET) analysis (35). The comparative gel electrophoresis showed that the 2-3-6 helical junction underwent ion-dependent folding into a stable conformation. The FRET analysis employed involved measuring the efficiency of energy transfer between donor and acceptor fluorophores attached at pairs of 5'-termini and can be used to investigate global conformation of nucleic acids (36). It showed that stem-loops III and VI are coaxially stacked on each other with stem-loop II subtending at an angle of 65º from stem-loop VI. The three stem-loops are in an almost planar arrangement. The stacking of stem-loops III and VI was supported by deletion analysis and UV-induced cross-linking between stem-loop II and III (19). The ion-induced folding of the 2-3-6 junction was shown not to change significantly in the context of the full-length ribozyme. (Figure 1.4) (37). These studies suggest that a cleft is formed between stem-loops II and VI that likely serves as a receptor for the substrate (Figure 1.4b) (24). The folding of the 3-4-5 helical junction was also investigated in isolation using comparative gel

electrophoresis and FRET analysis (38). The results indicated that the 3-4-5 helical junction undergoes an ion-dependent conformational change that involves movement of stem-loop V, leaving stem-loops III and IV coaxially stacked. Stem-loop V ends up being at a smaller angle to stem-loop III than to stem-loop IV. A model of the tertiary structure of the catalytic domain was derived by assembling the two three-way junctions through their common stem-loop III, using the dihedral angle between stem-loops II and V measured by gel electrophoresis (Figure 1.4) (38). The cleft between stem-loops II and VI is clearly seen in this model and is thought to be the docking site of the substrate (stem-loop I) (Figure 1.4b).

Stem-loops II and VI contain several base bulges and one internal loop (Figure 1.2). The importance of these bulges and this internal loop was studied by site-specific mutational analysis (37). It was found that the bulge at residue 652 was important though the identity of the residue was not. This was essentially the same result for the two-adenine bulge at positions 725 and 726. The internal loop containing nucleotide A730 is referred to as the A730 loop, and it is the most affected by the site-directed mutagenesis (37). The four unpaired bases are very sensitive to nucleotide substitution, with A756 being the most sensitive. The FRET analysis was repeated with A756 mutated to a G and this mutation was shown to have little or no effect on the global fold of the ribozyme. Residues A730 and A756 were found to be UV cross-linked to a 4-thio-uridine that was placed immediately downstream of the scissile bond (39). These observations lead to the A730 loop being proposed as the active site of the VS ribozyme (37,39).

U-Turn Motifs in the VS Ribozyme

U-turns (or more precisely uridine-turns) are RNA structural motifs that are characterized mainly by a sharp reversal in the RNA phosphodiester backbone following a uridine base. U-

turns were first found in the anticodon and TψC-loops of tRNA (40). They are also found in the hammerhead ribozyme (41,42), 23s rRNA (43), U2 snRNA (44), and HIV RNA (45). The sharp turn in the direction of the RNA backbone induced by U-turns results in an increased exposure to solvent for the bases located 3' of the turn, which allows these bases to form tertiary interactions (46). The characteristics of a canonical U-turn motif are: a UNR sequence (U = uracil, N = any base, R = purine), a non-Watson-Crick base pair involving the nucleotide 5' of the UNR sequence, a sharp turn in the backbone after the U, stacking of the bases 3' of the U, stacking of the U base with the 5'-phosphate group of the R, and two hydrogen-bonding interactions involving the U residue (40,46). These hydrogen-bonding interactions are between the U 2'OH and the R N7 and between the U H3 and the 3' phosphate group of the R (Figure 1.5).

The two terminal loops of stem-loops I and V are involved in the loop-loop interaction that is important for activity of the ribozyme. The residues U628-G630 of stem-loop I and U696-C699 of stem-loop V form U-turn consensus sequences. Despite the fact that the secondary structure of the VS ribozyme (Figure 1.2) shows that the base pairs 5' of U628 and U696 are Watson-Crick base pairs (involving G627 and C632 for stem-loop I and U695 and A701 for stem-loop V), there is strong evidence for the formation of U-turns at these positions. Using site-specific mutagenesis to replace the uridine base at positions U628 and U696 greatly affected activity (26). Using NAIM analysis to identify 2'-OH groups that are important for activity have implicated U628 and U696 in U-turn motifs where the U 2'-OH participates in a hydrogen-bond interaction. In the NMR structure of the wild-type stem-loop I, the terminal loop was very dynamic and a conformation was not identified (29). The proposed base pair between G627 and C632 was not seen in the structure. It is important to note here that the published structure was solved in the absence of magnesium ions.

Metal-Ion Binding in the VS Ribozyme

Metal ions have specific roles in the folding pathway, the tertiary structure and the chemical mechanism of ribozymes (47). All the small naturally-occurring nucleolytic ribozymes require metal ions for their activity, except for the hairpin ribozyme (48,49). The function of the metal ion could be to properly orient the molecule for activity and/or be directly involved in the chemistry of the reaction. There is one generally accepted reaction mechanism for all the nucleolytic ribozymes. The cleavage reaction proceeds by a nucleophilic attack of the 2'-oxygen on the 3'-phosphorus in an S_N2 reaction, with departure of the 5'-oxygen, leaving a 2'-3' cyclic phosphate. The ligation reaction is the reverse, in which the 5'-oxygen attacks the 3'-phosphorus of the cyclic phosphate (13). Metal ions may play a role in the reaction chemistry by acting as general acid-base catalysts when solvated, or forming inner-sphere complexes and acting as Lewis-acids by directly binding phosphate oxygen atoms and polarizing them (50).

The exact mechanism of the cleavage reaction of the VS ribozyme is unknown. There is a requirement for the presence of metal ions for activity with magnesium being the most preferred divalent metal ion (22). Other ions such as calcium, manganese and certain monovalent ions support catalysis, whereas cobalt hexammine supports the correct folding of the ribozyme, but not catalysis (22,51-53). The fact that the VS ribozyme can cleave in the absence of divalent metal ions has cast doubt on the idea of a metal ion-mediated mechanism (53). The lack of a pH dependence on the rate of the cleavage reaction argues against the mechanism proceeding by a general acid-base mechanism (22), but pH dependence corresponding to a pK_a of 5.6 has been observed for the ligation reaction (50). It is also possible that the rate-limiting step is not the chemistry of cleavage, but some conformational change that occurs within the ribozyme. It has

also been suggested by NAIM analysis that ionization of A756 in the A730 loop plays a role in catalysis (54).

Whereas the role of magnesium ions in the catalytic chemistry of the VS ribozyme is uncertain, there is no doubt about its importance in the folding of the ribozyme. The secondary structure of the VS ribozyme forms in the absence of divalent metal ions, and the addition of magnesium ions leads to the formation of the tertiary structure (20). Formation of the loop-loop interaction and the conformational change in the substrate is magnesium-ion dependent (26-28). The loop-loop interaction also forms in the presence of cobalt hexammine, but is not as stable as in the presence of magnesium ions (52). The K_D for magnesium-ion binding to the *trans*-cleaving ribozyme was reported to be 16 mM (55).

Metal-ion binding sites in the VS ribozyme were inferred from phosphorothioate substitution interference, using Mn^{2+} to rescue the reaction (56). This method identified important metal-ion binding sites that affected the activity of the VS ribozyme. The important divalent metal-ion binding sites in the VS ribozyme that were identified by this method are shown in Figure 1.6. Magnesium-ion binding to the substrate was also investigated using single 2'-deoxyribonucleoside substitution (55). It was found that there is a possible magnesium-ion binding site in the terminal loop of stem-loop I, with the K_D for magnesium ions of residues U628 and G630 being the most affected by the 2'-deoxyribonucleoside substitution.

It is interesting to note that in both stem-loops I and V, a phosphate group has been implicated in a divalent metal-ion interaction (Figure 1.6). These groups are the G690 3'P and the A698 3'P for stem-loops I and V, respectively. These two phosphates occupy equivalent positions within their respective proposed U-turn motifs, which is the R 3'P of the UNR sequence. These are the same phosphates that would be involved in one of the stabilizing

hydrogen bonds of the U-turn. This hydrogen bond between U H3 and R 3'P is found in all U-turns characterized to date, and is thought to be the crucial interaction that stabilizes the turn in the backbone (46). So it can be assumed that magnesium ions bind to the U-turn motifs in stem-loops I and V.

Study of RNA structure by NMR spectroscopy.

The two methods that are commonly used for determination of high-resolution structures of biological macromolecules are X-ray crystallography and solution state Nuclear Magnetic Resonance (NMR) spectroscopy. In recent years, there has been an explosion in the number of structures of proteins and nucleic acids that have been solved by these methods. Nucleic acid structures are lagging behind those of proteins because they are more difficult to crystallize and there is extensive chemical shift overlap in their NMR spectra (57). For small RNA molecules, it has been easier to overcome the latter problem in recent years than trying to crystallize nucleic acids.

Apart from the problem of crystallization, an advantage of using NMR to obtain high-resolution structure of RNA is that it is a solution structure. Crystal packing forces can affect the structure of RNA molecules. For examples, these forces can result in hairpins forming less biologically-relevant hairpin structures (58). A solution structure allows for information on dynamics and also to investigate flexible regions of the molecule. Dynamics in RNA can be studied at the atomic level and a wide range of timescales (from picosecond to hours) can be investigated (59).

There are several steps involved in solving a RNA structure by NMR spectroscopy (60). The first step involves choosing a suitable system that is amenable to NMR. Then, due to the

intrinsically low sensitivity of the technique, milligram amounts RNA need to be synthesized and purified. Because of the limited resolution of the naturally occurring nuclei the RNA has to be isotopically labeled. The resonance assignment of as many of the NMR-active nuclei in the sample as possible is the third step. The fourth step involves obtaining the structural restraints from NOE data (through space correlation), J-coupling data (through-bond correlations) and residual dipolar coupling data (vector alignment correlation). The fifth and final step is to use the restraint data to calculate and refine a high-resolution three-dimensional structure.

NMR methods can also be useful to elucidate magnesium-ion binding sites in RNA by chemical shift mapping (61). Chemical shift mapping involves looking at the effect that the addition of magnesium ions to a sample has on chemical shifts (62). Typically, only those residues that bind or are in proximity to the metal ion will have their chemical shifts greatly affected. Other metal ions can be used as probes for magnesium-ion binding. Manganese ions can be used as a probe for magnesium-ion binding (63). It is a paramagnetic metal that broadens the resonance linewidths of atoms that are within 10 Å of it. The effect is dependent on the distance between the nuclei and the ion, as well as, the degree of occupancy of the binding site. Cobalt hexammine can also be used to model RNA-magnesium ion interactions (64). The eighteen protons in the hexammine ligands give proton-proton NOEs to protons that are within proximity.

Purpose of Study

All of the small self-cleaving ribozymes, except for the VS ribozyme, recognize their substrate by forming base-pairing interactions between single-stranded regions of the catalytic domain and the complementary single-stranded substrate (14). The VS ribozyme is unique in that it

recognizes a hairpin as its substrate, and its mode of substrate recognition involves tertiary interactions between the catalytic domain and the substrate domain (21). This unique mode of substrate recognition makes the VS ribozyme an interesting model to study RNA tertiary interactions.

Since the loop-loop interaction is an important step in the substrate recognition of the VS ribozyme, it would be interesting to study it structurally. The VS ribozyme is the largest of the small catalytic RNA that has not had its tertiary structure solved by either NMR spectroscopy or X-ray crystallography. There have been three structures published of the substrate domain, but they all focused on the cleavage site and not the terminal loop that participates in the loop-loop interaction (29-31). There has been no high-resolution structural study published on stem-loop V. The terminal loop of stem-loop V was investigated using biochemical techniques and was postulated to contain a U-turn motif (26,33,65). With a five-member loop it is plausible that these biochemical observations could be explained by the formation of some other turn structures. So it is important to obtain a high-resolution structure of stem-loop V to determine conclusively the structural motif present in stem-loop V. The formation of the loop-loop interaction is magnesium-ion dependent since stem loop V can shift the stem-loop I conformation by itself but only in the presence of magnesium ions (28). A magnesium-ion binding site has been postulated in the terminal loop of stem-loop V (56). Structural investigation of the binding of magnesium ions to stem-loop V would give some insight into the formation of the loop-loop interaction.

The main goal of this project was to understand structurally the mode of substrate recognition, with emphasis placed on the formation of the loop-loop interaction and its dependence for magnesium ions. This would be achieved by solving a high-resolution structure

of stem-loop V by NMR in the absence of magnesium ions. Then the effect of magnesium ions on the structure of stem-loop V would be investigated and the binding sites of magnesium ions would be elucidated.

In Chapter 2 of this dissertation I will present the strategies used in obtaining the resonance assignments and structural restraints that were used in determining the structure of stem-loop V. In Chapter 3, I will present the high-resolution structure of stem-loop V in the absence of divalent metal ions. I will describe the structural motif that is found in the terminal loop of stem-loop V. In Chapter 4, I will present the high-resolution structure of stem-loop V in the presence of near-saturating amounts of magnesium ions. I will show the effect of the magnesium ions on the structural motif found in the terminal loop of stem-loop V. I will also use manganese ions to localize divalent metal ion binding sites in stem-loop V. In Chapter 5, I will conclude by summarizing the results of this project and the briefly discuss some future directions.

Figure 1.1 Rolling-circle replication of the *Neurospora* VS DNA plasmid (see text) (18) .

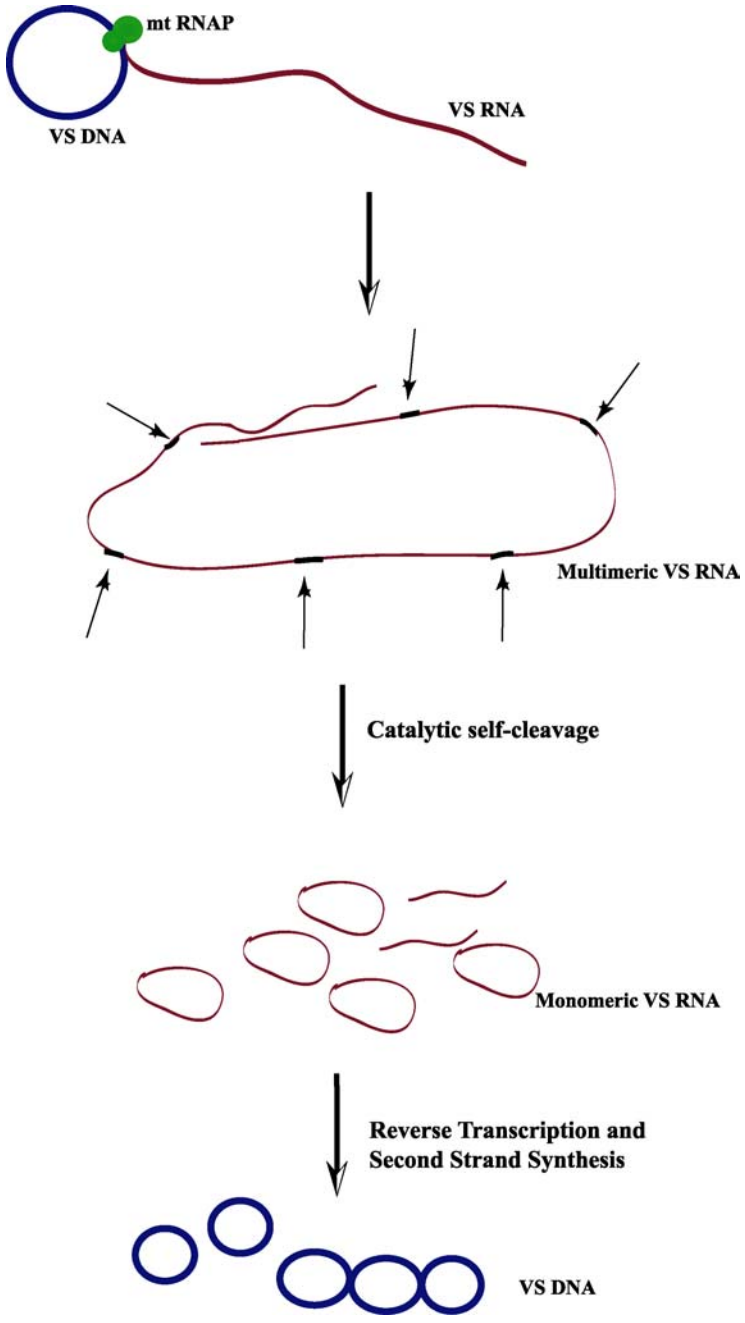


Figure 1.2 Secondary structure of the *Neurospora* VS ribozyme (20). Helices are represented by Roman numerals and lower-case letters represent non-native nucleotides. The arrow indicates the cleavage site and the small (blue) and large (green) rectangles designate of the substrate and catalytic domains, respectively.

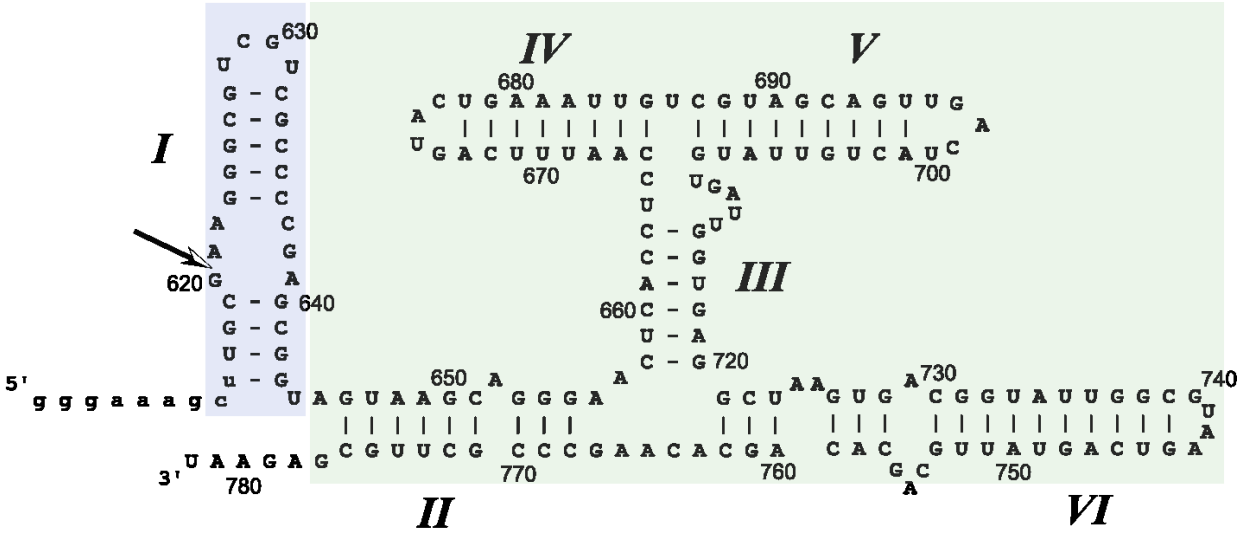


Figure 1.3. Loop-loop interaction in the *Neurospora* VS ribozyme. Proposed conformational change in stem-loop I (boxed) upon binding stem-loop V of the catalytic domain of the VS ribozyme in the presence of magnesium ions (26). The circled nucleotides have been proposed to participate in a loop-loop interaction.

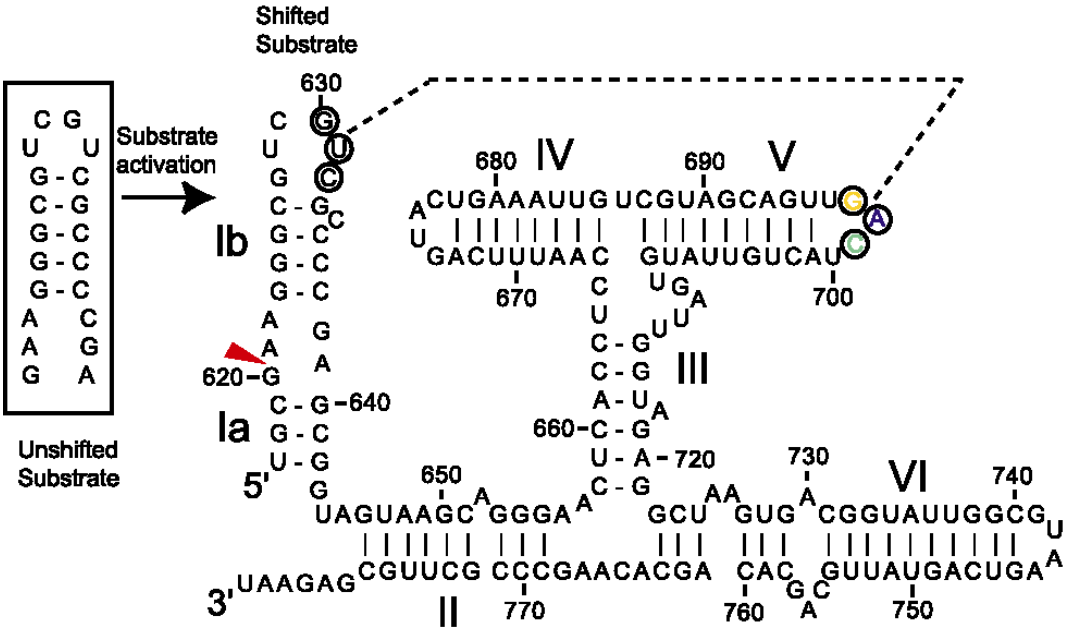


Figure 1.4. Tertiary model of the catalytic domain of the VS ribozyme. A) Schematic of the secondary structure of the VS ribozyme obtained from comparative gel electrophoresis and FRET analysis. B) A model for the complete catalytic domain of the VS ribozyme. Figure taken from reference (50).

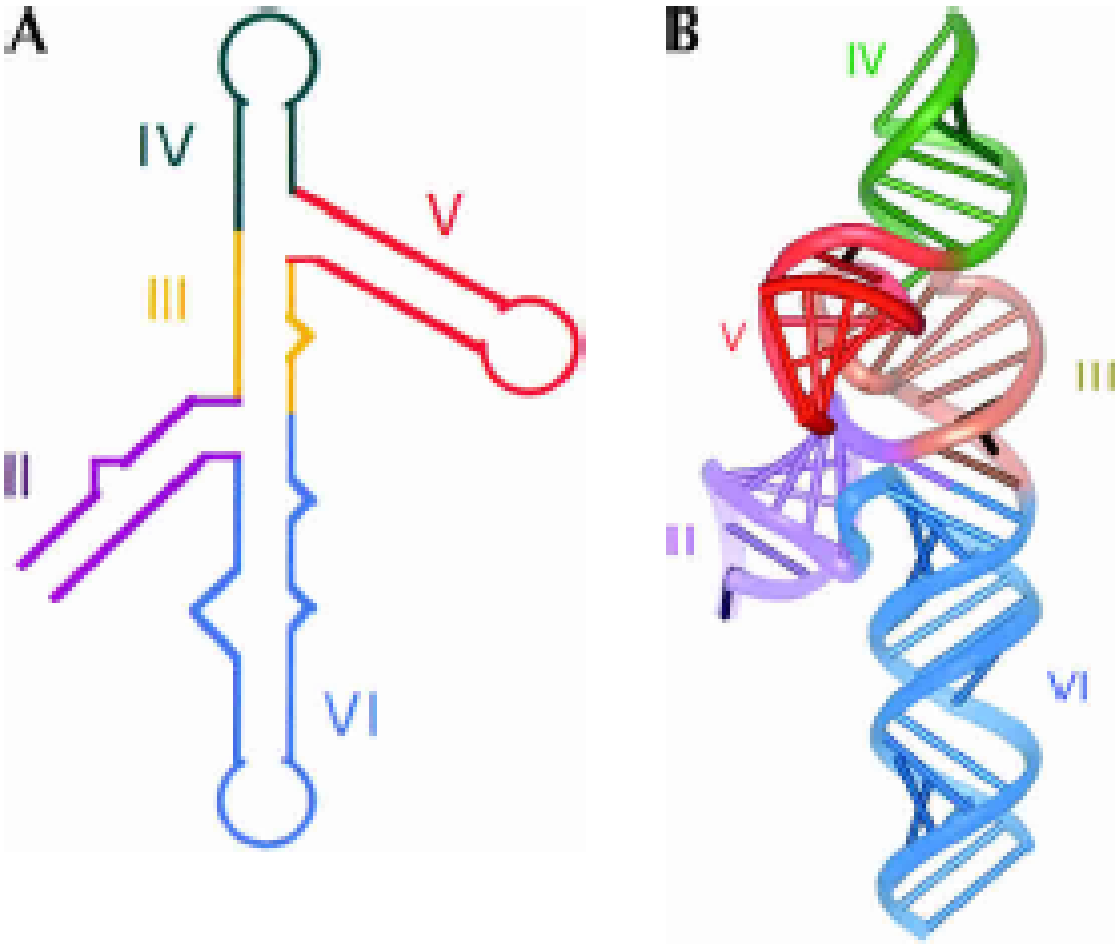


Figure 1.5. Schematic representation of a canonical U-turn motif.

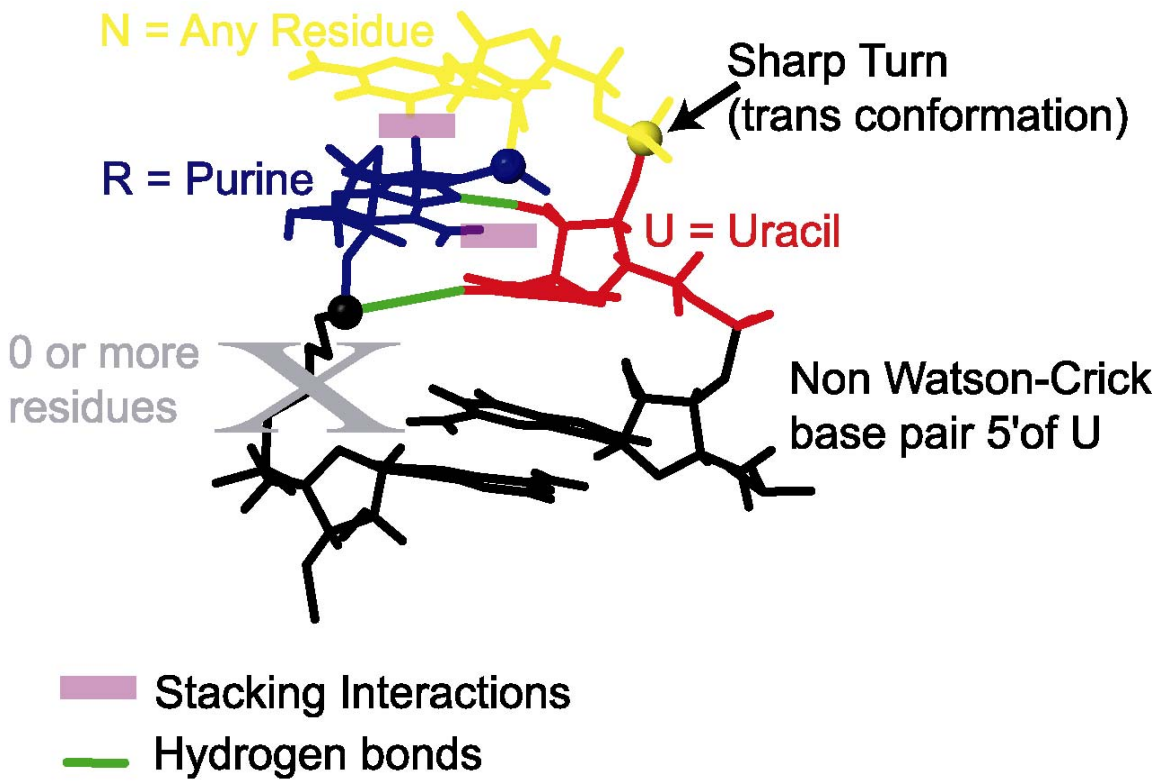
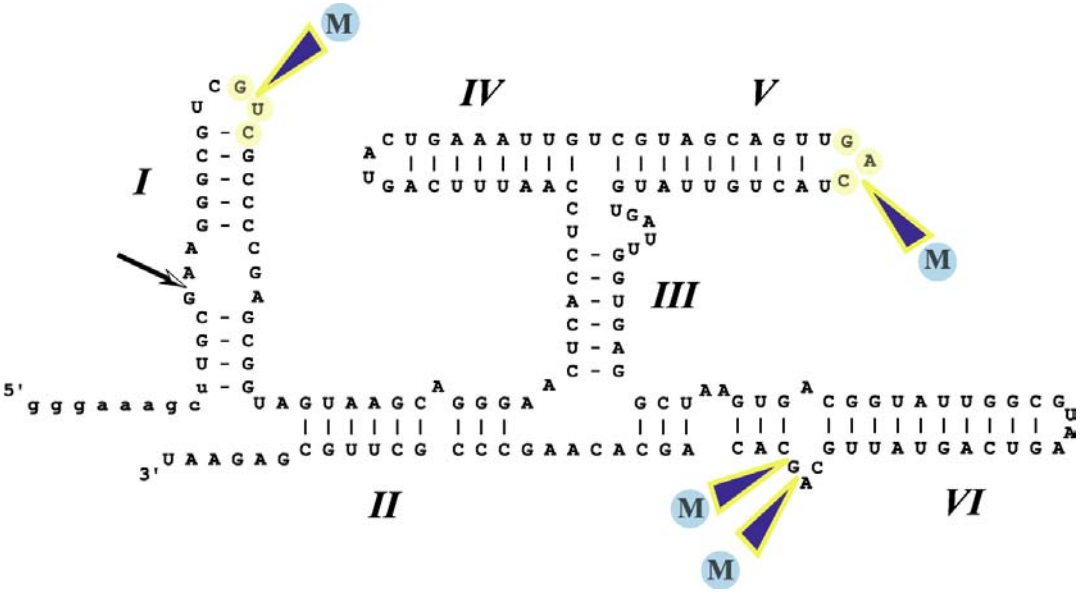


Figure 1.6. Metal-binding sites in the VS ribozyme. Important Mn^{2+}-coordination sites in the *Neurospora* VS ribozyme determined by phosphorothioate interference assays and manganese rescue (56). Sulfur substitutions of the non-bridging oxygen at specific phosphate groups inhibited the ribozyme activity, which could be rescued at certain sites by replacing magnesium ions by manganese ions in the catalytic reaction, suggesting direct divalent metal ion binding at these sites (circled M symbols).



References:

1. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, **35,** 849-857.

2. Zaug, A.J. and Cech, T.R. (1986) The intervening sequence RNA of Tetrahymena is an enzyme. *Science*, **231,** 470-475.

3. DeRose, V.J. (2002) Two decades of RNA catalysis. *Chem Biol*, **9,** 961-969.

4. Wilson, D.S. and Szostak, J.W. (1999) In vitro selection of functional nucleic acids. *Annu Rev Biochem*, **68,** 611-647.

5. Jaschke, A., Stuhlmann, F., Bebenroth, D., Keiper, S. and Wombacher, R. (2002) Ribozyme-catalysed carbon-carbon bond formation. *Biochem Soc Trans*, **30,** 1137-1140.

6. Cech, T.R. (1988) Ribozymes and their medical implications. *JAMA*, **260,** 3030-3034.

7. Lai, M.M. (1995) The molecular biology of hepatitis delta virus. *Annu Rev Biochem*, **64,** 259-286.

8. Buzayan, J.M., Gerlach, W.L. and Bruening, G. (1986) Non-enzymatic cleavage and ligation of RNAs complementary to a plant virus satellite RNA. *Nature*, **323,** 349-353.

9. Uhlenbeck, O.C. (1987) A small catalytic oligoribonucleotide. *Nature*, **328,** 596-600.

10. Saville, B.J. and Collins, R.A. (1990) A site-specific self-cleavage reaction performed by a novel RNA in *Neurospora* mitochondria. *Cell*, **61,** 685-696.

11. Muth, G.W., Ortoleva-Donnelly, L. and Strobel, S.A. (2000) A single adenosine with a neutral pKa in the ribosomal peptidyl transferase center. *Science*, **289,** 947-950.

12. Nissen, P., Hansen, J., Ban, N., Moore, P.B. and Steitz, T.A. (2000) The structural basis of ribosome activity in peptide bond synthesis. *Science*, **289,** 920-930.

13. Lilley, D.M. (1999) Structure, folding and catalysis of the small nucleolytic ribozymes. *Curr Opin Struct Biol*, **9,** 330-338.

14. Tanner, N.K. (1999) Ribozymes: the characteristics and properties of catalytic RNAs. *FEMS Microbiol Rev*, **23,** 257-275.

15. Doudna, J.A. and Cech, T.R. (2002) The chemical repertoire of natural ribozymes. *Nature*, **418,** 222-228.

16. Guo, H.C., De Abreu, D.M., Tillier, E.R., Saville, B.J., Olive, J.E. and Collins, R.A. (1993) Nucleotide sequence requirements for self-cleavage of *Neurospora* VS RNA. *J Mol Biol*, **232,** 351-361.

17. Saville, B.J. and Collins, R.A. (1991) RNA-mediated ligation of self-cleavage products of a *Neurospora* mitochondrial plasmid transcript. *Proc Natl Acad Sci U S A*, **88,** 8826-8830.

18. Kennell, J.C., Saville, B.J., Mohr, S., Kuiper, M.T., Sabourin, J.R., Collins, R.A. and Lambowitz, A.M. (1995) The VS catalytic RNA replicates by reverse transcription as a satellite of a retroplasmid. *Genes Dev*, **9,** 294-303.

19. Rastogi, T. and Collins, R.A. (1998) Smaller, faster ribozymes reveal the catalytic core of Neurospora VS RNA. *J Mol Biol*, **277,** 215-224.

20. Beattie, T.L., Olive, J.E. and Collins, R.A. (1995) A secondary-structure model for the self-cleaving region of *Neurospora* VS RNA. *Proc Natl Acad Sci U S A*, **92,** 4686-4690.

21. Guo, H.C. and Collins, R.A. (1995) Efficient trans-cleavage of a stem-loop RNA substrate by a ribozyme derived from *Neurospora* VS RNA. *EMBO J*, **14,** 368-376.

22. Collins, R.A. and Olive, J.E. (1993) Reaction conditions and kinetics of self-cleavage of a ribozyme derived from *Neurospora* VS RNA. *Biochemistry*, **32,** 2795-2799.

23. Zamel, R., Poon, A., Jaikaran, D., Andersen, A., Olive, J., De Abreu, D. and Collins, R.A. (2004) Exceptionally fast self-cleavage by a Neurospora Varkud satellite ribozyme. *Proc Natl Acad Sci U S A*, **101,** 1467-1472.

24. Lafontaine, D.A., Norman, D.G. and Lilley, D.M. (2002) Folding and catalysis by the VS ribozyme. *Biochimie*, **84,** 889-896.

25. Jones, F.D., Ryder, S.P. and Strobel, S.A. (2001) An efficient ligation reaction promoted by a Varkud Satellite ribozyme with extended 5'- and 3'-termini. *Nucleic Acids Res*, **29,** 5115-5120.

26. Rastogi, T., Beattie, T.L., Olive, J.E. and Collins, R.A. (1996) A long-range pseudoknot is required for activity of the *Neurospora* VS ribozyme. *EMBO J*, **15,** 2820-2825.

27. Andersen, A.A. and Collins, R.A. (2000) Rearrangement of a stable RNA secondary structure during VS ribozyme catalysis. *Mol Cell*, **5,** 469-478.

28. Andersen, A.A. and Collins, R.A. (2001) Intramolecular secondary structure rearrangement by the kissing interaction of the *Neurospora* VS ribozyme. *Proc Natl Acad Sci U S A*, **98,** 7730-7735.

29. Flinders, J. and Dieckmann, T. (2001) A pH controlled conformational switch in the cleavage site of the VS ribozyme substrate RNA. *J Mol Biol*, **308,** 665-679.

30. Hoffmann, B., Mitchell, G.T., Gendron, P., Major, F., Andersen, A.A., Collins, R.A. and Legault, P. (2003) NMR structure of the active conformation of the Varkud satellite ribozyme cleavage site. *Proc Natl Acad Sci U S A*, **100,** 7003-7008.

31. Michiels, P.J., Schouten, C.H., Hilbers, C.W. and Heus, H.A. (2000) Structure of the ribozyme substrate hairpin of Neurospora VS RNA: a close look at the cleavage site. *RNA*, **6,** 1821-1832.

32. Hiley, S.L. and Collins, R.A. (2001) Rapid formation of a solvent-inaccessible core in the *Neurospora* Varkud satellite ribozyme. *EMBO J*, **20,** 5461-5469.

33. Sood, V.D. and Collins, R.A. (2001) Functional equivalence of the uridine turn and the hairpin as building blocks of tertiary structure in the *Neurospora* VS ribozyme. *J Mol Biol*, **313,** 1013-1019.

34. Sood, V.D. and Collins, R.A. (2002) Identification of the catalytic subdomain of the VS ribozyme and evidence for remarkable sequence tolerance in the active site loop. *J Mol Biol*, **320,** 443-454.

35. Lafontaine, D.A., Norman, D.G. and Lilley, D.M. (2001) Structure, folding and activity of the VS ribozyme: importance of the 2-3-6 helical junction. *EMBO J*, **20,** 1415-1424.

36. Lilley, D.M. (2000) Analysis of global conformation of branched RNA species using electrophoresis and fluorescence. *Methods Enzymol*, **317,** 368-393.

37. Lafontaine, D.A., Wilson, T.J., Norman, D.G. and Lilley, D.M. (2001) The A730 loop is an important component of the active site of the VS ribozyme. *J Mol Biol*, **312,** 663-674.

38. Lafontaine, D.A., Norman, D.G. and Lilley, D.M. (2002) The global structure of the VS ribozyme. *EMBO J*, **21,** 2461-2471.

39. Hiley, S.L., Sood, V.D., Fan, J. and Collins, R.A. (2002) 4-thio-U cross-linking identifies the active site of the VS ribozyme. *EMBO J*, **21,** 4691-4698.

40. Quigley, G.J. and Rich, A. (1976) Structural domains of transfer RNA molecules. *Science*, **194,** 796-806.

41. Doudna, J.A. (1995) Hammerhead ribozyme structure: U-turn for RNA structural biology. *Structure*, **3,** 747-750.

42. Pley, H.W., Flaherty, K.M. and McKay, D.B. (1994) Three-dimensional structure of a hammerhead ribozyme. *Nature*, **372,** 68-74.

43. Lebars, I., Yoshizawa, S., Stenholm, A.R., Guittet, E., Douthwaite, S. and Fourmy, D. (2003) Structure of 23S rRNA hairpin 35 and its interaction with the tylosin-resistance methyltransferase RlmAII. *EMBO J*, **22,** 183-192.

44. Stallings, S.C. and Moore, P.B. (1997) The structure of an essential splicing element: stem loop IIa from yeast U2 snRNA. *Structure*, **5,** 1173-1185.

45. Puglisi, E.V. and Puglisi, J.D. (1998) HIV-1 A-rich RNA loop mimics the tRNA anticodon structure. *Nature Struct Biol*, **5,** 1033-1036.

46. Gutell, R.R., Cannone, J.J., Konings, D. and Gautheret, D. (2000) Predicting U-turns in ribosomal RNA with comparative sequence analysis. *J Mol Biol*, **300,** 791-803.

47. Pyle, A.M. (2002) Metal ions in the structure and function of RNA. *J Biol Inorg Chem*, **7,** 679-690.

48. Earnshaw, D.J. and Gait, M.J. (1998) Hairpin ribozyme cleavage catalyzed by aminoglycoside antibiotics and the polyamine spermine in the absence of metal ions. *Nucleic Acids Res*, **26,** 5551-5561.

49. Fedor, M.J. (2002) The role of metal ions in RNA catalysis. *Curr Opin Struct Biol*, **12,** 289-295.

50. Lilley, D.M. (2004) The Varkud satellite ribozyme. *RNA*, **10,** 151-158.

51. Collins, R.A. (2002) The *Neurospora* Varkud satellite ribozyme. *Biochem Soc Trans*, **30,** 1122-1126.

52. Maguire, J.L. and Collins, R.A. (2001) Effects of cobalt hexammine on folding and self-cleavage of the *Neurospora* VS ribozyme. *J Mol Biol*, **309,** 45-56.

53. Murray, J.B., Seyhan, A.A., Walter, N.G., Burke, J.M. and Scott, W.G. (1998) The hammerhead, hairpin and VS ribozymes are catalytically proficient in monovalent cations alone. *Chem Biol*, **5,** 587-595.

54. Jones, F.D. and Strobel, S.A. (2003) Ionization of a critical adenosine residue in the neurospora Varkud Satellite ribozyme active site. *Biochemistry*, **42,** 4265-4276.

55. Tzokov, S.B., Murray, I.A. and Grasby, J.A. (2002) The role of magnesium ions and 2'-hydroxyl groups in the VS ribozyme-substrate interaction. *J Mol Biol*, **324,** 215-226.

56. Sood, V.D., Beattie, T.L. and Collins, R.A. (1998) Identification of phosphate groups involved in metal binding and tertiary interactions in the core of the *Neurospora* VS ribozyme. *J Mol Biol*, **282,** 741-750.

57. Wijmenga, S.S. and van Buuren, B.N.M. (1998) The use of NMR methods for conformational studies of nucleic acids. *Prog NMR Spectr*, **32,** 287-387.

58. Chattopadhyaya, R., Grzeskowiak, K. and Dickerson, R.E. (1990) Structure of a T4 hairpin loop on a Z-DNA stem and comparison with A-RNA and B-DNA loops. *J Mol Biol*, **211,** 189-210.

59. Palmer, A.G., 3rd, Kroenke, C.D. and Loria, J.P. (2001) Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzymol*, **339,** 204-238.

60. Pardi, A. (1995) Multidimensional heteronuclear NMR experiments for structure determination of isotopically labeled RNA. *Methods Enzymol*, **261,** 350-380.

61. Feigon, J., Butcher, S.E., Finger, L.D. and Hud, N.V. (2001) Solution nuclear magnetic resonance probing of cation binding sites on nucleic acids. *Methods Enzymol*, **338,** 400-420.

62. Butcher, S.E., Allain, F.H. and Feigon, J. (2000) Determination of metal ion binding sites within the hairpin ribozyme domains by NMR. *Biochemistry*, **39,** 2174-2182.

63. Feig, A.L. (2000) The use of manganese as a probe for elucidating the role of magnesium ions in ribozymes. *Met Ions Biol Syst*, **37,** 157-182.

64. 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.

65. Sood, V.D., Yekta, S. and Collins, R.A. (2002) The contribution of 2'-hydroxyls to the cleavage activity of the *Neurospora* VS ribozyme. *Nucleic Acids Res*, **30,** 1132-1138.

CHAPTER 2

RESONANCE ASSIGNMENT AND STRUCTURAL RESTRAINTS OF SL5

Introduction

In order to calculate a high-resolution structure from NMR, experimental restraints need to be obtained. These restraints are used by a structure calculation program to arrive at a set of structures that satisfy the conditions of the experimental restraints. The precision and accuracy of the structure is dependent on the restraints used in the calculation, so it is of utmost importance to have a reliable set of restraints. The first step in obtaining these restraints is the accurate and unambiguous resonance assignment of as many NMR-active nuclei as possible. In order to increase the amount of NMR-active nuclei in RNA, the RNA had to be enriched with NMR-active isotopes. This involved synthesizing the RNA *in vitro* in the presence of isotopically labeled nucleoside triphosphates (NTPs). This synthesis of the RNA and NTPs was done in-our laboratory house. This meant that we had to modify the NMR construct used in this study to improve synthesis by T7 RNA polymerase. Unlabeled, ^{15}N-labeled, and ^{13}C/^{15}N-labeled RNA was synthesized for the purpose of this study.

In this chapter I will first go through the assignment procedure used to obtain the resonance assignments of SL5 in the absence of magnesium. I will then go through the process of obtaining structural restraints from the resonance assignments. The principles explained here were also used in the structure determination of SL5 in the presence of magnesium.

PART A: Resonance assignment of the ^{13}C, ^{15}N, and ^{1}H chemical shifts of SL5

A.1 Atomic composition of stem-loop V (SL5)

The stem-loop V RNA used for NMR studies (SL5) is a small RNA molecule composed of 17 nucleotides (Figure 2.1). There are eight purines (5 guanosines, 3 adenines) and nine pyrimidines (5 cytosines, 4 uracils) in the molecule (Figure 2.2). The atomic composition of SL5 consists of hydrogens, carbons, nitrogens, phosphorous, and oxygens. In all there are 199 hydrogen, 161 carbon, 63 nitrogen, and 17 phosphorous atoms in SL5. The hydrogen, carbon, and nitrogen atoms can be further classified as seen in Table 2.1. The numbering scheme used to identify the atoms is shown in Figure 2.2.

A.2 Assignment procedure for SL5

In order to determine a NMR structure, the very first objective is to obtain the resonance assignment of all the ^{1}H nuclei (protons) in the molecule of interest. It is very important to identify the individual protons because the main experiments used to obtain distance and torsion angle restrains require knowledge of the ^{1}H assignments. For small organic compounds it is usually sufficient to obtain 1D spectra of the protons and get the resonance assignment in this straightforward manner. However, a close inspection of Tables 2.1 and 2.2 will show that it is not that simple when dealing with a RNA molecule. The large number of protons coupled with the narrow chemical shift range in which they resonate results in a 1D ^{1}H spectrum of SL5 that has a tremendous amount of overlap (Figure 2.3). Multidimensional NMR methods has helped to a great extent alleviated this problem, by adding extra dimensions which enable us to resolve proton resonances (1,2). These extra dimensions are usually other nuclei frequencies that are in

some way connected to the resonance of interest. These other nuclei can be heteroatoms (such as ^{13}C and ^{15}N) or other protons, and they can be connected either through bonds or through three-dimensional space. So it is also important to obtain the resonance assignment of these associated nuclei, in order to obtain the proton resonances. In the case of SL5 it was helpful to resolve the proton resonances via ^{13}C and ^{15}N resonances that were present in the ^{13}C and/or ^{15}N-labeled sample. In order to do this, we also had to obtain assignment of the ^{13}C and ^{15}N resonances. Another very important reason for obtaining the assignments of the ^{15}N and ^{13}C heteroatoms in SL5 is that there are experiments that will correlate the NMR signal from these heteroatoms to a connected proton. These experiments reveal the connection between the two atoms that could be useful in determining a structure. The best examples of these types of experiments are those used to determine the torsion angles in RNA.

Even with the most perfectly resolved multidimensional NMR spectra it is still a challenge to decipher which resonances belong to which nuclei. I will go through the process of assigning the NMR-active nuclei using unlabeled, ^{15}N-labeled and $^{13}C/^{15}N$-labeled SL5 RNAs. The assignment of the various resonances is a cooperative event, that is, the assignment of one type of nuclei (for example a proton) may also give you the assignment of another type (say a carbon) depending on the experiment that produced the assignments. Also, the assignment of a particular nucleus may not be confirmed until after other nuclei of the same type are all assigned, and in some cases, until after the majority of the molecule is assigned. So for these reasons it is very difficult to describe the assignment procedure chronologically. Instead, I have decided to group the assignment strategy according to the types of nuclei as classified in Table 2.1. As such, within a particular section there may be references to assignments made in an earlier or later

section. I will try to explain the rational behind the assignment, as well as, give a brief description of the experiment used to make the assignment.

A.2.1 Exchangeable protons

The assignment of the NMR active nuclei in SL5 started with the assignment of the numerous hydrogen nuclei (protons) in the molecule. I intended to use proton-proton NOEs to solve the structure of SL5, so it was very important to obtain as many proton assignments as possible. Since protons are NMR-active, this meant that the assignment could proceed with an unlabeled RNA molecule. A 1D ^{1}H spectrum in H_2O is one of the simplest and quickest spectrums to collect (Figure 2.3). From Table 2.1 and Figure 2.3, it is seen that the most resolved protons are the exchangeable protons. Exchangeable protons are those protons that are in chemical exchange with the solvent and there are usually directly attached to an electronegative atom such as nitrogen and oxygen in the case of RNA. Depending on the rate of exchange, the proton resonance may be broadened (3). The resonance may be so broad that it may not be observed at all. The nitrogen-bonded exchangeable protons are the imino and amino protons, and they are found on the bases. The oxygen-bonded exchangeable protons are the hydroxyl groups found on the ribose moiety. The exchangeable protons that are bonded to oxygen in SL5 are in such fast exchange with the solvent that there were not detected at all on the NMR time scale of the experiments used in this study. The nitrogen-bonded exchangeable protons that are involved in interactions that slow down rotation about the C-N bond, and hence, slow down the rate of chemical exchange may be observed by NMR. The most common example of such an interaction is the base pair. Base pairing interactions are frequent in RNA as they stabilize the secondary and tertiary structures of the molecules. They involve both imino and amino protons with Watson-

Crick base pairs being the most commonly occurring types (Figure 2.4). Imino protons that are involved in base pairing interactions are typically observed by NMR. The amino protons of guanines and adenines are often too broad to be detected even when these residues are involved in base pairing interactions. The amino protons of cytosines are usually observed when they are involved in base pairing interactions. From the secondary structure of the SL5 construct (Figure 2.1) we expect to find six Watson-Crick base pairs, meaning that six imino protons signals and eight amino signals (from the base paired cytosines) that may be detectable.

A.2.1.1 Imino protons

The ^{1}H assignment of SL5 was initiated on an unlabeled sample at 25 ºC and pH 7.0 in 10 mM d_{11}-Tris, 50 mM NaCl, 0.2 mm EDTA, and 0.05 mM NaN_3. The assignment of SL5 was done at this temperature, pH and buffer conditions. The first step was to collect a complete 1D proton spectrum in H_2O. As can be seen from Table 2.2 the resonances for the imino protons of RNA are well resolved from the rest of the protons in the molecule. This provided an ideal starting point for the proton assignments. For SL5 at least six imino proton signals are expected in the 1D ^{1}H spectrum based on the proposed secondary structure (Figure 2.1.), as each Watson-Crick base pair should contribute one imino proton to the spectrum. The imino proton region of a 1D proton spectrum of SL5 is shown in Figure 2.5. Only four signals were observed. Two of the six imino signals were not observed because the base pairs at the base of stems and at the base of loops are not as stable as those in the stem.

In order to assign the imino protons, a 2D ^{1}H-^{1}H flip-back watergate NOESY (4,5) spectrum (collected in 90% H_2O:10% D_2O, mixing time 400 ms) was used. The 2D ^{1}H-^{1}H flip-back watergate NOESY spectrum correlates protons that are close together in three-dimensional

space. This is a through space correlation and is caused by a dipole-dipole relaxation mechanism (6). The resulting cross peaks are referred to as NOE (nuclear Overhauser effect) cross peaks. The spectrum is shown in Figure 2.6. The four imino proton signals are clearly seen in the diagonal. Diagonal peaks are peaks that have the same chemical shift in both dimensions. To start, a ^{1}H-^{15}N HSQC (7) optimized for the imino nitrogens (Figure 2.25) was used to determine the base type associated with each imino protons. One U imino proton is observed at 14.13 ppm and the other three imino protons belong to guanosine residues (Figure 2.25). Also it is known that the imino protons of A-U base pairs are usually downfield of those of G-C base pairs (8). The peak at 14.13 ppm is first assigned to the imino proton of U703. In a helix, imino protons have NOE cross peaks to imino protons that are in adjacent base pairs. So in this case there should be two cross peaks from the imino proton of U703 to the imino protons of G692 and G694. There are indeed two cross peaks, one at 13.47 and the other at 12.04 ppm. To determine which resonance belongs to which residue, the imino proton of G694 is first examined. The imino proton of G694 should exhibit only one cross peak to the imino proton of U703. The peak at 13.47 ppm was assigned to the imino proton of G694 since, there is only one cross peak for 13.47 and that is with 14.13 ppm (U703). That leaves the peak at 12.04 ppm to be assigned to the imino proton of G692. This imino proton gives an NOE to the imino proton of G705, which belongs to the adjacent G705-C691 base pair (Figure 2.6).

A.2.1.2 Amino protons

Since the observed imino protons are involved in base pairs then the next step in the assignment process was to use the same 2D ^{1}H-^{1}H flip-back watergate NOESY spectrum and find cross peaks between the imino protons and the amino protons of the residues involved in each base

pair. I was expecting to see the amino protons of the cytosines in stable G-C base pairs, as those of adenines and guanines exchange too fast to be observed in this spectrum (9). The amino protons of the cytosines that are base paired with the guanosines gave NOE cross peaks to the observable imino protons of the guanosines (Figure 2.7). The imino proton of G692 gave cross peaks to the amino protons of C704. The imino proton of G694 gave cross peaks to the amino protons of C702. The imino proton of G705 gave cross peaks to the amino protons of C691. It is important to note here that in the same region of the spectrum the H2 of adenines give cross peaks with the imino proton of their base-paired uracils. Hence, the intense cross peak to the imino proton of U703 in Figure 2.7 is the H2 of A693. This NOESY experiment only showed peaks for three of the five cytosine amino protons in the molecule and no peaks for the amino protons belonging to the three adenines and five guanosines in the molecule. The other two cytosines had their amino protons assigned from a 2D H(NCCC)H spectrum (10). This experiment correlates the exchangeable protons with the non-exchangeable H6 protons of cytidines via the J-coupling of the intervening bonds. As can be seen from Figure 2.8, all the amino protons of the cytosines were seen in this experiment. The unknown amino protons were assigned from their correlation to their H6 resonances that were assigned later. For four out of five cytosines, we observed two signals for the two amino protons, which indicate that at least one amino proton is hydrogen bonded. It is interesting to note here that there is a weak single resonance for the amino protons of C699, which suggests that these protons are not involved in any hydrogen bonding interactions. The single resonance is due to chemical exchange caused by rotation about N4-C4 bond.

The amino protons for two of the three adenines and for two of the five guanosines were assigned from a 2D ^{1}H-^{15}N CPMG-NOESY spectrum (9). This experiment allows for the

detection of NOEs involving exchanged-broadened amino protons. It achieves this by first detecting proton-proton NOEs and then transferring the magnetization to the attached nitrogen(s) (Figure 2.9). In order for these peaks to be observed the spectrum had to be collected at 5 ºC. At low temperature the rate of chemical exchange about the C4-N4 bond of the amino protons is reduced. When the experiment was repeated at 25 ºC, only the amino peaks of G697 and A698 were observed (not shown). These peaks were assigned by their attached ^{15}N resonance (Table 2.2) and by cross peaks that they gave with protons that were already assigned. The amino protons of A698 and A701 were observed in the spectrum but the amino protons of A693 could not be unambiguously assigned. The amino protons of guanosine G694 and G697 were assigned. For the residues in the stem that had their amino protons assigned, there were separate signals seen for the protons, whereas for the residues in the loop that had their amino protons assigned only a single signal was seen. Only a single amino proton signal was observed for the bases G697, A698, and C699, which indicates that these amino protons are not involved in any base pairing interactions.

A.2.2 Non-exchangeable protons

A.2.2.1 Base protons

From the 2D H(NCCC)H experiment (Figure 2.8) it was shown that the amino protons of the cytosines can be correlated to their H6 resonances. There are other experiments that give through-bond correlations between non-exchangeable protons and the exchangeable protons via the J-couplings of the intervening bonds. These experiments have the added advantage of being nucleotide specific. These experiments are the 2D H(NC)-TOCSY-(C)H for guanosines (11), and 2D H(NCCC)H for uracils (10). Figures 2.10 and 2.11 shows the assignments obtained from

these experiments. In the case of these experiments however, only the four observable imino protons seen previously were observed in these spectra.

Another helpful nucleotide-specific experiment is the 2D HNC-TOCSY-CH (12) for adenines. This experiment correlates the H2 and H8 with the N6 via the J-coupling of the intervening bonds. This is helpful for I already have the N6 and H2 assignments from the 2D ^{1}H-^{15}N CPMG-NOESY (Figure 2.6). Also this will be helpful in the sequential walk, which will be used to assign all the H8 and H6 atoms in SL5 in a sequence-specific manner (Section A.2.2.2), as the 2D HNC-TOCSY-CH helps correlates intra-residue H2s and H8s of the adenines (Figure 2.12).

Just before I get to the sequential walk, two other experiments must be analyzed. They are the 2D ^{1}H-^{15}N MQ-(HC)N(C)H (13) and the 2D DQF-COSY. The 2D ^{1}H-^{15}N MQ-(HC)N(C)H experiment correlates the H1' and the H6 (for pyrimidines) or H8 (for purines) with the N1 (for pyrimidines) or N9 (for purines). These correlations are through-bond correlations. The advantage of this experiment is the fact that the purine resonances are resolved from the pyrimidines due to the different chemical shift range of the N1 and N9 chemical shifts (Table 2.2). Also the cytosines are resolved from the uracils, due to the different chemical shift ranges of the N1 chemical shifts. Figure 2.13 shows the various regions of the 2D ^{1}H-^{15}N MQ-(HC)N(C)H spectrum. Figure 2.13A shows the cytosine region of the spectrum. The H6 of all the cytosines were assigned from the 2D H(NCCC)H spectrum (Figure 2.8). The H6 of C704 and C691 clearly correlates unambiguously with their respective H1'. However, at the H6 resonance for C702 (7.50 ppm) there are two signals (7.50 ppm and 7.51 ppm-the error in that dimension is such that the peaks are considered to be at the same frequency). Also, at the N1 frequency of 151.07 ppm there are two H1' and two H6 peaks at that frequency. To determine which H1'

frequency is paired with which H6 resonance, we reexamined the 2D ^{1}H-^{1}H flip-back watergate NOESY (Figure 2.16). The region of the spectrum to look at is the H1'-H6/H8 region. At the H6 frequency of 7.63 ppm there is a cross peak at 5.86 ppm, but not at 5.46 ppm. At the H6 frequency of 7.51 ppm there is a cross peak at 5.46 ppm, but not at 5.86 ppm. This means that in Figure 2.13A, the peak at 7.63 ppm correlates with 5.86 ppm and the peak at 7.51 ppm connected with 5.46 ppm. The same procedure was followed for the uracils and the purines (Figures 2.13B and 2.13C respectively). The H8/H6 resonances were paired with their respective H1' resonances. If there was any question as to the pairing, the 2D ^{1}H-^{1}H flip-back watergate NOESY was checked for cross peaks. By looking at the 2D ^{1}H-^{1}H flip-back watergate NOESY, the peaks in the 2D ^{1}H-^{15}N MQ-(HC)N(C)H have been unambiguously correlated.

The second experiment is the 2D DQF-COSY (Figure 2.14). This spectrum gives cross peaks for protons that are J-coupled (three-bond coupling). The H5 and H6 protons of the pyrimidines are separated by three bonds (Figure 2.1) so H5-H6 cross peaks appear in this spectrum. The resonances of the H5 protons of all the cytosines can be assigned since the H6 resonances are known at this point. The H5 proton of U703 was also assigned similarly. The importance of this experiment at this point is that it indicates the position of the H5 protons, which will prevent any confusion during the sequential walk. The confusion could arise from the fact that the H5 protons resonate at the same frequency as the H1' protons that are used in the sequential walk.

A.2.2.2 Sequential walk

The sequential walk is an intra- and inter-nucleotide NOE correlation pattern that can be observed between the ribose protons (H1', H2') of one nucleotide and base proton (H6/H8) of

this nucleotide and of the nucleotide that is directly 3' to it (Figure 2.15). There are two connectivity pathways that are commonly used to complete a sequential walk, one utilizes the cross peaks formed between the H8/H6 and H1' resonances and the other utilizes the cross peaks formed between the H8/H6 and H2' resonances. Since the H1' resonances are the most resolved of the ribose protons, the connectivity pathway involving the H8/H6 and H1' cross peaks was the one used to complete a sequential walk in SL5. The walk is very important in assigning the resonances sequentially since the NOE cross peaks can only be formed in the 5'-3' direction (Figure 2.15). The walk does not take place in the reverse direction for the H1'-H8/H6 distance is too large (> 7 Å) in this direction for transfer of magnetization to take place.

The sequential walk is completed by connecting the intra-residue H6/H8-H1' cross peaks with the inter-residue ones. Before commencing the walk, it was beneficial to indicate on the spectrum, the cross peaks that were already known (Figure 2.16). All the H5-H6 cross peaks, were previously identified from the 2D DQF-COSY (Figure 2.14). From the 2D ^{1}H-^{15}N MQ-(HC)N(C)H (Figure 2.13) all the H1'-H6/H8 intra-residue cross peaks, as well as the nucleotide types were known. And finally since the adenosine H2s are known from the 2D HNC-TOCSY-CH (Figure 2.12) and the H1' are known from the 2D ^{1}H-^{15}N MQ-(HC)N(C)H (Figure 2.13), the intra-residue H1'-H2 cross peaks are labeled. These cross peaks were first highlighted in the 2D NOESY collected in D_2O (Figure 2.16). The position of the C691 H1'-H6 cross peak is known from the combination of 2D ^{1}H-^{15}N MQ-(HC)N(C)H (Figure 13A) and 2D H(NCCC)H (Figure 2.8), and from assignment of amino protons (Section A.2.1.2). By looking along the H6 resonance, three peaks are seen that are correlated to the C691 H6 (these peaks appears as doublets in F2 due to the H5-H6 coupling) (Figure 2.16). Two of these cross peaks are already known (the intra-residue H5-H6 and H1'-H6 at 5.27, 7.86 and 5.64, 7.86 ppm respectively), and

the third one occurs at a frequency that corresponds to a guanine H1' and C691 H6 cross peak (5.70, 7.86 ppm) (Figure 2.13C). This may be G690, which would be the starting point for the sequential assignments. To ensure this, the H8 frequency is checked for other cross peaks at 8.00 ppm, since as the first residue it would only give a single intra-residue H1' cross peak. There are no other peaks at that frequency, so this cross peak was assigned to G690. Starting from the G690 H1'-H8 cross peak, the G690 H8/C691 H1' cross peak is found. Then in a vertical direction (along the C691 H6 resonance) the highlighted C691 H1'/H6 cross peak is found. The steps are then repeated to continue the walk. For the section of the spectrum that is shown in Figure 2.16, the vertical lines connect an inter-residue $H1'_{i-1}$-$H8/H6_i$ cross peak to an intra-residue $H1'_i$-$H8/H6_i$ cross peak, whereas horizontal lines connect an intra-residue $H1'_i$-$H8/H6_i$ cross peak to an inter-residue $H1'_i$-$H8/H6_{i+1}$ cross peaks. By knowing the sequence and highlighting the known cross peaks the walk was completed very easily. If these cross peaks were not highlighted completing the walk would have been difficult as the H1' and H5 resonate in the same region of the spectrum. Another way to complete the sequential walk would be to do the walk in 3D using a 3D ^{13}C-edited HMQC-NOESY (14). The 3D ^{13}C-edited HMQC-NOESY resolves proton-proton NOEs by the resonance of their attached carbons. This means that a $^{13}C/^{15}N$-labeled SL5 sample was needed. This would clear up any ambiguity that may occur during the 2D sequential walk. Figure 2.17 shows an example of an assignment using a 3D sequential walk.

A.2.2.3 Ribose protons

Once the resonance assignment of the base protons was completed the next step was to assign the ribose protons. Table 2.2 shows that all the ribose protons resonate in a narrow frequency range of the ^1H spectrum. This means that it would be difficult to assign them based on homonuclear

experiments, as the resolution of the ^{1}H spectrum is very poor. Another problem is the generally low values of the $J_{H1'-H2'}$ coupling constants in RNA. This makes the use of homonuclear J-coupling experiments difficult. Hence, multidimensional heteronuclear experiments were used, which were the 3D HCCH-COSY (15) and the 3D HCCH-TOCSY (15). This meant that a $^{13}C/^{15}N$ labeled SL5 sample was used. These experiments are through-bond experiments that correlate carbons that belong to the same spin system. The proton resonances are then assigned by one-bond correlation to their attached carbons. The difference between the two experiments lies in the number of bonds that separate the carbons. The HCCH-COSY experiment transfers magnetization between carbons that are directly bonded together, whereas a HCCH-TOCSY transfers the magnetization throughout the entire carbon spin system. By taking a plane of the 3D spectrum of the HCCH-COSY or HCCH-TOCSY at the C1' resonance of a particular residue, all the protons that are attached to carbons that are J-coupled to the C1' will give cross peaks to the H1' of that residue. Figure 2.18 shows a 2D $^{1}H(F1)$ $^{1}H(F3)$ plane of a 3D HCCH-COSY spectrum of SL5. The 2D $^{1}H(F1)$ $^{1}H(F3)$ plane shown is at the C1' chemical shift of G697 (93.6 ppm). Since the spectrum is in 3D the "diagonal" peak shown is known as an autocorrelation peak. In this spectrum the autocorrelation peak represent that proton directly attached to the C1', which is the H1'. The cross peak shown is to the attached proton of the carbon that is one-bond J-coupled to the C1'. That carbon is the C2', hence the cross peak represents the H2' resonance. If the plane was taken at the C2' resonance, then there would be two cross peaks. This is because the C2' is one-bond J-coupled to two carbons, the C1' and the C3'. Figure 2.19 shows the same C1' plane of G697 of the 3D HCCH-TOCSY spectrum. The 3D HCCH-TOCSY transfers magnetization through the entire ribose carbon spin system. Hence, cross peaks to the H2', H3', H4', H5' and H5" were expected.

To aid in the assignment, the 3D HCCH-TOCSY can be collected in a manner in which the H5'/H5" cross peaks are in anti-phase to the other cross peaks. This readily identifies the H5'/H5" resonances as shown in Figure 2.19B. Usually in an A-form helix the H5" resonance occurs upfield of the H5' resonance, which explains the choice of assignments in the figure, but it must be noted that this is not always the case (16). Five cross peaks were expected but only four are seen in the spectrum suggesting that there is an overlap of resonances. I know the overlap does not occur with the H5' or H5" resonances by looking at Figure 2.19B, because if this was the case the peaks would cancel out and not be observed because they were in opposite phases. To determine what two peaks are overlapped the spectrum in Figure 2.19A can be displayed differently. In the Figure 2.20 the 2D plane of the 3D HCCH-TOCSY spectrum of SL5 is displayed with ^{13}C(F2) dimension replacing the ^{1}H(F1) dimension. This resolves the protons according to the resonances of their attached carbons. From Figure 2.20 it is seen that the overlap in Figure 2.19A occurs between the H2' and H3' resonances.

A.2.3 Protonated carbons

Only carbons that are protonated were considered for the study of SL5. This is a consequence of the fact that the carbon resonances were used to resolve the proton resonances. This makes assigning the protonated carbons fairly straightforward. By collecting experiments that correlate the resonances of the carbon to its attached protons, then the resonance of the carbon is obtained. One such experiment is the 2D ^{1}H-^{13}C CT-HSQC (17,18). This spectrum gives a through-bond correlation between all protonated carbons and their directly attached protons. The full and unfolded 2D ^{1}H-^{13}C CT-HSQC spectrum of SL5 is shown in Figure 2.21. The resonances are resolved enough to distinguish the different classes of protonated carbons. There is also the

added benefit that some groups of signals have opposite phase compared to others in the spectrum with careful selection of the constant-time period, which is 9.5 ms. In instances where there is a overlap of proton resonances, a 3D HCCH-COSY and a 3D HCCH-TOCSY were used to resolve any ambiguity. Figures 2.21-2.24 illustrate assignment of the ^{13}C resonances from the 2D ^{1}H-^{13}C CT-HSQC.

A.2.4 Nitrogens

A.2.4.1 Protonated nitrogens

We used 2D ^{1}H-^{15}N HSQCs (7) and 2D ^{1}H-^{15}N HMQC (19) as simple ways to assign resonances for protonated nitrogens. The experiments can be collected in H_2O to observe the exchangeable protons directly or, in cases where the exchangeable protons are not observed, in D_2O where the protonated nitrogens are correlated to non-exchangeable protons. Figure 2.25 shows the result of a 2D ^{1}H-^{15}N HSQC collected in H_2O and optimized to observe the imino protons. Once again only the four imino protons that were previously observed in the 1D ^{1}H spectrum (Figure 2.5) are observed here. The uridine N3 resonates downfield from the guanosine N1. Figure 2.26 shows the spectrum from the same experiment but it is optimized for the amino protons. The amino protons of the cytosine residues are all observed except for C699. As shown in the 2D H(NCCC)H spectrum (Figure 2.8), the amino protons of C699 give a single, weak peak at 6.77 ppm. It is possible that the single resonance expected for C699 is overlapped by the strong C704 signal. Other peaks seen in the spectrum, but not shown in the figure are the amino protons of G697 and A698. They both produce single peaks suggesting that the amino protons of these two residues are not involved in any base-paring interactions. It should be noted here that for each of the peaks assigned in Figure 2.26, there is a smaller signal that resonates slightly upfield in the

^{15}N dimension. This small peak is due to what is known as the isotope-shift effect. The experiment is collected in 90% H_2O and 10% D_2O (The D_2O is used to lock the NMR spectrometer). Since the amino protons are in exchange with the solvent then at any given time a small percentage of the amino nitrogens will have a deuterium atom in place of a proton. The presence of the deuterium (D) and the proton in NHD will cause a shift in the nitrogen resonance. And since the NHD is a second but smaller population of amino signals, there is an extra smaller peaking representing it. Since the amino protons of most of the purines were seen in the 2D ^{1}H-^{15}N CPMG-NOESY (Figure 2.9), the attached nitrogens were assigned from it. The only uracil to have its imino proton nitrogen resonance assigned was U703 (Figure 2.25). The experiment used to assign the imino nitrogens of the other uracils was a 2D ^{1}H-^{15}N HMQC (19) collected in D_2O and optimized for J = 7 Hz (19). Since it was collected in D_2O, the imino protons would not be observed. The 2D ^{1}H-^{15}N HMQC allowed for the correlation of uridine H5 to the N1 and N3, and cytosine H5 to the N1. The spectrum is shown in Figure 2.27 and all the uracil N3 are observed. An interesting note at this point is that the N3 resonances of U703 and U695 are downfield of U696 and U700. This may indicate that these nitrogens are involved in hydrogen-bonding interactions, as forming such an interaction would lead to the deshielding of the N3 nuclei leading to a more downfield chemical shift. We know that this is the case for U703 as it was shown to be involved in a Watson-Crick base pair with A693. So this suggests that U695 also forms a base-pairing interaction.

A.2.4.2 Non-protonated nitrogens

Although these nitrogens are not directly bonded to protons they can still be correlated to protons in the molecule. This is achieved through transferring the magnetization from a proton in the

molecule to the nitrogen that has to be detected as illustrated in Figure 2.27. This is made possible by experiments that exploit two-bond and three-bond J_{HN} coupling constants. I have already shown how a 2D ^{1}H-^{15}N HMQC was used to assign the pyrimidine N1 and N3 by their correlation to the H5. Another experiment that does this type of correlation is a modified HNN-COSY (20). From this experiment the N9 and N7 of the purines were correlated to their H8 and the N1 and N3 of the adenines were correlated to their own H2 resonances. All these correlation were through-bond ^{1}H-^{15}N correlations. Figures 2.28 and 2.29 show the results.

A.2.5 ^{31}P resonances.

The phosphorous resonances were assigned by collecting a 1D ^{31}P spectrum and a 2D ^{1}H-^{31}P HETCOR (21). The ^{1}H-^{31}P HETCOR experiment correlates the phosphorous resonance to the H3' resonance of the residue that is 5' to the phosphorous. This experiment also provided another connectivity pathway to do sequential assignments. Since this sequential assignment is made through bonds instead of through space, it provides unambiguous validation of the sequential assignments. Figure 2.30 shows the result of the experiment and the assignments of the 16 phosphates in SL5. The peaks have a multiplet pattern due to splitting associated with the active and passive J-coupling of the H3' and the phosphorus. The resulting pattern is similar to that seen in a homonuclear COSY experiment (Figure 2.14.).

PART B: Obtaining the structural restraints for the structure calculation of SL5

B.1. Restraints for structure calculation

There were two types of restraints used in the structure calculation of SL5, distance restraints and torsion angle restraints. The distance restraints were based on calibrated distances obtained from the NOE data. These gave a through-space relation between the protons in the NMR structure of SL5. The most important distance restraints were the inter-nucleotide distance restraints, as these really were the ones that impacted the structure as most intra-residue proton distances are fixed. The torsion angle restraints were angular restraints that restrained the torsion angles of the SL5 RNA. I will now describe how both sets of restraints were obtained for SL5, starting with how distance restraints are obtained from NOE data. Then I will describe how the torsion angles gamma, and delta were defined from J-coupling and NOE data. I will finish up by describing how the hydrogen bonds of the base pairs were detected and used as restraints for the structure calculation.

B.2 NOE distance restraints

B.2.1 Selection of NOESY mixing times

Before collecting the 3D ^{13}C-edited HMQC-NOESY (14) spectra used to obtain the restraints several 2D ^{1}H ^{1}H NOESY spectra with various mixing times were collected. These were used to ascertain which mixing times would be used for the 3D datasets. Two mixing times for the NOE dataset needed to be chosen, a short mixing time to collect all the short-range distance information, and a longer mixing time NOE to obtain long-range distance information. For SL5, 2D ^{1}H ^{1}H NOESY with the following mixing times were analyzed; 30, 60, 90, 120, 180, 240,

and 360 ms. At low mixing times the NOE intensities increased proportionally with the mixing time. At higher mixing times the intensities began to level off and even decrease. The decrease is due to spin-diffusion in which the nuclei begin to loose their magnetization due to relaxation processes, including NOE transfer to other protons (22). The more the intensity of an NOE builds up the greater chance it has of being detected in the spectra at that particular mixing time. At 90 ms there were a significant amount of signals seen and the signal to noise ration was very good. So this was chosen as the low mixing time for SL5. Also a long mixing time was needed to get long-range distance information between protons separated by longer distances (4-5 Å). From the assignment process in which a 3D ^{13}C-edited HMQC-NOESY with a mixing time of 240 ms was used for the sequential walk, it was noted that a lot of spin diffusion was seen in the spectrum. NOEs typically occur between protons that are within 5 Å of each other. The relaxation process that produces spin-diffusion gives NOE cross-peaks between protons that are greater than 5 Å apart, giving the appearance that they are closer together in the structure than they really are. For example, H_A may give an NOE to H_C without necessarily being close to it in space (> 5 Å) by direct NOE transfer from H_A to H_B and H_B to H_C. This obviously will lead to problems during calibration and assignment of the NOEs. So it is best to collect a long mixing time that minimizes on the amount of spin diffusion that occurs while maximizing signal intensity. The mixing time chosen was 180 ms. There will still be a small amount spin diffusion at the longer mixing time but they will be corrected for by giving the restraints a large upper bound.

B.2.2 Peak picking the NOESY spectra

Two 3D ^{13}C-edited HMQC-NOESY spectra were collected, with mixing times of 90 ms and 180 ms. All the resonances seen in these spectra were assigned previously. The process of choosing the peaks to use for the restraints and obtaining their intensities is referred to as peak picking. In order to pick a peak, the diagonal peak for the attached proton at the corresponding carbon plane is found. The carbon plane is then changed to find the plane at which the cross peaks are at their maximum intensities. This is the plane that the peaks are picked on. The program I used to pick the peaks was NMRView (23). All cross peaks are selected even if they represent overlapping resonances. Care was taken for peaks that are close to the diagonal. Due to the decoupling scheme used in collecting the experiments, there will be symmetrical artifacts about the diagonal. Once all the peaks are picked, they are then assigned as unambiguously as possible. After the peaks were picked, the intensities were extracted. Care should also taken at this step because there might be overlap between planes which would lead to greater than true intensities. Also finally, for as many peaks as possible, the peaks were paired with their symmetrical peak. What this means is that for every pair of protons that produce a NOE cross peaks, the intensities are obtained for the cross peaks at the plane of each of the attached carbons.

B.2.3 Calibration of the NOEs

B.2.3.1 Non-exchangeable protons

These NOEs were obtained from 3D ^{13}C-edited HMQC-NOESY (14) experiments collected in D_2O. The advantage of working in D_2O is to remove the intense water signal, which interferes with observation of ribose protons. Once the intensities were obtained distance information was inferred from them. This was done by calibrating the intensities of the NOEs in terms of

distances. For each NOE, its intensity will represent a distance range. The NOEs were then separated into classes based on their intensities. Each NOE that belongs to the same class will be given the same range of distances, but different classes will have different ranges. The closest approach for two non-bonded atoms is the sum of their van-der Waal's radii. This turns out to be approximately 1.8 Å for protons. Hence, for all the NOEs the closest distance possible is 1.8 Å. This is called the lower bound of the calibration. All the classes of NOEs will have the same lower bound for their range of distances. The differences between the classes of NOEs will be the upper bound of the range.

First the 90 ms dataset was looked at, as this is the dataset that provided the short-range distance classes. The NOEs between atoms that had known distances were first analyzed. This included all the intra-residue H5 to H6, H6/H8 to H1', H2 to H1', and H1' to either H2', H3' and H4' NOEs. The distances or distance ranges for these proton pairs are all known (2). The intensities of the symmetrical peaks were averaged. The most intense of the initial set of NOEs that were analyzed belonged to the H5 to H6 and H1' to H2' NOES. The H5 to H6 distance is fixed at 2.4 Å, whereas the H1' to H2' distance has a maximum distance of approximately 2.8 Å. For theses NOEs it was noted that the intensities were all greater than 0.8 (the units are arbitrary). This gave a starting point for the calibration as peaks with intensities greater or equal to 0.8 units were given an upper bound of 3.0 Å. All peaks that were put in this class of NOEs were labeled as VERY STRONG NOEs. Irrespective of the conformation of the nucleotide, the maximum distance between the intra-residue H1' and the H8/H6 atoms is approximately 3.9 Å (2). The intensities of the H6/H8 to H1' NOEs were generally between 0.2 and 0.3 units, hence, all NOEs with an intensity between 0.2 and 0.4 units were given an upper bound of 4.1 Å. These NOEs were placed in a class of NOEs labeled as STRONG NOEs. Those NOEs that had

intensities greater than 0.4 units but less than 0.8 units were also placed in the STRONG NOEs class. Continuing this analysis lead to all other NOEs that were seen in the 90 ms dataset were given and upper bound of 5.5 Å and placed in the class of MEDIUM NOES. For those NOEs which had higher than true intensities due to overlap, and their assignment was not to a known intra-residue distance, were put into the class with an upper bound of 5.5 Å. This ensures that they were not over-constrained during the structure calculation. From the 180 ms mixing time dataset another class of NOEs were assigned. These were the NOEs that were seen only in the 180 ms dataset and not the 90 ms dataset. These NOEs may constitute spin diffusion NOEs, but structural information can still be obtained from them. These NOEs are given an upper bound of 7.0 Å to account for the possible spin diffusion. This class of NOEs was referred to as WEAK NOEs. Table 1 lists the different classes of NOEs.

B.2.3.2 Exchangeable protons

These constraints were obtained from a 2D ^{1}H-^{1}H flip-back watergate NOESY (4,5) and 2D ^{1}H-^{15}N CPMG-NOESY spectrum (9) collected in H_2O. The mixing time for the 2D ^{1}H-^{15}N CPMG-NOESY was 150 ms and for the 2D ^{1}H-^{1}H flip-back watergate NOESY was 400 ms. The reason for the long mixing times was to maximize the intensity of imino proton to imino proton NOEs, which are separated by a large distance (4-5 Å) in A-form helices. Due to the long mixing time and hence, the accompanying spin diffusion it was difficult to assign the NOEs from this spectrum into different classes based on intensity. So all NOEs obtained from this spectrum was regarded as WEAK NOEs (1.8-7.0 Å) irrespective of intensity. The 2D ^{1}H-^{15}N CPMG-NOESY experiment allows for the detection of NOEs involving exchanged-broadened amino protons (Section A.2.1.2). This experiment was first collected using different temperatures (5, 15, and 25

°C) and various mixing times (150, 240, and 400 ms) in order to determine the best condition under which to collect the experiment. The spectrum with the most signals was the one collected at 5 ºC with a mixing time of 150 ms (Figure 2.9). The resonances of the signals changed only slightly at 5 ºC but assignment was still possible. Comparison of the signals seen at 5 ºC and 25 ºC showed that the SL5 RNA was still in the same conformation. At 25 ºC and 150 ms, most of the amino resonances seen at 5 ºC were not seen. This was because the higher temperature increased the rate of chemical exchange of the amino protons. Also at 25 ºC, the amino protons of G697 and A698 were observed. This meant that the data at 5 ºC can be used to obtain restraints for the stem and the data at 25 ºC was used to obtain restraints for the loop. The more intense NOEs were placed in the MEDIUM NOEs class (1.8-5.5 Å), and all the rest were placed in the WEAK NOEs class (1.8-7.0 Å).

B.3. Torsion angle restraints

In RNA there are twelve torsion angles that define the nucleotide conformation in RNA. α, β, γ, δ, ε, ζ, $\nu_{(0-4)}$, χ. The first six are the angles that define the backbone of RNA. $\nu_{(0-4)}$ defines the sugar pucker, the pseudorotation angles $\nu_{(0-4)}$ can be used to more precisely define the exact phase and amplitude of the sugar pucker, and χ is the angle of the glycosidic bond which links the sugar to the base (Figure 2.31). All these dihedral angles can also be constrained in the structure calculation. Constraints for the dihedral angles were obtained by either measuring coupling constants between certain protons or doing a comparative analysis of relative NOE intensities between certain protons. The experiments that were used to obtain coupling constants were those that gave through bond correlation of J-coupled protons. A 2D

DQF-COSY, and a 3D HCCH E.COSY were used in an attempt to restrain the torsion angles in SL5. Only the δ, and γ torsions angles were used in the structure calculation of SL5.

B.3.1 Sugar pucker and the backbone angle δ

The two most common conformations of the ribose furanose rings are the C3'-endo (N-type) and the C2'-endo (S-type). These conformations are commonly referred to as the sugar pucker. The C3'-endo conformation is most commonly found in A-form RNA and DNA, whereas the C2'-endo conformation is commonly found in B-DNA.

Two coupling constants that are particularly sensitive to the sugar pucker are $^{3}J_{H1'/H2'}$ and $^{3}J_{H3'/H4'}$. In the C3'-endo conformation $^{3}J_{H1'/H2'}$ is relatively small (< 2 Hz), while $^{3}J_{H3'/H4'}$ is relatively large (8 > Hz) (24). The opposite is true for C2'-endo sugar conformation. So by measuring the values of these couplings an estimate of the sugar puckering of the SL5 residues were obtained. For SL5 several methods were used to estimate the sugar puckers. First, a 2D DQF-COSY was used to look qualitatively at the $^{3}J_{H1'/H2'}$ coupling constants (Figure 2.32). The DQF-COSY is an experiment that correlates protons that are ^{3}J-coupled. The transfer of magnetizations between the coupled protons depends on the size of the coupling constant between them. The larger the coupling constant, the more magnetization is transferred, and this leads to more intense cross peaks. So by looking at the intensity of the H1'/H2' cross peaks for each residue the sugar pucker can be inferred; if the cross peak is intense then the sugar pucker is C2'-endo, if it is weak then the pucker is C3'-endo. Cross peaks of intermediate intensity indicate that the pucker is in an equilibrium conformation between the two conformations. It is also possible that the pucker is in some other less favorable sugar pucker conformation. An attempt was made to corroborate the sugar pucker conformation by doing a similar qualitative analysis of

the $^{3}J_{H3'/H4'}$ couplings. However, this was not possible as there was severe overlap in the H3'-H4' region of the spectrum. This was overcome by resolving the cross-peaks by their attached carbons. A 3D HCCH E. COSY (25) provided the necessary resolution and was used to calculate directly the value of the $^{3}J_{H1'/H2'}$ and $^{3}J_{H3'/H4'}$ coupling constants. The multiple splitting in the DQF-COSY spectrum made it difficult measure the couplings directly. The E. COSY pattern of the 3D HCCH E.COSY allowed for the coupling constants to be measured directly from the spectrum.

NOE intensity analysis was also used to determine the sugar puckers. This was achieved by using the distances seen in the idealized sugar pucker conformations and then doing comparative analysis of the intensities of the cross peaks that are affected by a change in conformation. Some of the intra-residue distances that are known to be affected by the sugar pucker are the H1'-H2', H3'-H4' and if the χ angle is known, H8/H6-H2' and H8/H6-H3' (24). The H1'-H2' and H3'-H4' NOE cross peaks were too overlapped to use them all in this analysis. Instead this analysis relied heavily on the H8/H6-H2' and H8/H6-H3' intensities. Essentially, if the pucker is C3'-endo then the H8/H6-H3' distance is shorter than the H8/H6-H2' distance, and vice versa for the C2'-endo conformation. The NOE method is the least realizable of the methods because the intensity of the NOE cross peaks can be affected by several factors other than the distance between the protons that gave rise to the NOE cross peaks. And the most important reason is that cross peaks can appear more intense than they really are because of overlap. Table 2.4 summarizes the results of the experiments used to determine the sugar pucker. It can be seen that all the residues in the stem, as well as, U696 are in the C3'-endo conformation, whereas G697-U700 are in a dynamic equilibrium. Since only the residues 690-696 and 701-706 had their

sugar puckers determined by this analysis, only they had their δ angle constrained. Their δ values were constrained to 86 +/- 10º (C3'-endo conformation).

B.3.2 γ torsion angle

The γ angles were constrained using comparative NOE intensity analysis. In this case the intra-residue H8/H6 -H5'/H5" NOE cross peaks were analyzed. This analyses would also help to confirm the stereo specific assignment of the H5'/H5" protons. There are three energetically favorable conformations for the γ angle: gauche$^+$, gauche$^-$, and trans. The gauche$^-$ conformation is usually never found due to steric hindrance. In the gauche$^+$ and trans conformers of C3'-endo and anti residues, the intra-residue H8/H6-H5'/H5" distances varies significantly (24). These distances are fairly even and long in the gauche$^+$ conformer (3.7 – 4.5 Å), whereas the H8/H6 – H5" distance is much shorter (2.5 – 2.9 Å) than the H8/H6- H5' distance in the trans conformer. So if there was not a significant difference in the intensities of the H8/H6-H5' and H8/H6-H5" NOEs (measured at the C5' plane) then the γ angle for that residue was constrained as gauche$^+$. For SL5 all the residues except for G697-U700 were restrained as C3'-endo, so these were the only residues for which we could not perform this comparative analysis. All these residues were found to have their gamma angles in the gauche$^+$ conformation and were restrained to a value of 60 +/- 15º. As for stereo-specific assignment of the H5'/H5" resonances, if a residue was found to have its γ angle in the trans conformation then, the peak with higher intensity would be assigned to the H5" resonance. However, since this did not happen the H5'H5" were assigned according to Shugar and Remin's rule which states that the H5' resonates downfield from the H5" (16).

B.4. Hydrogen bonds

A very important set of constraints for defining the secondary structure of RNA is distance constraints that define hydrogen-bonding interactions of base pairs. A complete account of the procedure used is given in Chapter 3 and will not be discussed here.

Table 2.1. ^{1}H, ^{13}C, ^{15}N composition of SL5.

17 residues	8 Purines	3 Adenines
		5 Guanines
	9 Pyrimidines	5 Cytosines
		4 Uracils
^{1}H Total = 199	Non-exchangeable Total = 131	Aromatic: H8 (5G, 3A); H2 (3A); H6 (5C, 4U); H5 (5C, 4U) Total = 29
		Sugar: H1'/H2'/H3'/H4'/H5'/H5" (17) Total = 102
	Exchangeable Total = 68	Imino: H1 (5G); H3 (3U) Total = 8
		Amino: NH2 (3A, 5G, 5C) Total = 26
		Sugar: OH (17 * 2) Total = 34
^{13}C Total = 161	Protonated Total = 114	Base: C8 (8 Pur); C2 (3A); C5 (9 Pyr); C6 (9 Pyr) Total = 29
		Sugar: C1'/C2'/C3'/C4'/C5' (17) Total = 85
	Non-protonated Total = 47	Purines: C4 (8); C5 (8); C6 (8); C2 (5G) Total = 29
		Pyrimidines: C2 (9); C4 (9) Total = 18
^{15}N Total = 63	Protonated Total = 22	Purines: N1 (5G); N2 (5G); N6 (3A) Total = 13
		Pyrimidines: N3 (4U); N4 (5C) Total = 9
	Non-protonated Total = 41	Purines: N9 (8 Pur); N7 (8 Pur); N3 (8 Pur); N1 (3A) Total = 27
		Pyrimidines: N3 (5C); N1 (9 Pyr) Total = 14
Total Number of ^{1}H, ^{13}C, ^{15}N = 423		

Table 2.2. Chemical shift ranges for ^{1}H, ^{13}C, and ^{15}N nuclei in RNA (19,26).

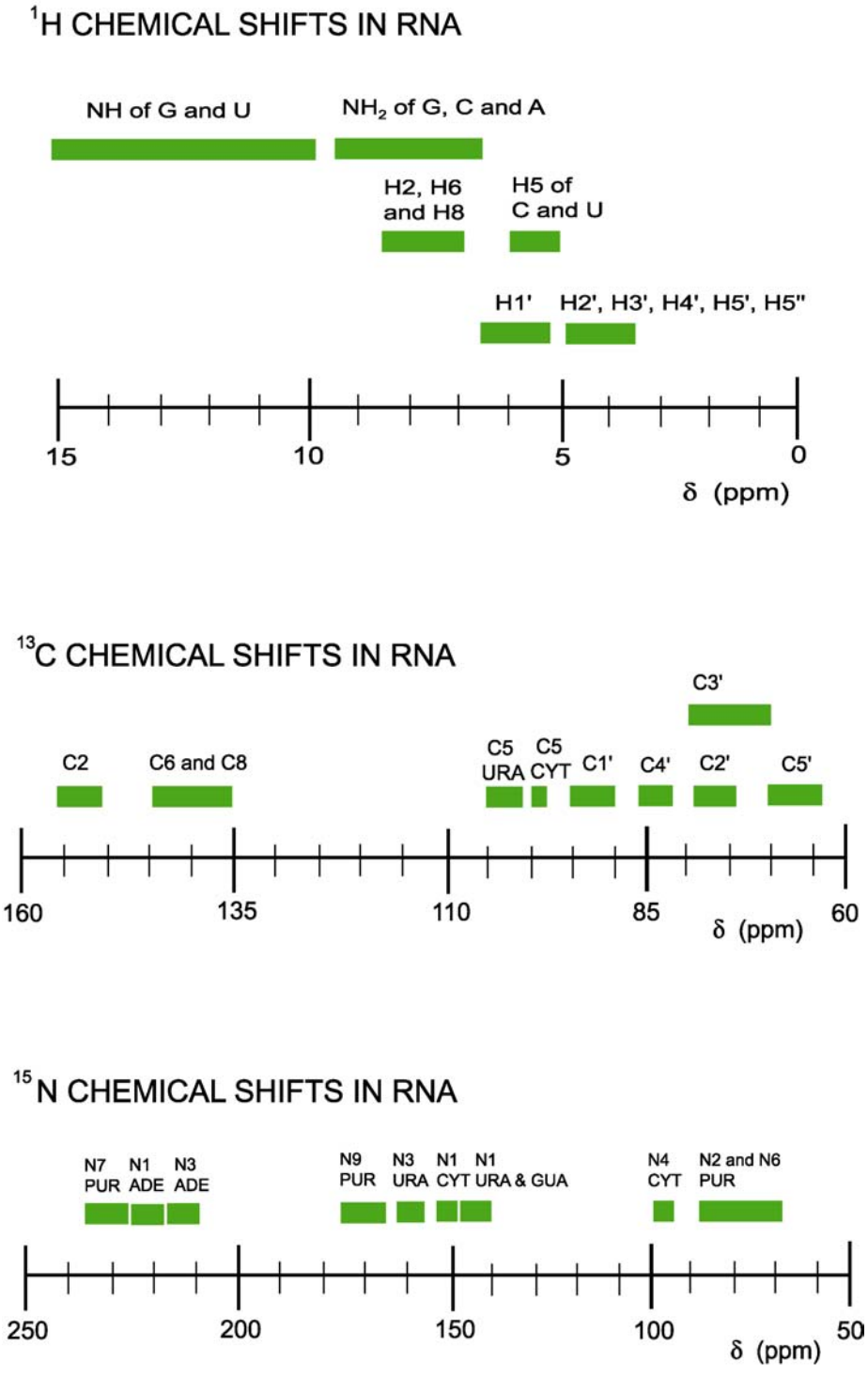


Table 2.3. Calibration of the NOE intensity (***I***) in various classes.

NOE Intensity (***I***) / units	Distance Limits / Å	Restraint Class
$I > 0.8^{A}$	1.8-3.0	VERY STRONG NOEs
$0.8 > I > 0.2^{A}$	1.8-4.1	STRONG NOEs
$0.2 > I^{B}$	1.8-5.5	MEDIUM NOEs
I^{C}	1.8-7.0	WEAK NOEs

[A] from the ^{13}C-edited HMQC-NOESY collected with a mixing time of 90 ms, [B] same as A or from ^{1}H-^{15}N CPMG-NOESY, [C] from the ^{13}C-edited HMQC-NOESY collected with a mixing time of 180 ms only or from the ^{1}H-^{15}N CPMG-NOESY and ^{1}H ^{1}H flip-back watergate NOESY.

Table 2.4. Results of the sugar pucker analysis for the 17 residues of SL5..

Residue	DQF-COSY [a]		HCCH E.COSY [b]			NOE [c]	RESULT [d]
	I (H1'/H2')	*Pucker* [e]	$^3J_{(H1'/H2')}$	$^3J_{(H3'/H4')}$	*Pucker*	*Pucker*	*Pucker*
G690	Weak	N	1.1	3.6	N	N	N
C691	Weak	N	0.9	8.2	N	N	N
G692	Weak	N	1.1	7.7	N	U	N
A693	Weak	N	1.6	9.3	N	N	N
G694	Weak	N	1.4	U	N	N	N
U695	Medium	N/S	3.1	U	N	N	N
U696	Medium	N/S	1.3	3	N	N	N
G697	Strong	S	1.4	1.6	N/S	S	N/S
A698	Strong	S	4.1	1.8	N/S	S	N/S
C699	Strong	S	4.5	0.9	N/S	S	N/S
U700	Strong	S	3.2	1.3	N/S	N	N/S
A701	Medium	N/S	1.9	U	N	N	N
C702	Weak	N	2.3	U	N	N	N
U703	Weak	N	1.6	7.2	N	N	N
C704	Weak	N	2.7	U	N	N	N
G705	Weak	N	0.9	7.2	N	U	N
C706	Medium	N/S	0.8	4.3	N	N	N

[a] Intensity (*I*) classification based on size of H1'-H2' cross peaks relative to H5-H6 cross peaks. [b] Coupling constants were measured in the indirect dimension due to phasing problems in the direct dimension. [c] Assumes no conformational averaging. [d] Result of combining the conclusions from all the analyses. [e] N = predominantly C3'-endo, S = predominantly C2'-endo, N/S = equilibrium of N and S, U = undetermined from analysis

Figure 2.1. The SL5 RNA used for the NMR structural studies of stem-loop V. The upper case letters represent wild-type nucleotides. The lower-case letters represent substitutions used to improve synthesis of the RNA by T7 RNA polymerase. The solid lines indicate the six Watson-Crick base pairs present in the proposed secondary structure.

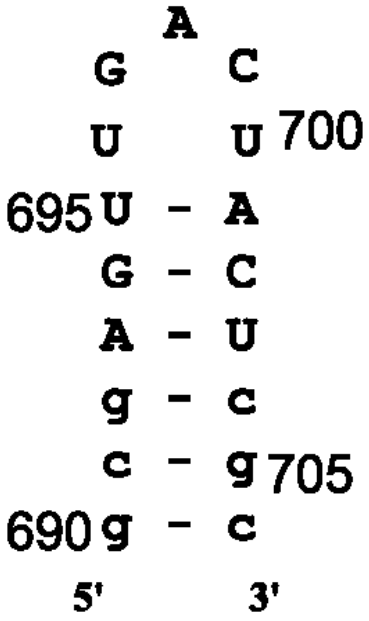


Figure 2.2. Illustration of the four bases and the sugar moiety typically found in RNA.

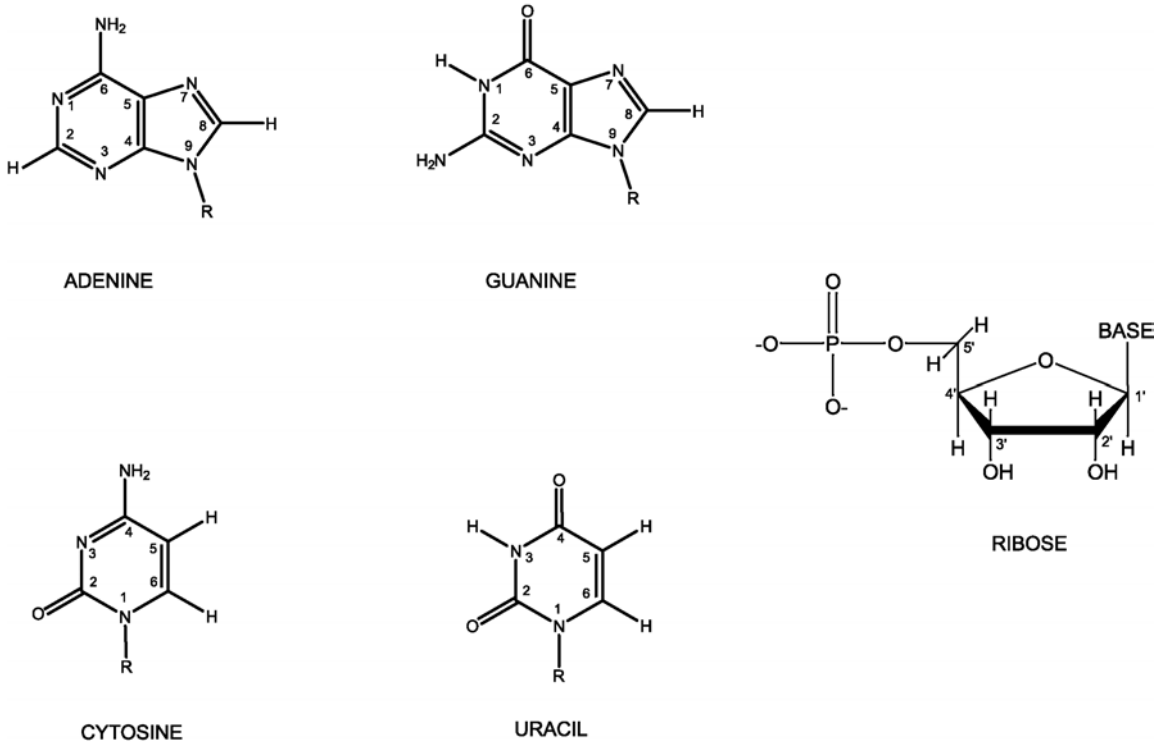


Figure 2.3. 1D watergate ^{1}H spectrum of SL5. The experiment was collected at 25 ºC in 10 mM d_{11}-Tris pH 7.0, 50 mM NaCl, 0.2 mM EDTA, 0.05 mM NaN_3, 90% H_2O:10% D_2O. Unless noted otherwise all spectra shown in this chapter were obtained at 25 ºC and at pH 7.0 under the same buffer conditions.

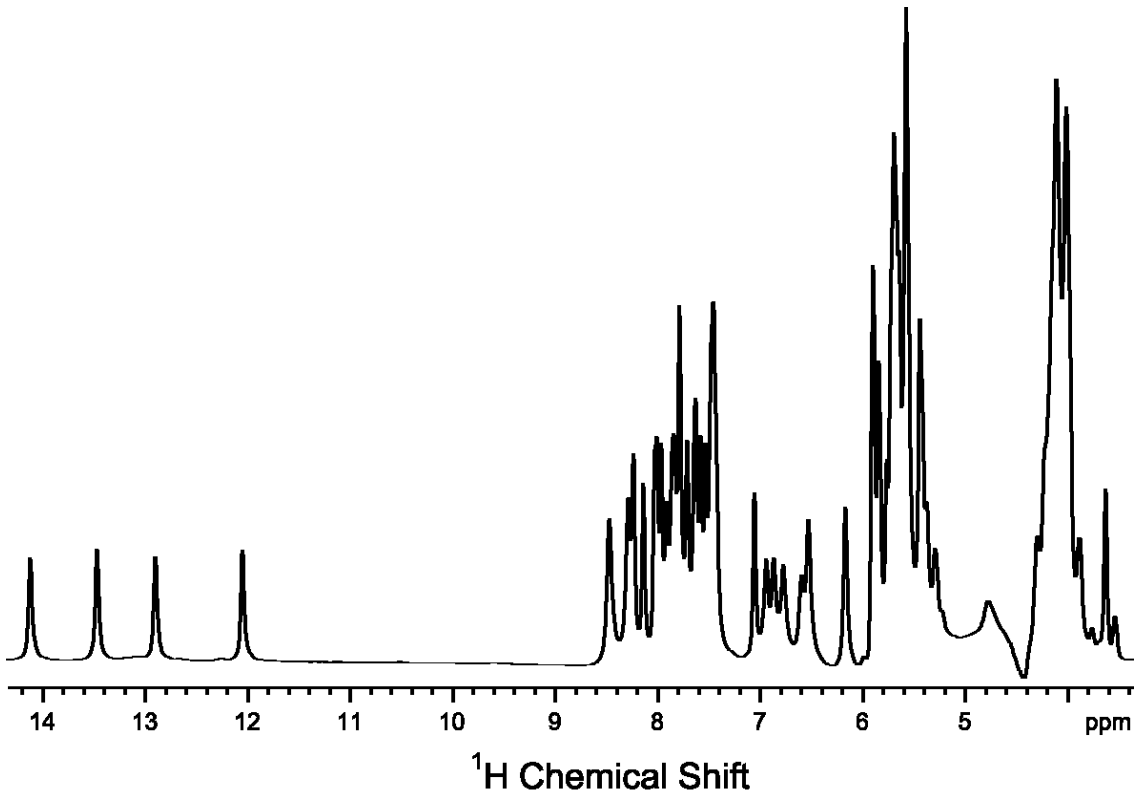


Figure 2.4. Watson-Crick base pairs found in RNA. The imino and amino protons are shown in red.

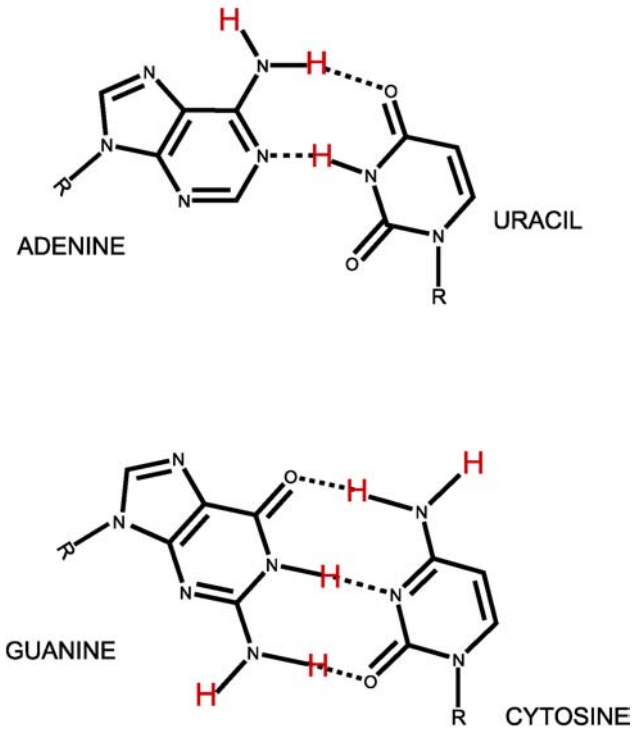


Figure 2.5. The imino proton region of 1D watergate ^{1}H spectrum of SL5 in H_2O. Assignments were made as described in text.

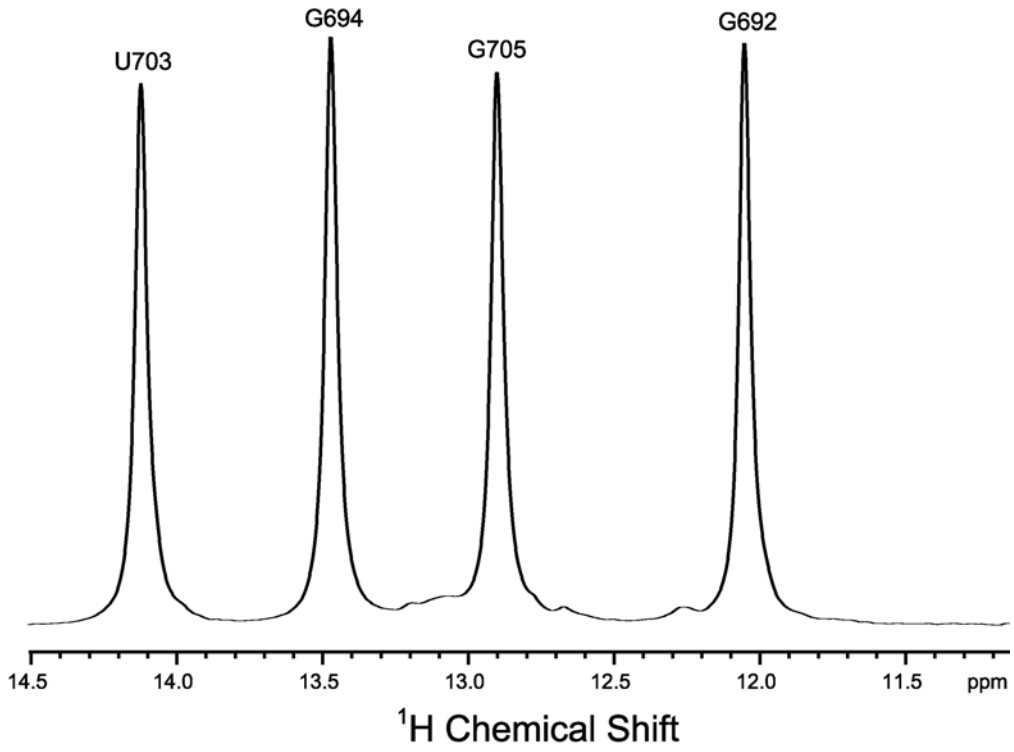


Figure 2.6. A 2D ^{1}H-^{1}H flip-back watergate NOESY (4,5) spectrum of SL5. The experiment was collected with a mixing time of 400 ms, showing the assignments of the imino proton region. The connectivities between the observed imino protons in the stem is shown.

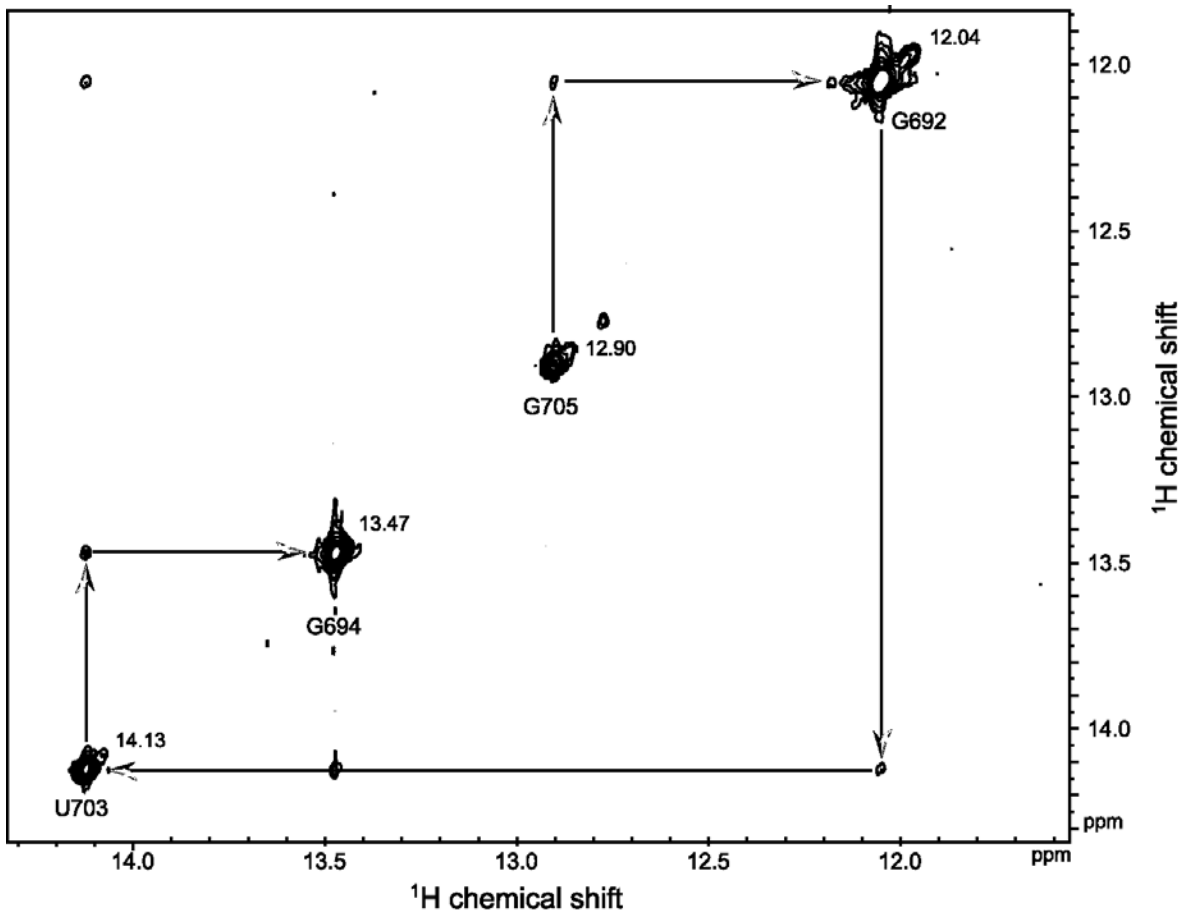


Figure 2.7. Imino-amino region of a 2D ^{1}H-^{1}H flip-back watergate NOESY (4,5) spectrum of SL5. The cross peaks between the imino protons of the guanosines and amino protons of the cytosines are shown. The cross peak between the imino proton of U703 and the H2 of A693 is also shown.

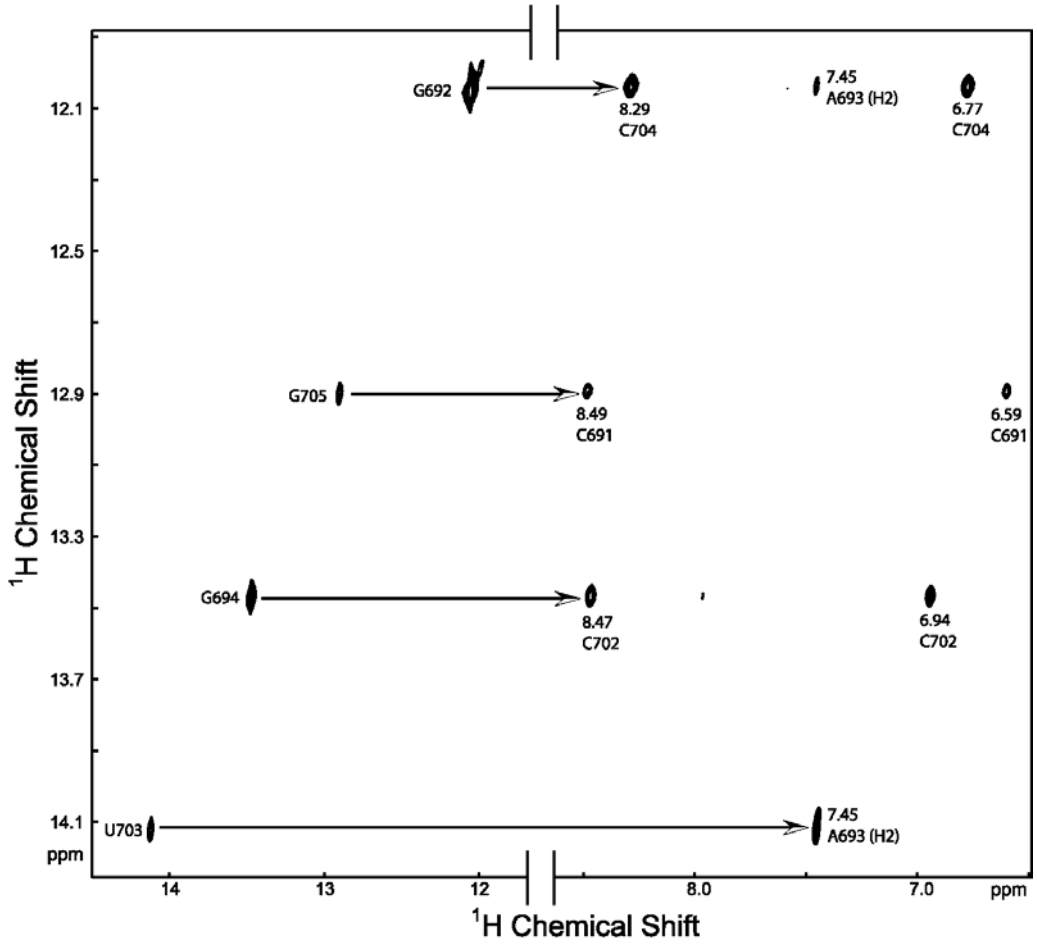


Figure 2.8. The cytidine-specific 2D H(NCCC)H (10) spectrum of SL5. This experiment correlates the amino protons of the cytosines to their H6 protons. In this case all the amino resonances of the cytosines are observed.

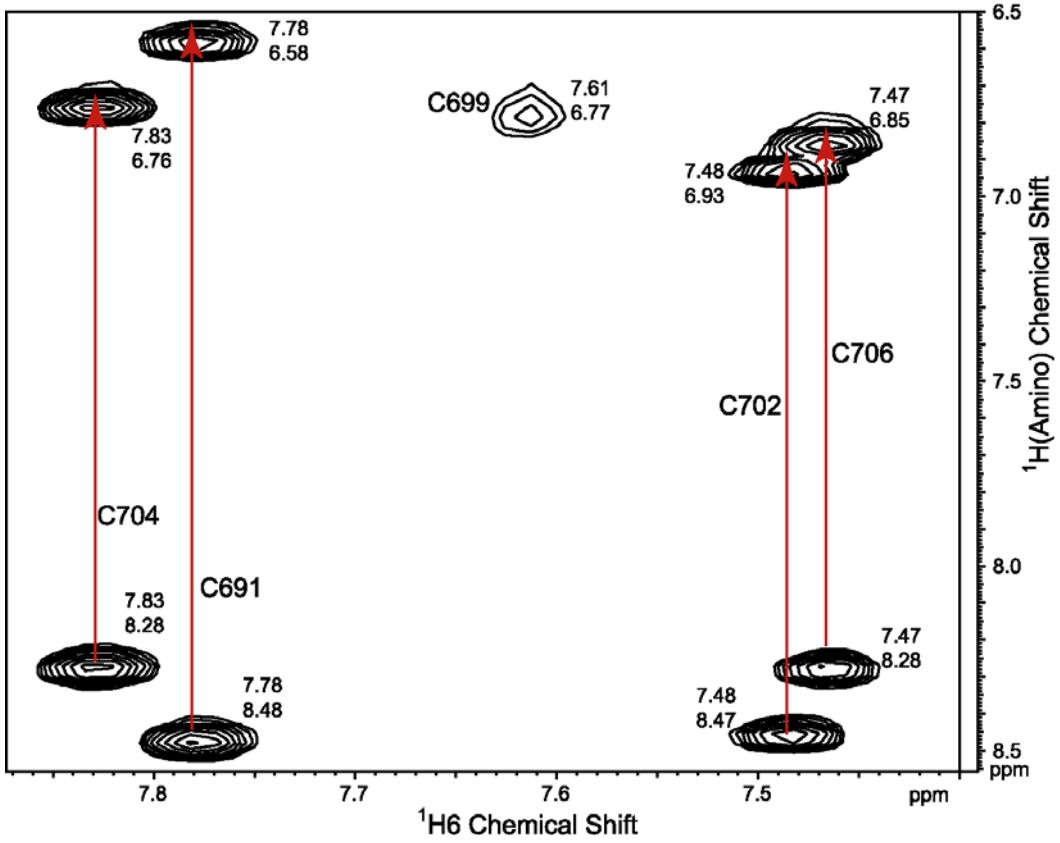


Figure 2.9. The purine nitrogen amino region of a 2D ^{1}H-^{15}N CPMG-NOESY (9) spectrum of SL5. The experiment was collected at 5 ºC. The amino protons of two of the five guanosines are observed. The amino protons of two of the three adenines are assigned.

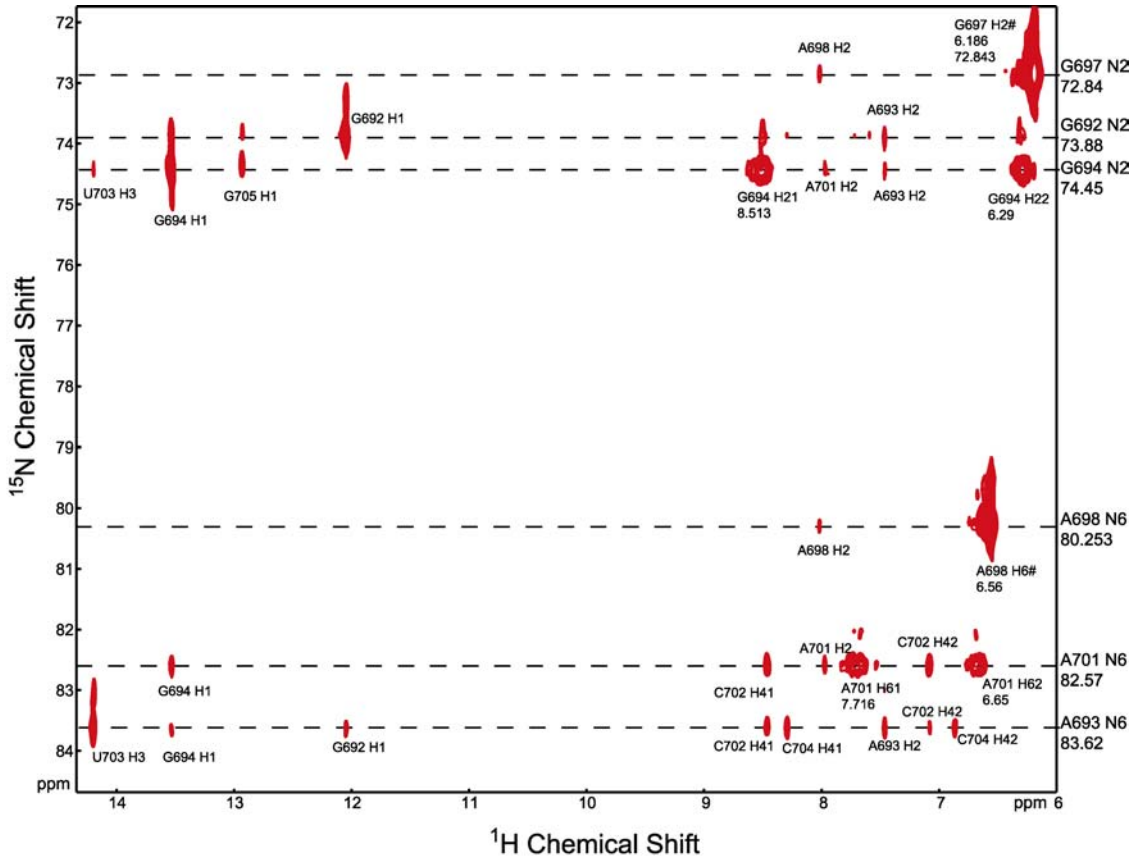


Figure 2.10. The guanosine-specific 2D H(NC)-TOCSY-(C)H (11) spectrum of SL5. The spectrum shows the correlation between the imino protons and the H8 protons. For the 2D H(NC)-TOCSY-(C)H experiment, correlations are only seen if the imino protons are observable, so only the guanosine imino protons observed in the 2D ^{1}H-^{1}H flip-back watergate NOESY (Figure 2.6) are observed here.

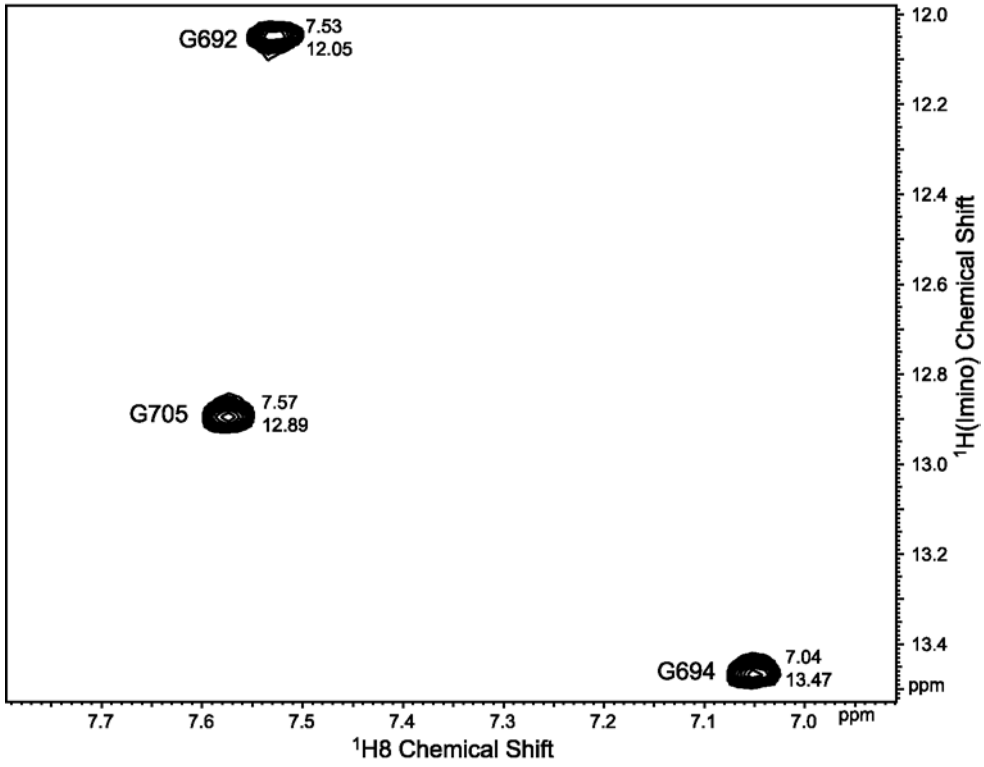


Figure 2.11. The uracil-specific 2D H(NCCC)H (10) spectrum of SL5. The spectrum shows the correlation between the imino proton and the H6 of U703. For the 2D H(NCCC)H experiment only observable imino protons give correlations. Hence, only the H6 of U703 can be assigned with this experiment.

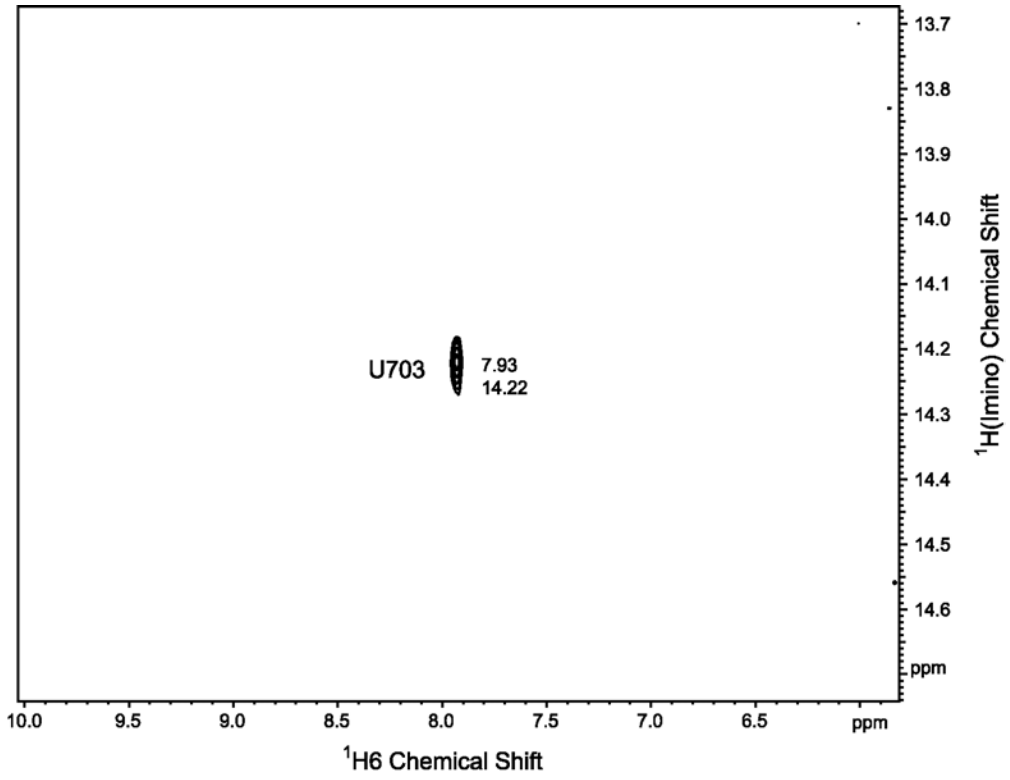


Figure 2.12. The adenine-specific 2D HNC-TOCSY-CH (12) spectrum of SL5. The spectrum shows the correlation between the H2 and H8 with the N6. The N6 and H2 assignments were made from the 2D ^{1}H-^{15}N CPMG-NOESY spectrum (Figure 2.9). The H8 resonances are labeled in red and the H2 resonances are labeled in black. The difference in the resonances between the 2D HNC-TOCSY-CH and the 2D ^{1}H-^{15}N CPMG-NOESY is due to the fact that the 2D HNC-TOCSY-CH was collected at 25 ºC, whereas the 2D ^{1}H-^{15}N CPMG-NOESY was collected at 5 ºC.

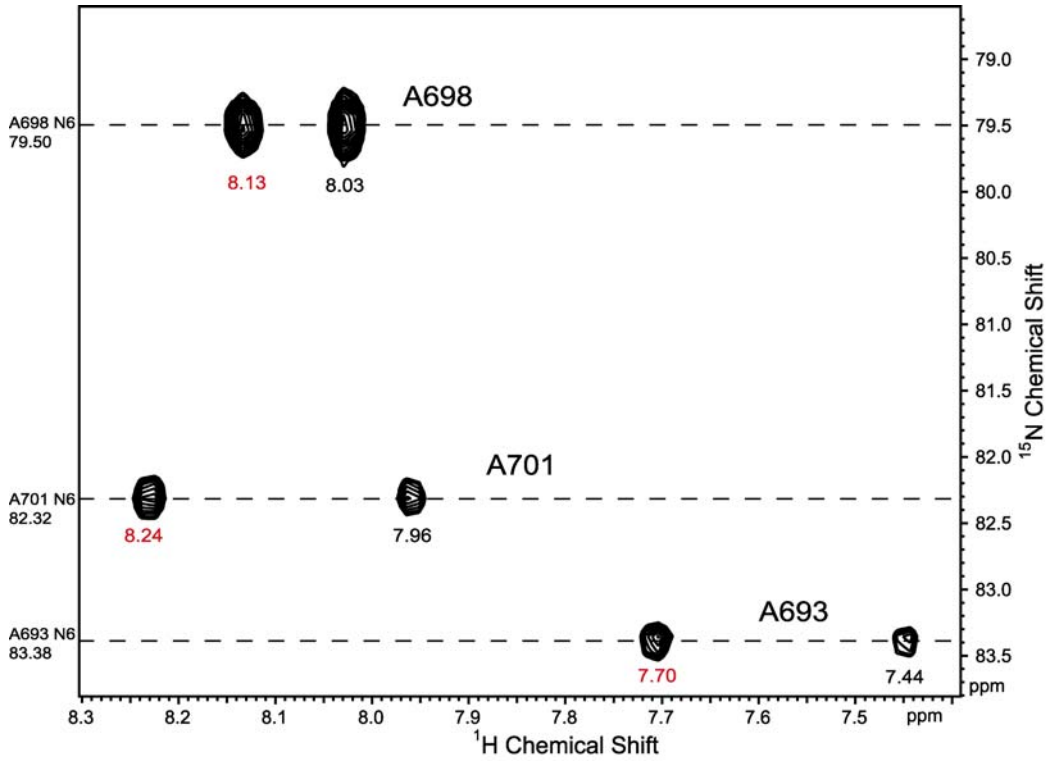
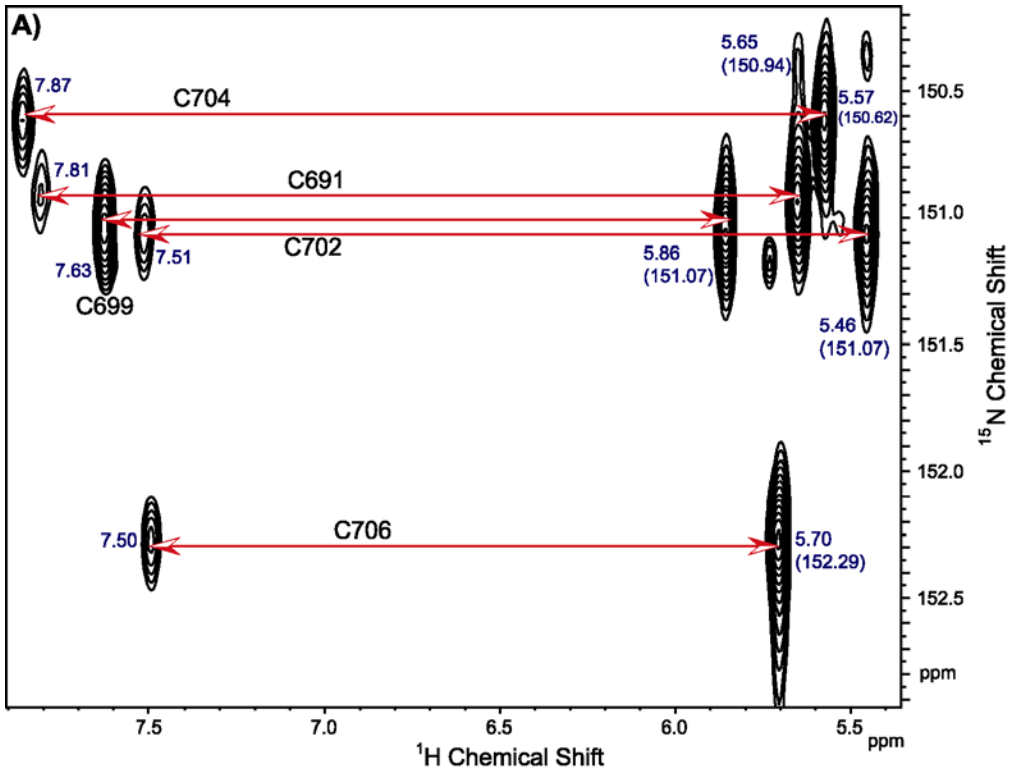
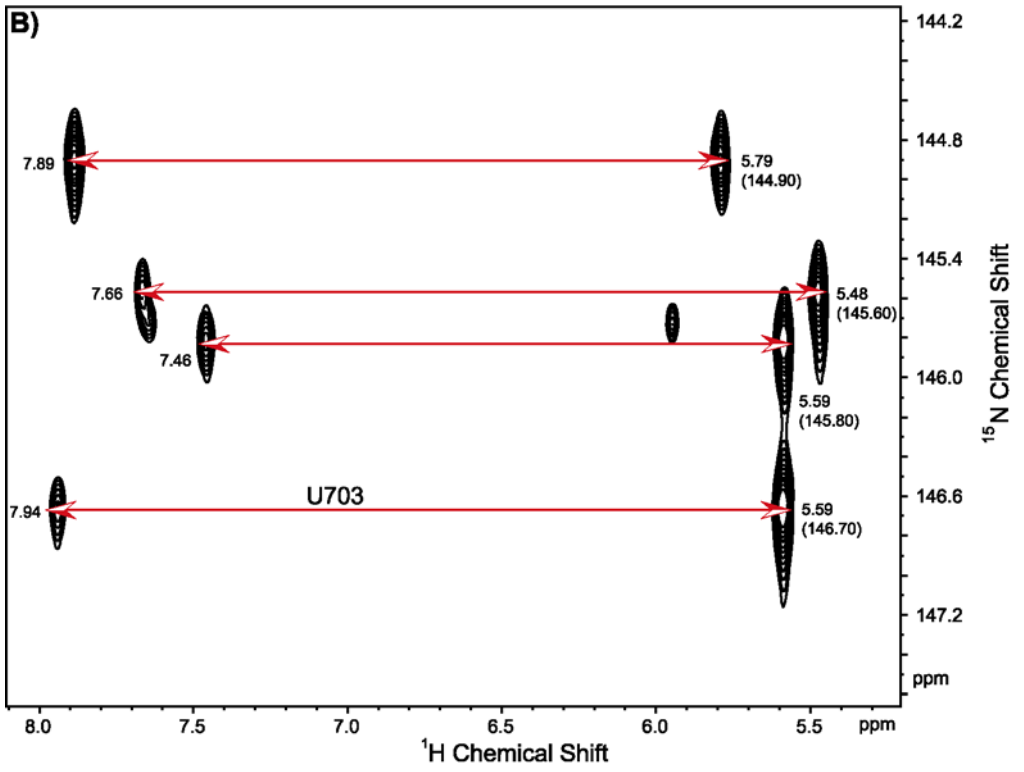


Figure 2.13. The 2D ^{1}H-^{15}N MQ-(HC)N(C)H (13) spectrum of SL5. The spectrum shows the correlation between the H1' and the intra-nucleotide H6/H8 via the N1/N9 nuclei. The protons are at the frequency of the N1/N9 nitrogen. The H1' protons are upfield of the H6/H8 protons. The correlated protons are connected with arrows. The known resonances are assigned. The nitrogen resonances are given in parentheses beside the H1' proton resonance. **A)** The cytosine region of the 2D ^{1}H-^{15}N MQ-(HC)N(C)H. **B)** The uracil region of the 2D ^{1}H-^{15}N MQ-(HC)N(C)H. **C)** The purine region of the 2D ^{1}H-^{15}N MQ-(HC)N(C)H. From the 2D HNC-TOCSY-CH (Figure 2.12) it is known which peaks are adenosine H8s and they are highlighted in orange.




B)
7.89
5.79
(144.90)
7.66
7.46
5.48
(145.60)
5.59
(145.80)
7.94
U703
5.59
(146.70)
144.2
144.8
145.4
146.0
146.6
147.2
ppm
15N Chemical Shift
8.0
7.5
7.0
6.5
6.0
5.5
ppm
1H Chemical Shift

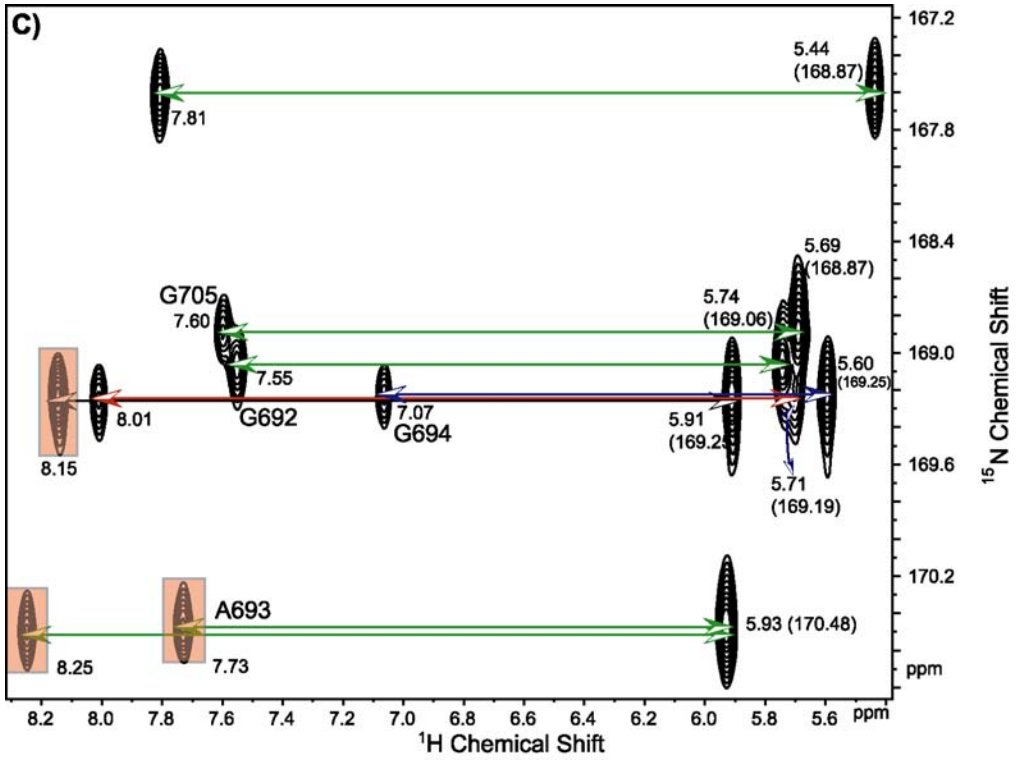

C)
5.44
(168.87)
7.81
5.69
(168.87)
G705
7.60
5.74
(169.06)
7.55
5.60
(169.25)
8.01
G692
7.07
G694
5.91
(169.25)
8.15
5.71
(169.19)
A693
5.93 (170.48)
8.25
7.73
167.2
167.8
168.4
169.0
169.6
170.2
ppm
15N Chemical Shift
8.2
8.0
7.8
7.6
7.4
7.2
7.0
6.8
6.6
6.4
6.2
6.0
5.8
5.6
ppm
1H Chemical Shift

Figure 2.14. H5/H6 region of the 2D DQF-COSY spectrum of SL5. The spectrum shows the H5/H6 correlations of all the pyrimidines. The cytosines and U703 cross peaks were assigned from the 2D ^{1}H-^{15}N MQ-(HC)N(C)H (Figure 2.13). The remaining cross peaks were assigned once the sequential walk was completed.

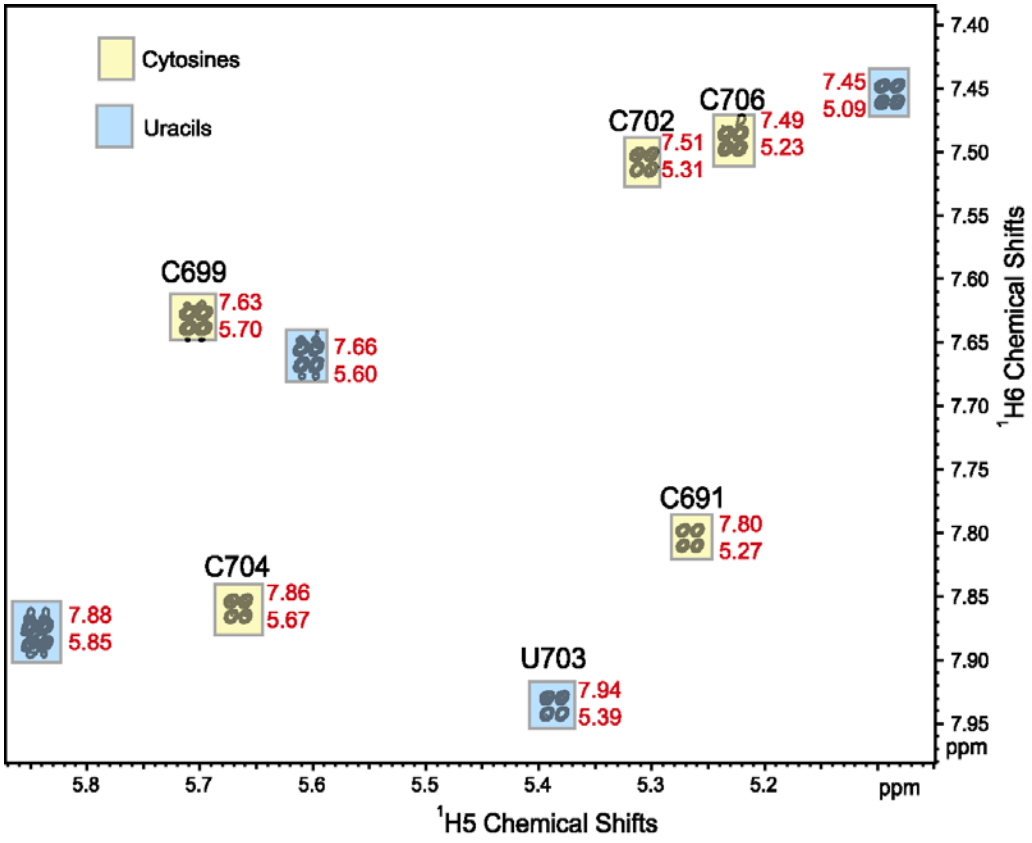


Figure 2.15. Diagram illustrating the H6/H8 –H1' sequential walk. The H1' proton has an intra-nucleotide correlation with its own H8/H6 (red arrows), as well as, with the H8/H6 proton of the nucleotide that is 3' sequential to it (blue arrow).

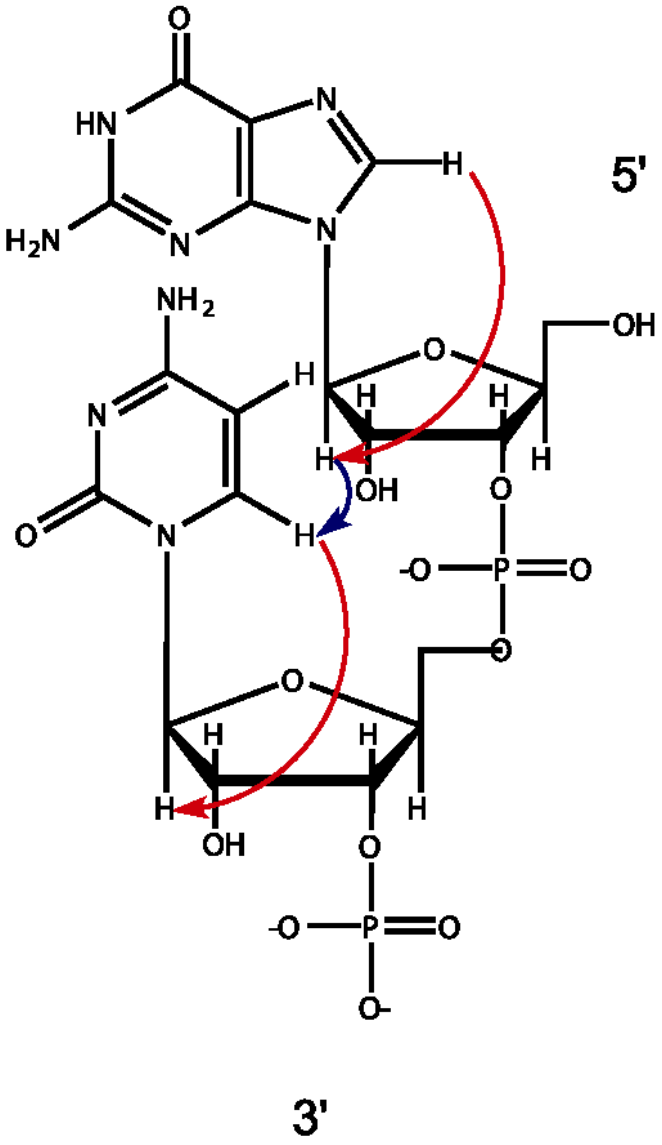


Figure 2.16. The SL5 sequential walk. Shown is the H6/H8-H1' region of the 2D ^{1}H-^{1}H NOESY spectrum of SL5. The experiment was collected in D_2O with a mixing time of 360 ms and shows the complete sequential walk.

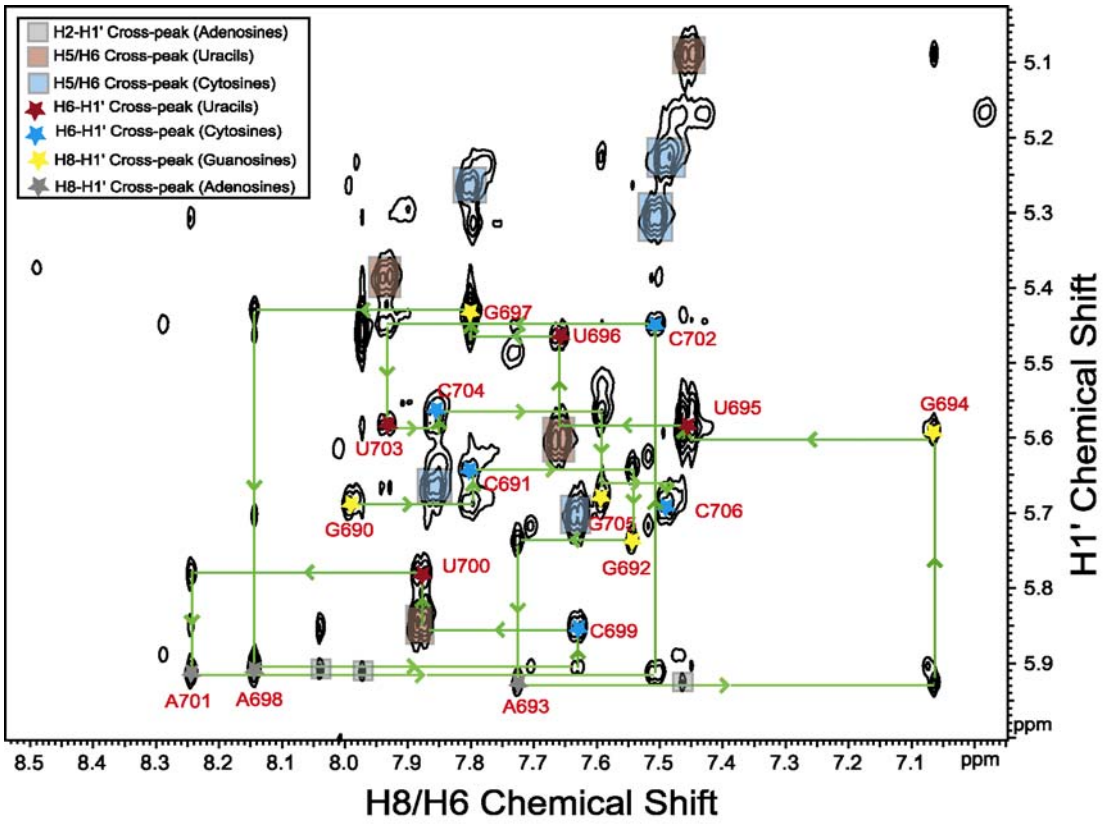
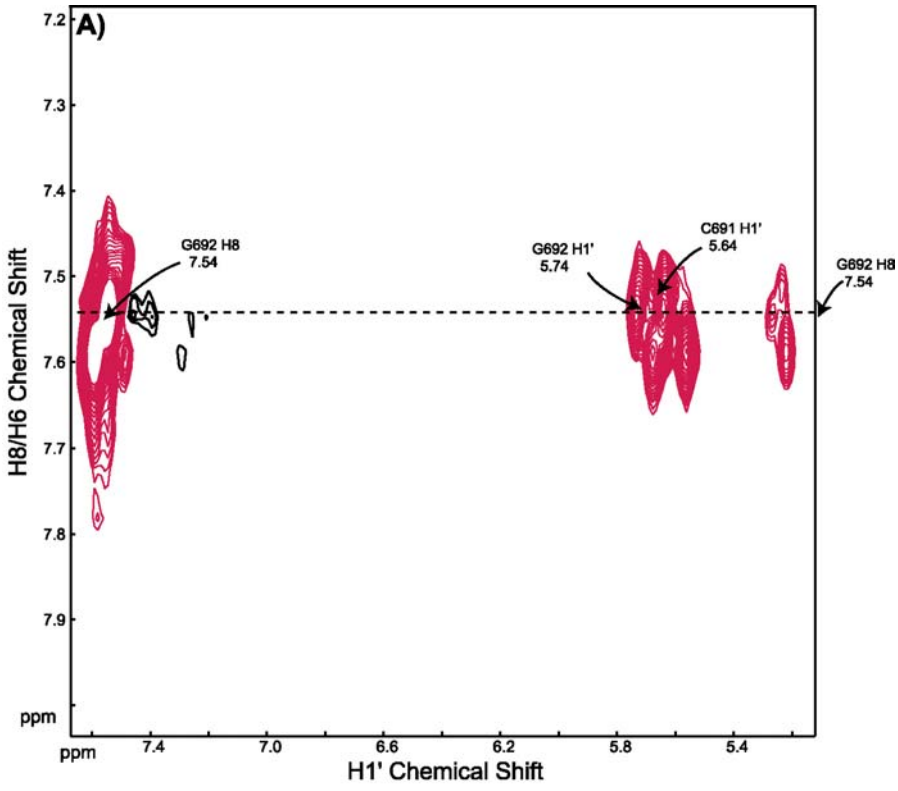


Figure 2.17. 3D sequential walk of SL5. Two planes from the 3D ^{13}C-edited HMQC-NOESY (14) spectrum with a mixing time of 240 ms showing how the 3D sequential walk is done for residue G692. **A)** At the plane that is the G692 C8 resonance (7.55 ppm). The cross peaks are the intra-residue cross peak to G692 H1' and the inter-residue cross peak to the H1' of the previous residue (in this case C691 H1'). **B)** At the plane that is the G692 C1' resonance (92.9 ppm). The cross peaks are intra-residue H1'-H8 and the inter-residue cross peak to the sequential residue's H8 (A693). The walk is then continued by going to the plane that represents the C8 resonance of A693.



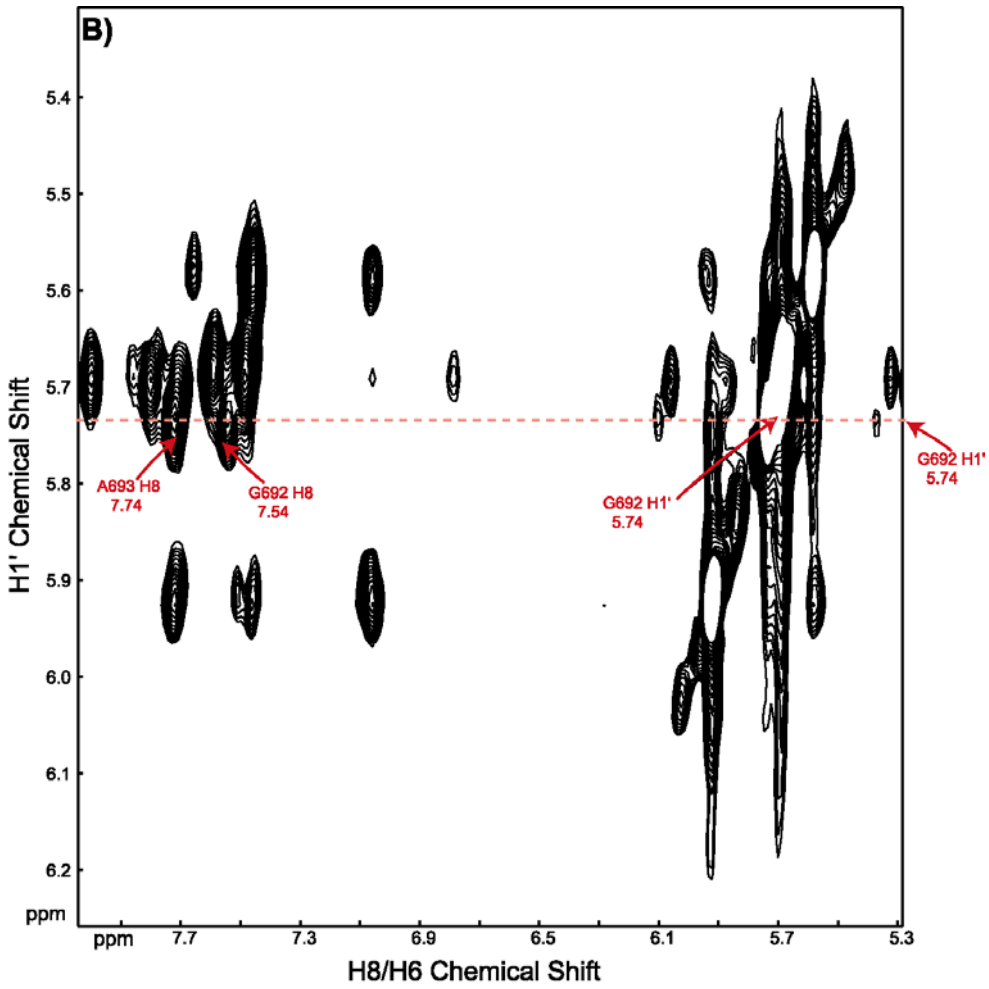

B)
H1' Chemical Shift
5.4
5.5
5.6
5.7
5.8
5.9
6.0
6.1
6.2
ppm
A693 H8
7.74
G692 H8
7.54
G692 H1'
5.74
G692 H1'
5.74
ppm
7.7
7.3
6.9
6.5
6.1
5.7
5.3
H8/H6 Chemical Shift

Figure 2.18. A 3D HCCH-COSY (15) spectrum of SL5. The 2D ^{1}H(F1) ^{1}H(F3) plane shown is at the C1' chemical shift of G697 in F2 (93.6 ppm). In this experiment the autocorrelation peak represents the resonance of the directly attached proton (in this case H1').

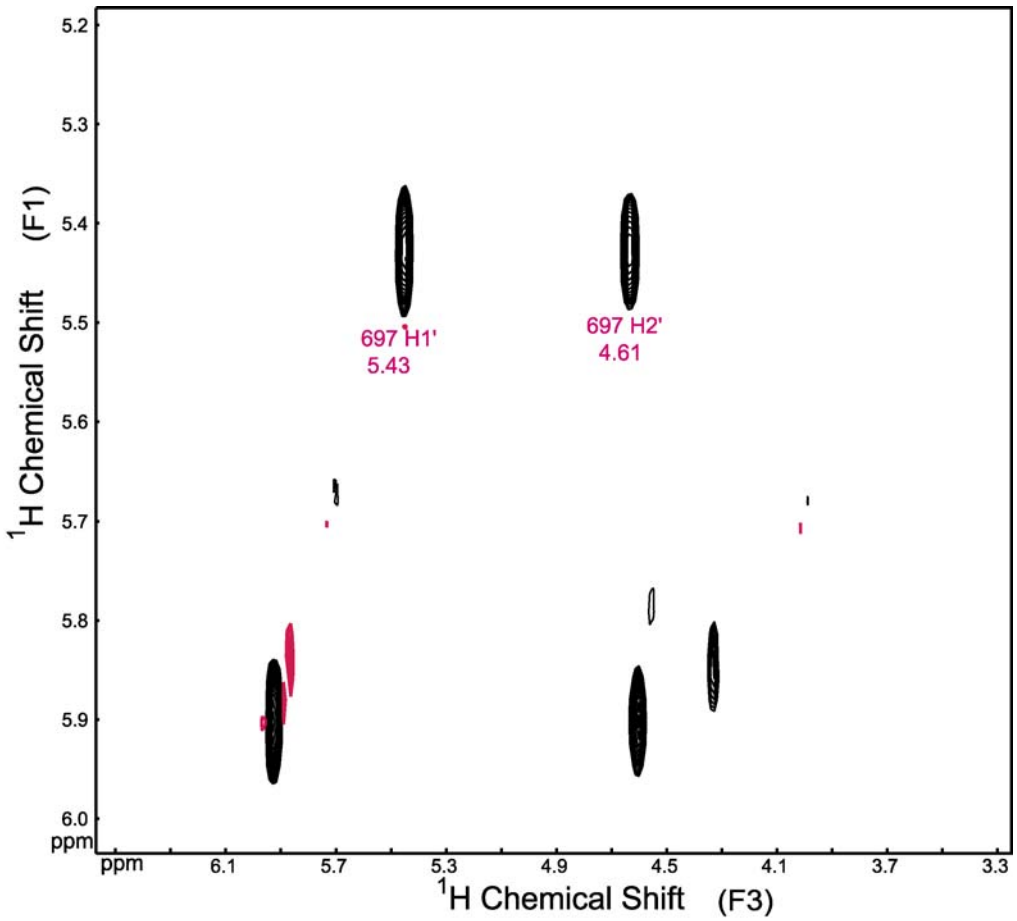
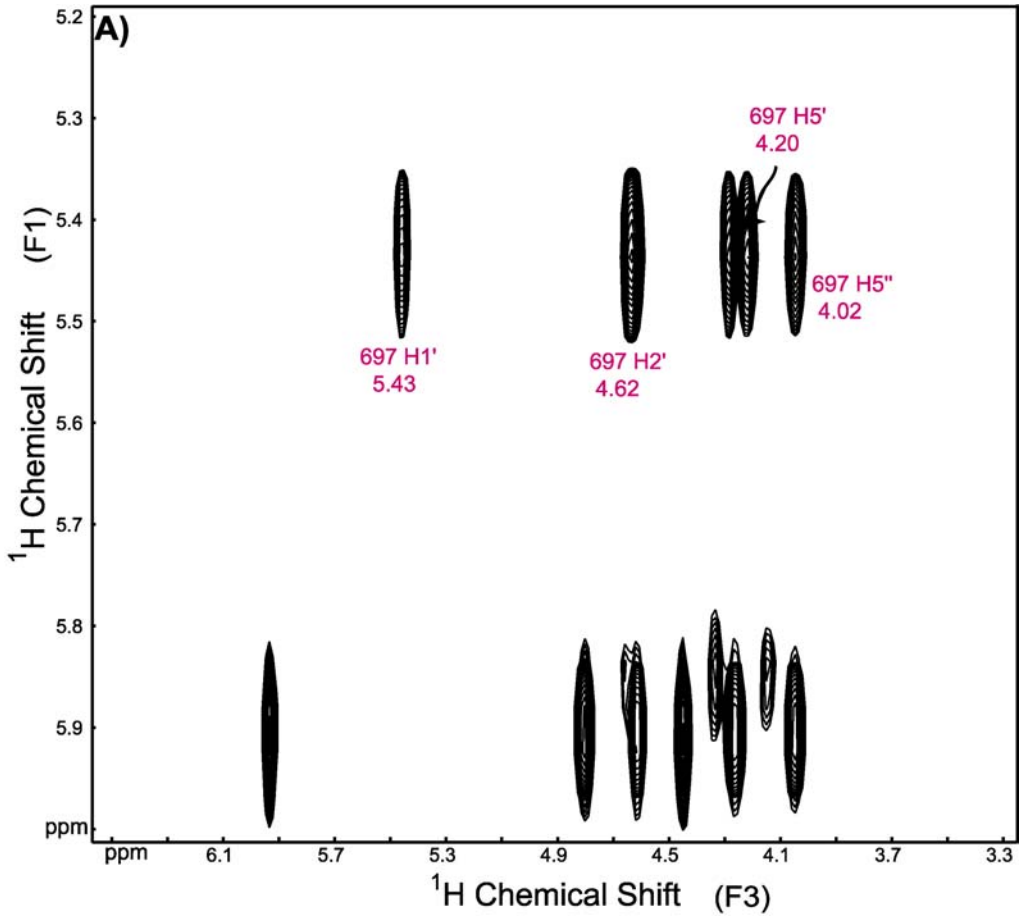


Figure 2.19. 3D HCCH-TOCSY (15) spectrum of SL5. The 2D ^{1}H(F1) ^{1}H(F3) planes shown are at the C1' chemical shift of G697 in F2 (93.6 ppm). The autocorrelation peak represents the resonance of the attached proton, H1'. The cross peaks to all the other ribose protons which are directly attached to the carbons that are J-coupled to C1' are shown. **A)** All the protons are in the same phase. **B)** The H5'/H5" protons are in anti-phase to the other protons.



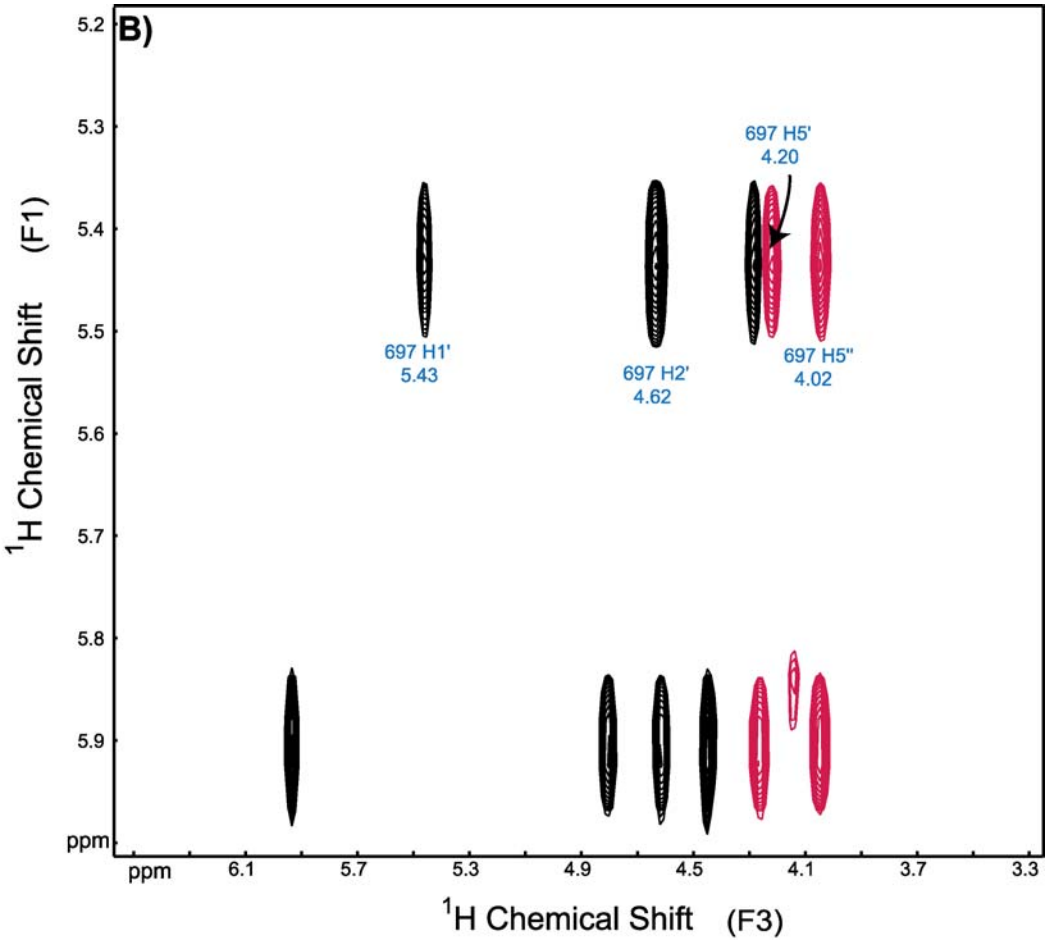

B)
697 H5'
4.20
697 H1'
5.43
697 H2'
4.62
697 H5"
4.02
5.2
5.3
5.4
5.5
5.6
5.7
5.8
5.9
ppm
^{1}H Chemical Shift (F1)
ppm
6.1
5.7
5.3
4.9
4.5
4.1
3.7
3.3
^{1}H Chemical Shift (F3)

Figure 2.20. 3D HCCH-TOCSY (15) spectrum of SL5. The 2D ^{1}H(F1) ^{13}C(F2) plane shown is at the H1' chemical shift of G697 in F3 (5.43 ppm).

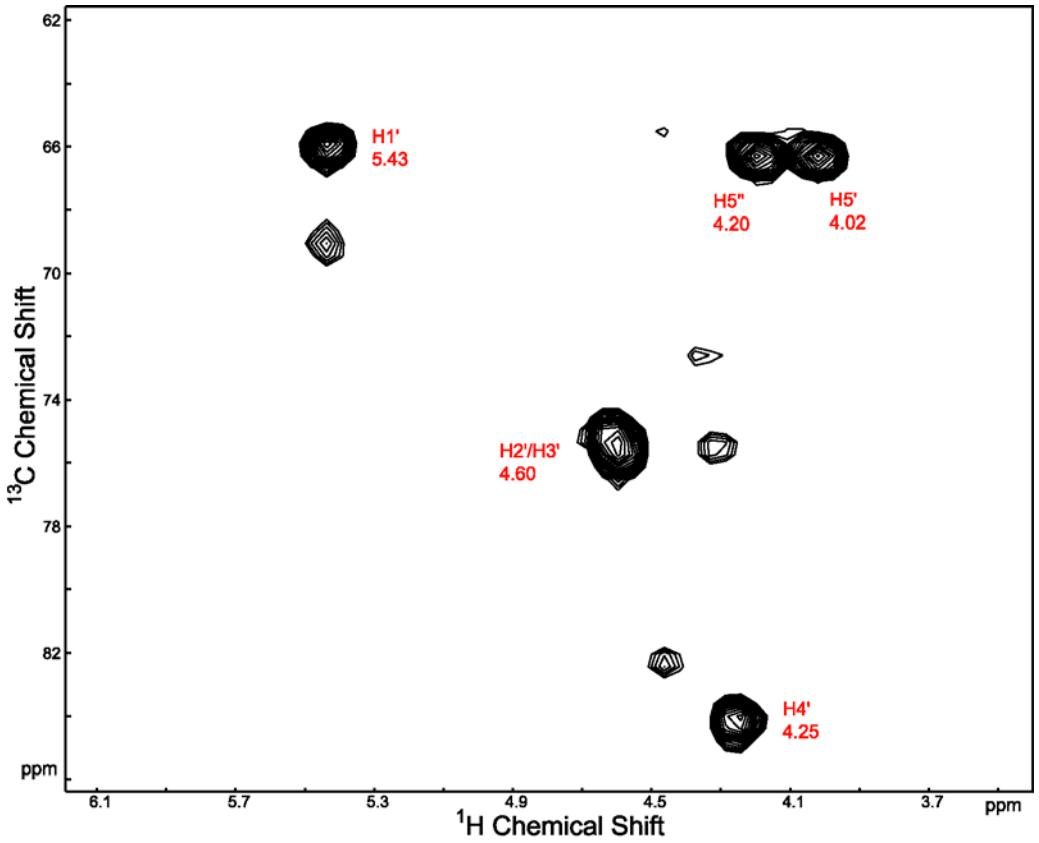


Figure 2.21. 2D ^{1}H-^{13}C CT-HSQC (17,18) spectrum of SL5. The constant time period was 9.5 ms. The spectrum shows the different groups of protonated carbons. The carbon (and proton) resonances are resolved according to Table 2.2.

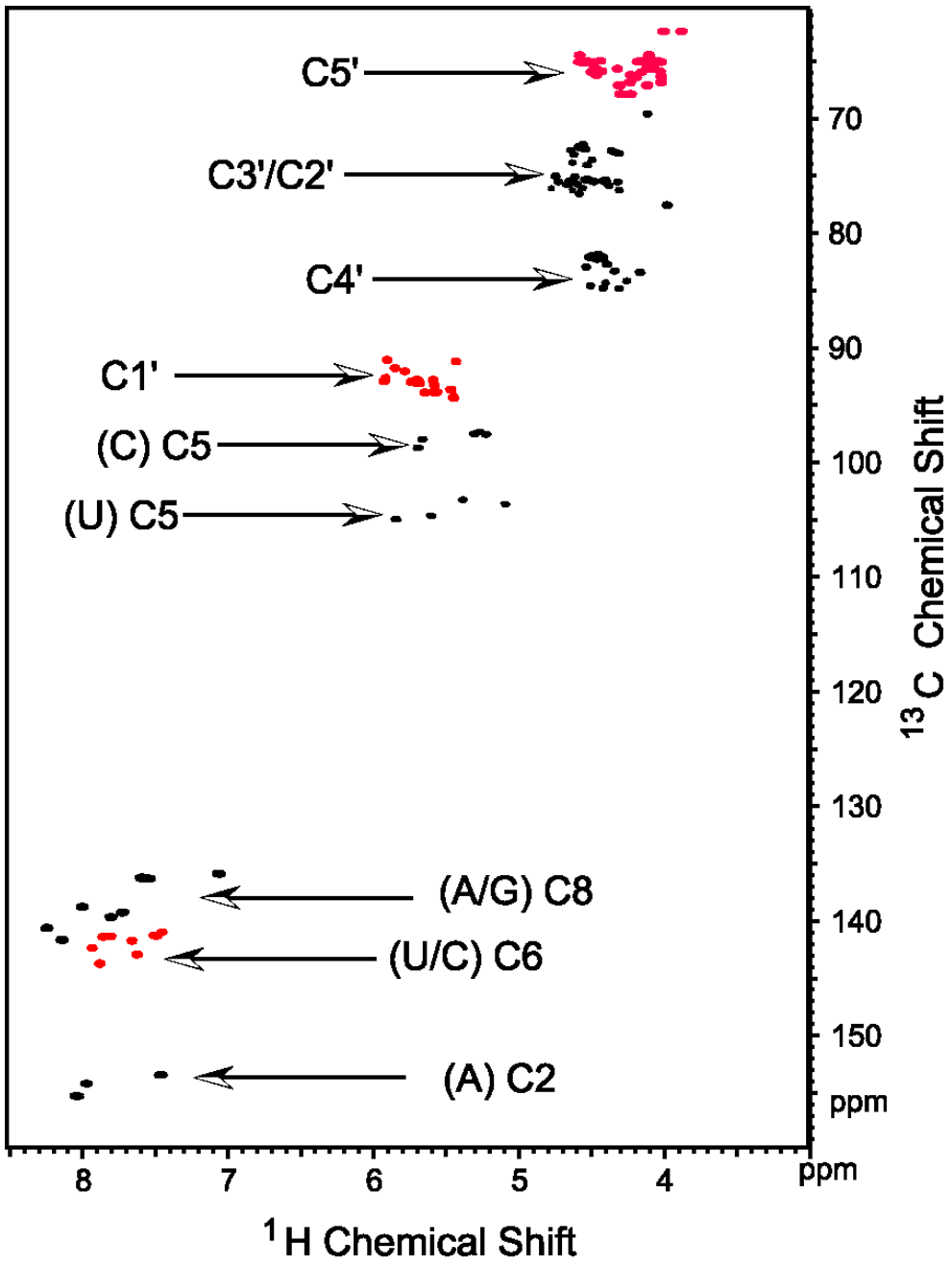


Figure 2.22. The H8/H6/H2-C8/C6/C2 region of the 2D ^{1}H-^{13}C CT-HSQC (17,18) spectrum of SL5. The complete ^{13}C resonance assignments of the peaks are shown. The H6-C6 peaks have negative phase and appear as red peaks in this figure. The three downfield carbon resonances are the C2 for the adenines.

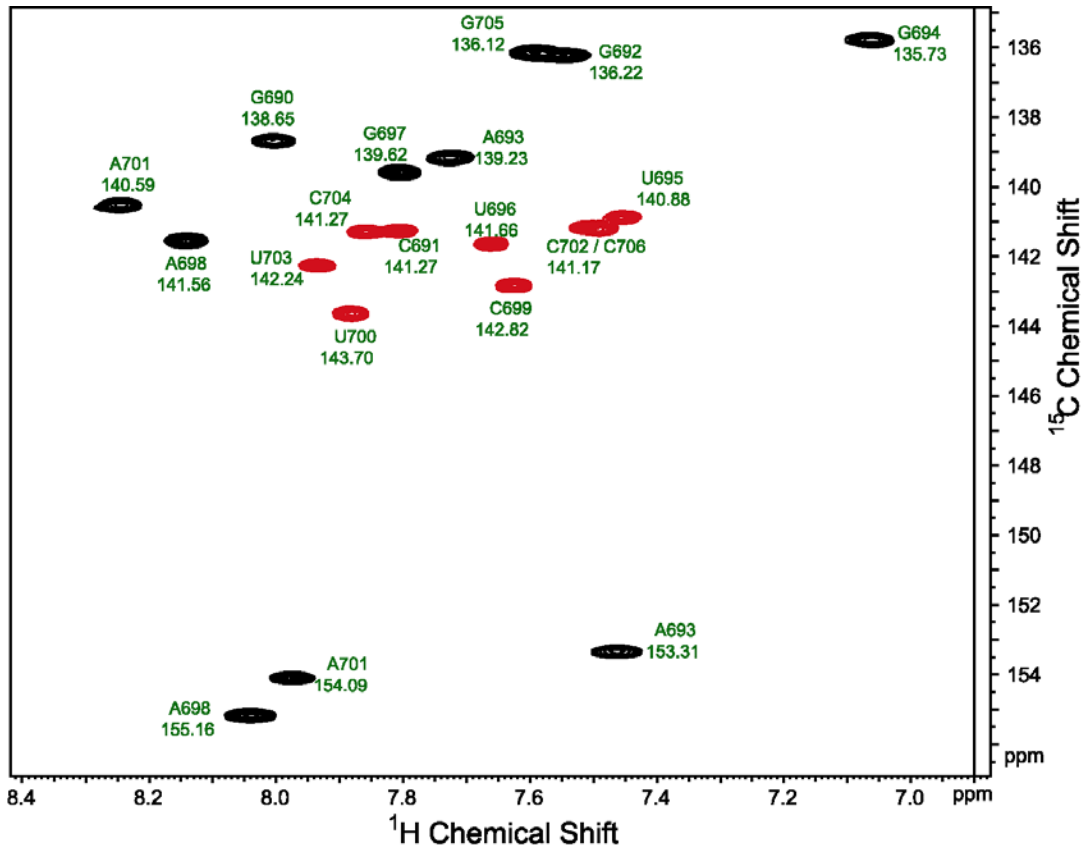


Figure 2.23. The H5-C5 region of the 2D ^{1}H-^{13}C CT-HSQC (17,18) spectrum of SL5. The spectrum shows the peak assignments. The uracil C5 carbons (downfield peaks in the ^{13}C dimension) are resolved from the cytosine C5 carbons (upfield peaks in the ^{13}C dimension).

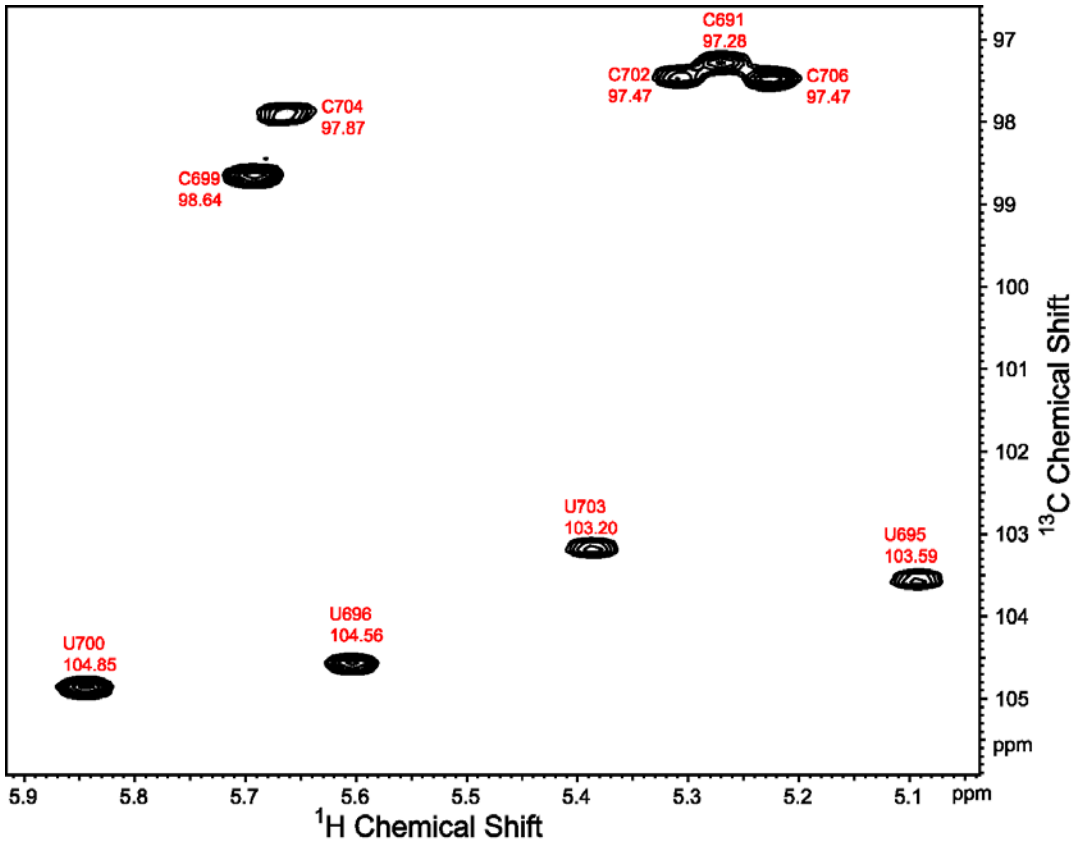


Figure 2.24. The H1'-C1' region of the ^{1}H, ^{13}C CT-HSQC (17,18) spectrum of SL5. The complete assignments are shown.

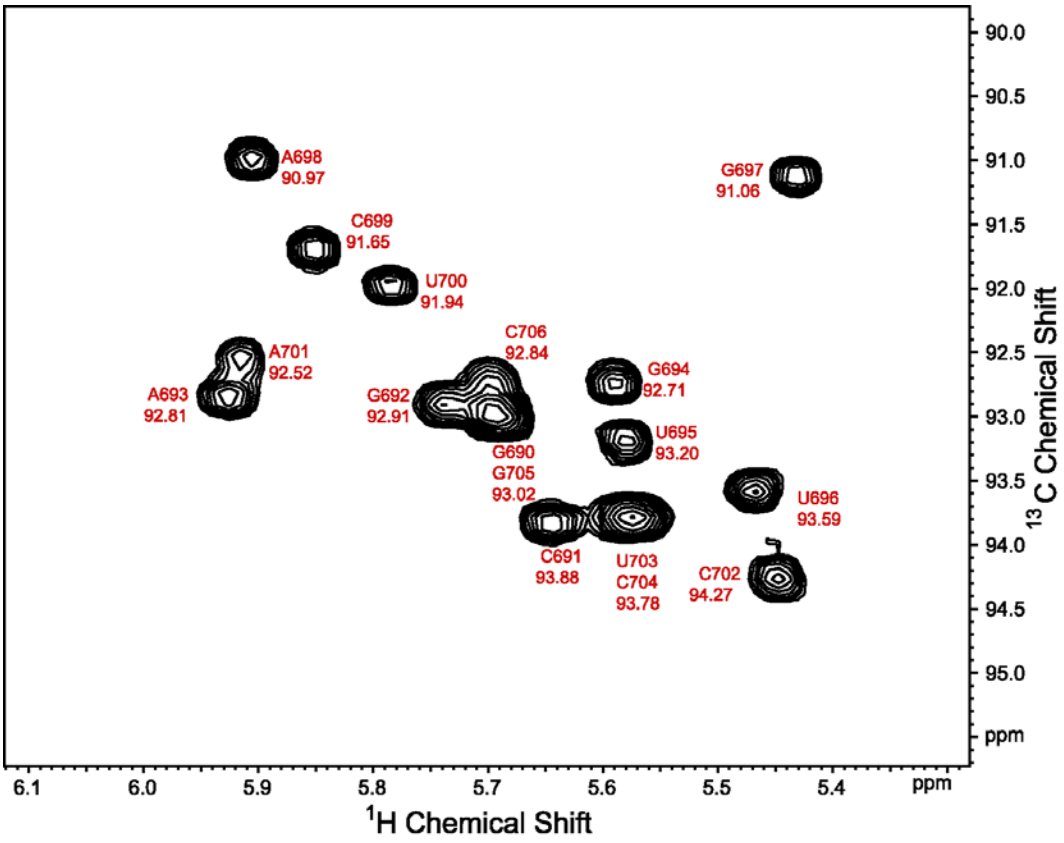


Figure 2.25. The imino proton region of a 2D ^{1}H-^{15}N HSQC (7) spectrum of SL5. The experiment was collected in H_2O and optimized for detection of the imino protons and their attached nitrogens. The nitrogen resonance assignments are shown.

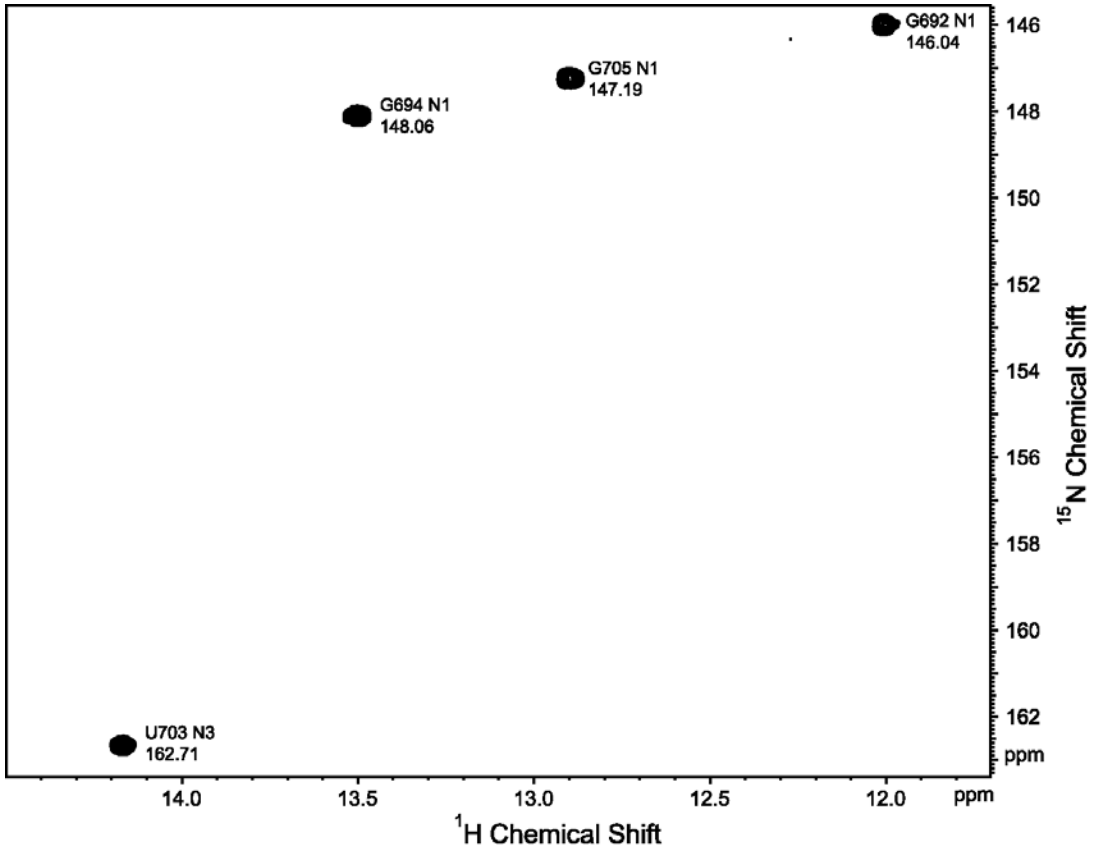


Figure 2.26. The amino proton region of a 2D ^{1}H-^{15}N HSQC (7) spectrum of SL5. The experiment was collected in H_2O and was optimized for detection of the amino protons. The only amino protons seen in this region of the spectrum were the cytosine amino protons. The nitrogen resonances are given, the amino proton resonances are in parentheses.

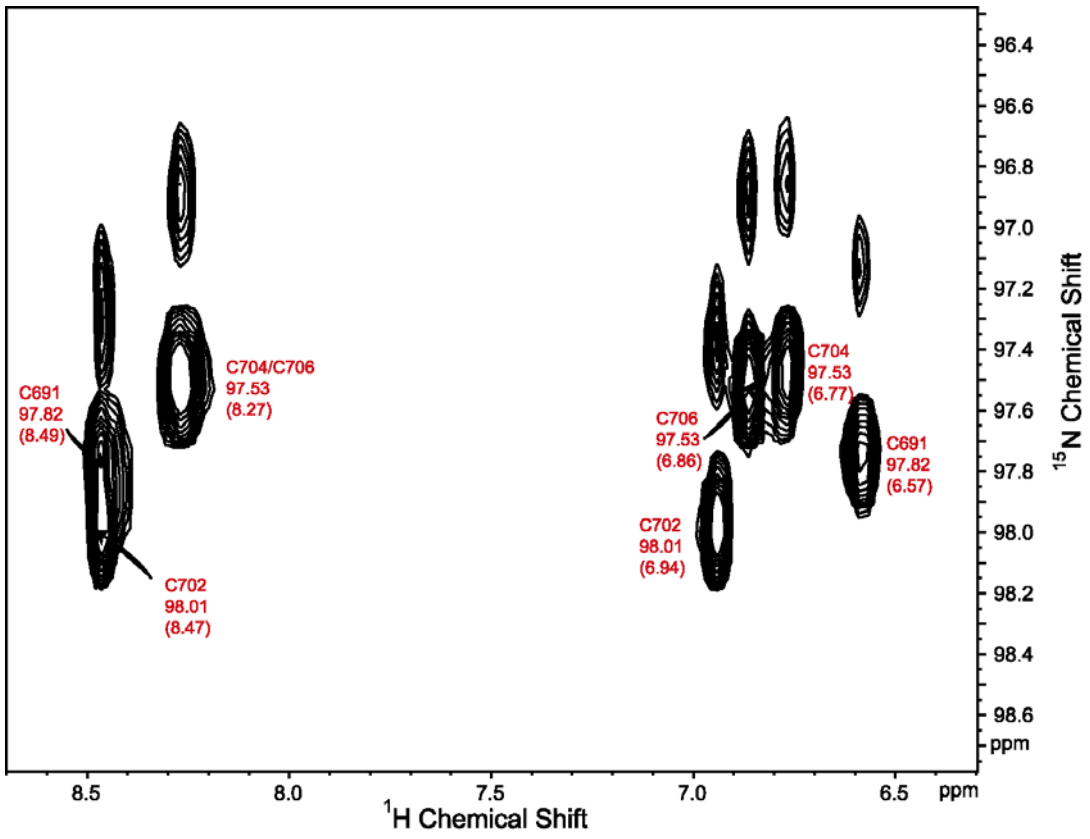


Figure 2.27. 2D ^{1}H-^{15}N HMQC (19) spectrum of SL5. The experiment was collected in D_2O and optimized for J = 7 Hz (19). The correlations between the uridine H5 and N1 and N3, as well as the correlation between cytosine H5 and N1 are shown. From Table 2.2 it is known that the N3s resonate downfield of the N1s, hence the lower four peaks are the uracil N3s.

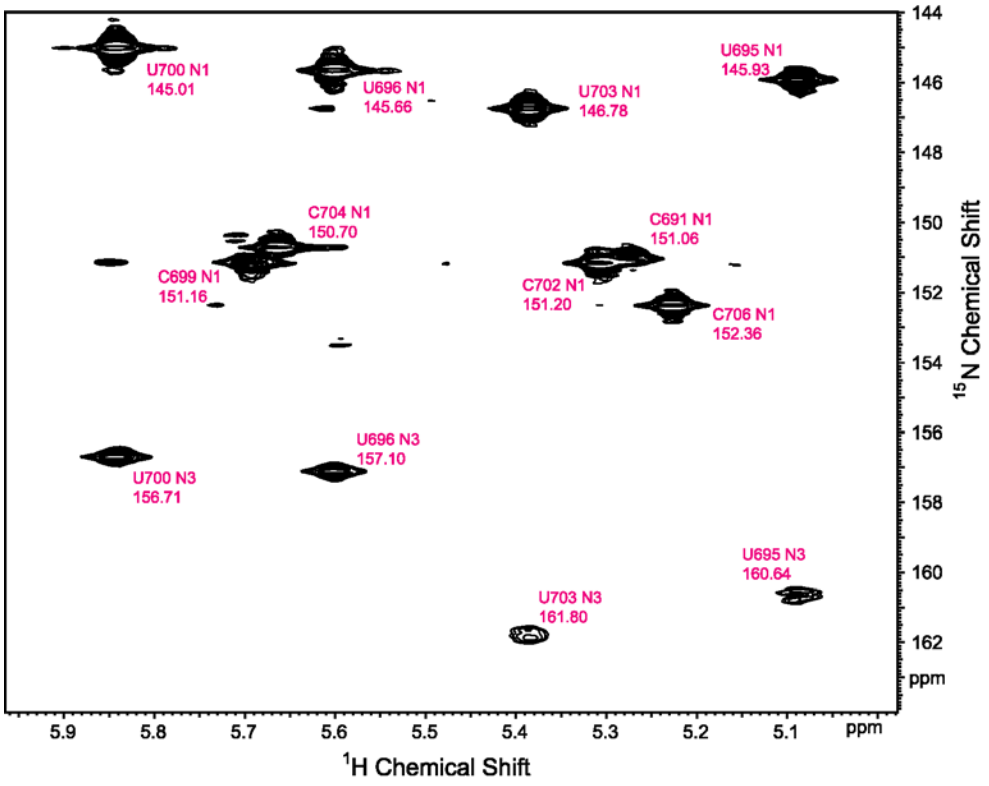


Figure 2.28. Adenine N1/N3 region of a modified HNN-COSY (20) spectrum of SL5 collected in D_2O. The spectrum shows the correlation between the adenine H2 signals and the coupled N1 and N3 signals. The N3 signals resonate upfield of the N1s.

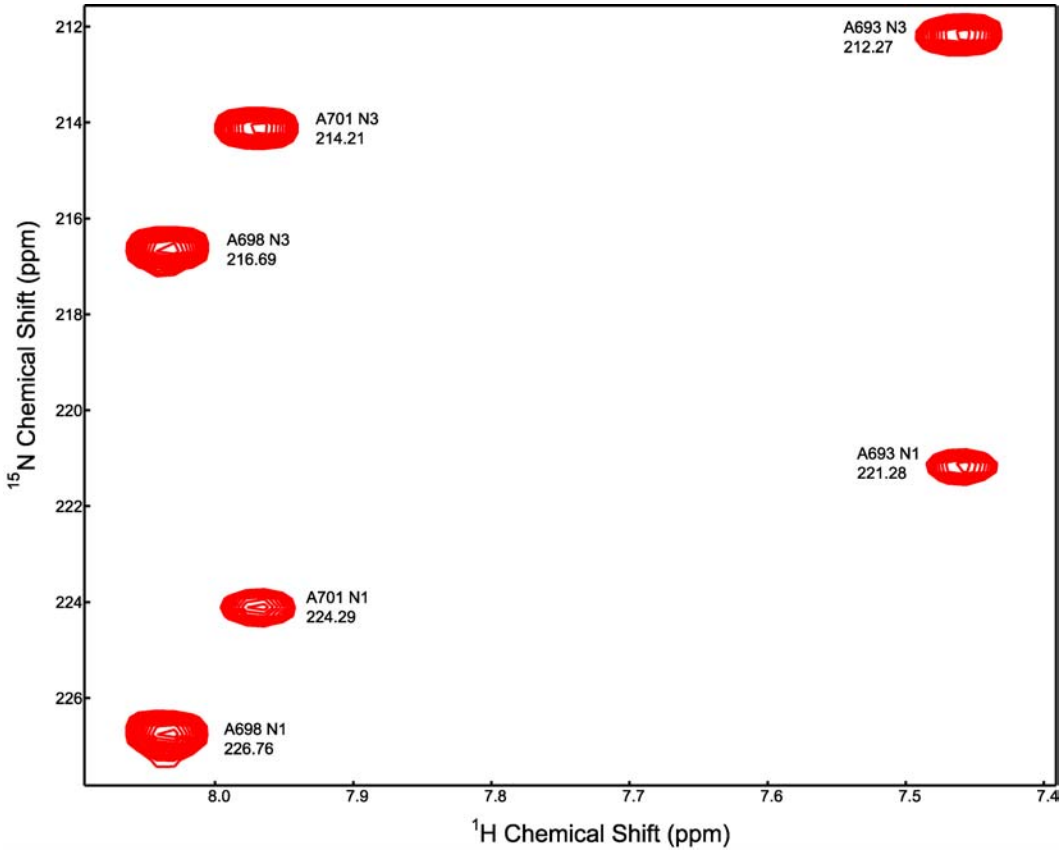


Figure 2.29. Purine H8-N7 region of a modified HNN-COSY (20) spectrum of SL5. The spectrum shows the correlation between the purine H8 and N7.

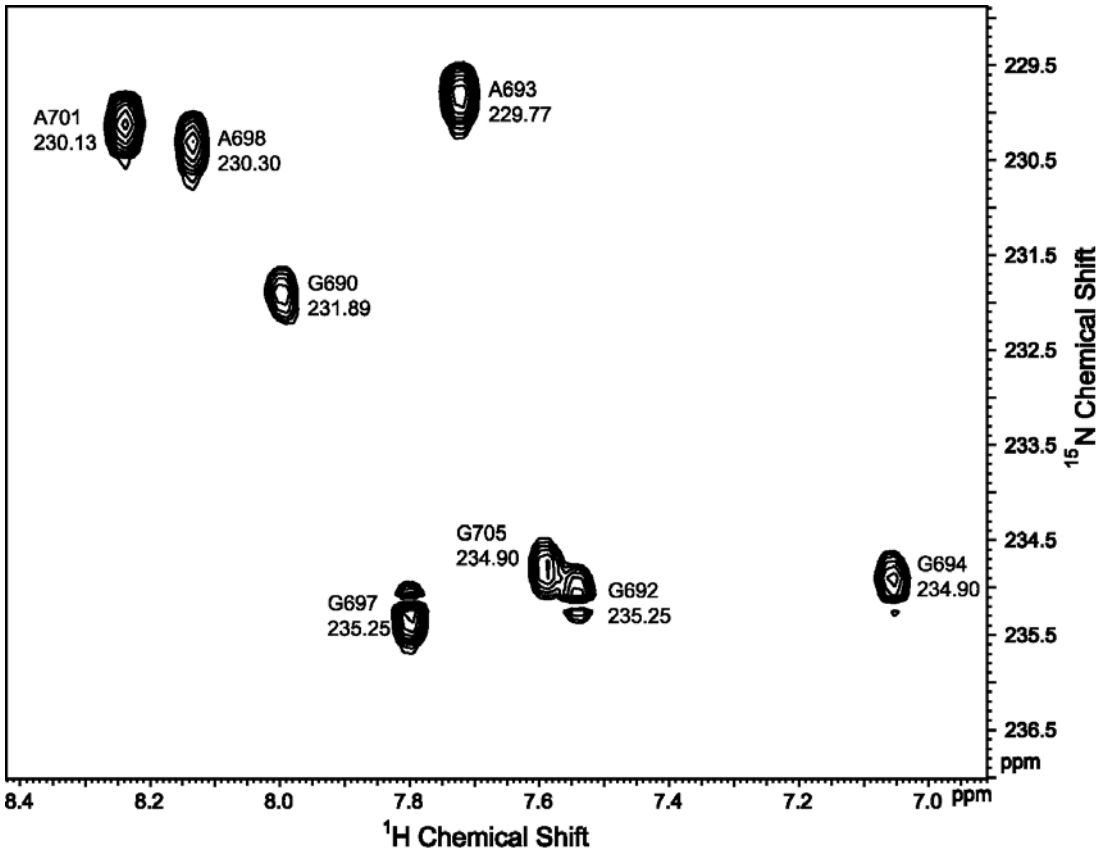


Figure 2.30. 2D ^{1}H ^{31}P-HETCOR (21) spectrum of SL5. The spectrum shows the ^{31}P assignments for all 16 SL5 phosphorous signals. The ^{31}P resonances were assigned by their correlation to the H3' of the previous residue.

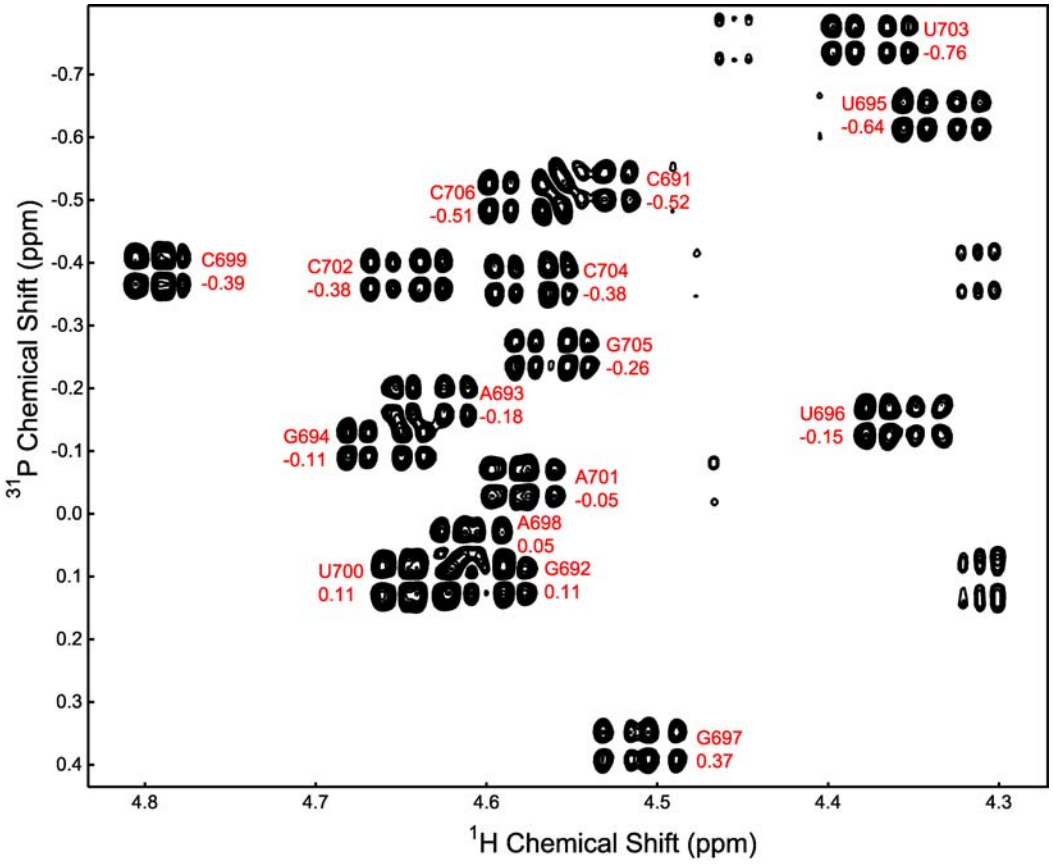


Figure 2.31. Definition of the torsion angles in RNA. The backbone torsion angles are α, β, γ, δ, ε, and ζ. The endocyclic torsion angles are $\nu_{(0-4)}$ and the glycosidic torsion angle χ (2).

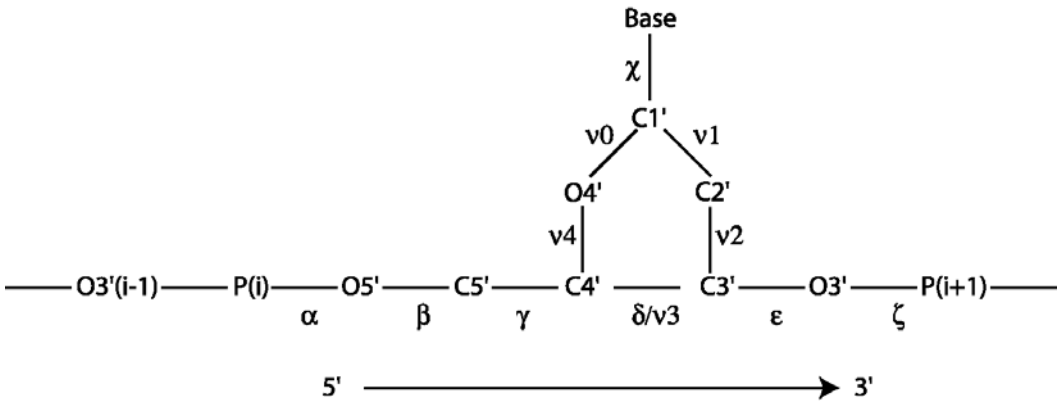
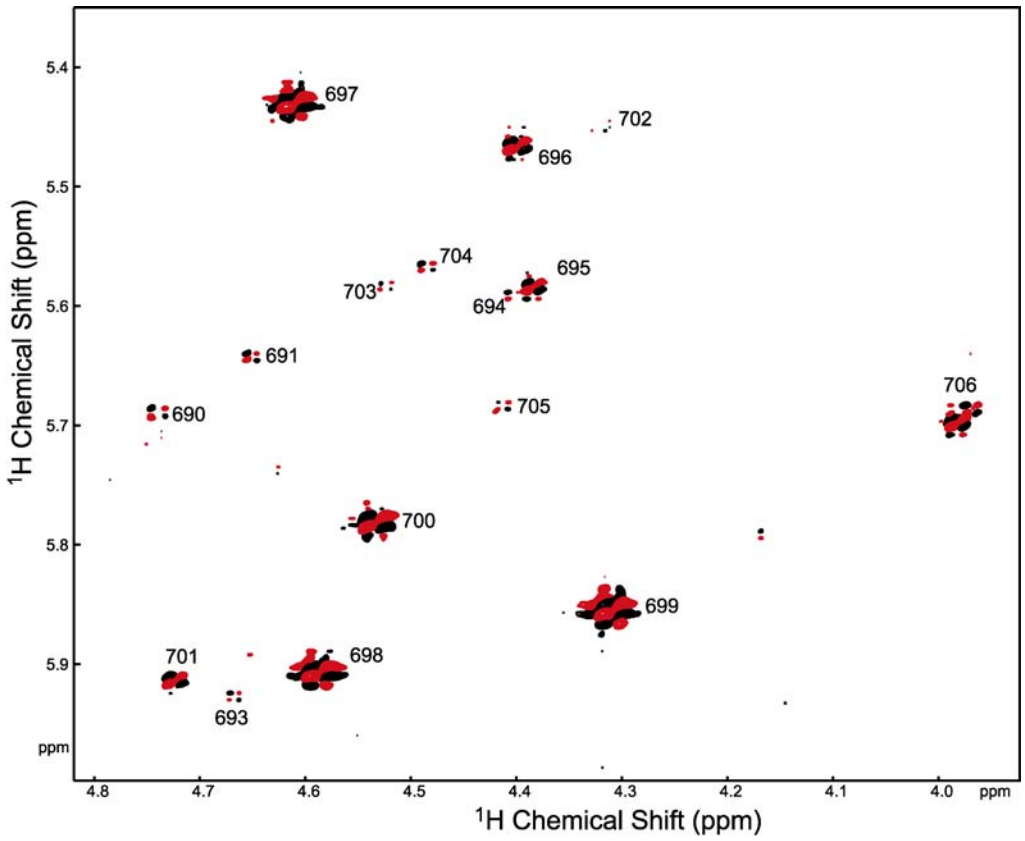


Figure 2.32. H1'-H2' region of the DQF-COSY spectrum of SL5. The intensity of the cross peaks was used to qualitatively deduce the $^{3}J_{H1'/H2'}$ coupling constant.



References

1. Pardi, A. (1995) Multidimensional heteronuclear NMR experiments for structure determination of isotopically labeled RNA. *Methods Enzymol*, **261,** 350-380.

2. Wüthrich, K. (1986) *NMR of proteins and nucleic acids*. Wiley, New York.

3. Hore, P.J. (1995) *Nuclear magnetic resonance*. Oxford University Press, Oxford England ; New York.

4. Piotto, M., Saudek, V. and Skleñár, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR*, **2,** 661-665.

5. Grzesiek, S. and Bax, A. (1993) The importance of not saturating water in protein NMR. Application to sensitivity enhancement of NOE measurements. *J Am Chem Soc*, **115,** 12593-12594.

6. Rahman, A.-U. (1986) *Nuclear magnetic resonance : basic principles*. Springer-Verlag, New York.

7. Kay, L.E., Keifer, P. and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J Am Chem Soc*, **114,** 10663-10665.

8. Roberts, G.C.K. (1993) *NMR of macromolecules : a practical approach*. IRL Press at Oxford University Press, Oxford England ; New York.

9. Mueller, L., Legault, P. and Pardi, A. (1995) Improved RNA structure determination by detection of NOE contacts to exchange-broadened amino protons. *J Am Chem Soc*, **117,** 11043-11048.

10. Simorre, J.P., Zimmermann, G.R., Pardi, A., Farmer, B.T. and Mueller, L. (1995) Triple resonance HNCCCH experiments for correlating exchangeable and nonexchangeable cytidine and uridine base protons in RNA. *J Biomol NMR*, **6,** 427-432.

11. Simorre, J.P., Zimmermann, G.R., Mueller, L. and Pardi, A. (1996) Correlation of the guanosine exchangeable and nonexchangeable base protons in C-13-(IN)-N-15-labeled RNA with an HNC-TOCSY-CH experiment. *J Biomol NMR*, **7,** 153-156.

12. Simorre, J.P., Zimmermann, G.R., Mueller, L. and Pardi, A. (1996) Triple-resonance experiments for assignment of adenine base resonances in C-13/N-15-labeled RNA. *J Am Chem Soc*, **118,** 5316-5317.

13. Skleñár, V., Dieckmann, T., Butcher, S.E. and Feigon, J. (1998) Optimization of triple-resonance HCN experiments for application to larger RNA oligonucleotides. *J Magn Reson*, **130,** 119-124.

14. Ikura, M., Kay, L.E., Tschudin, R. and Bax, A. (1990) Three-dimensional NOESY-HMQC spectroscopy of a ^{13}C-labeled protein. *J Magn Reson*, **86,** 204-209.

15. Pardi, A. and Nikonowicz, E.P. (1992) Simple procedure for resonance assignment of the sugar protons in ^{13}C-labeled RNAs. *J Am Chem Soc*, **114,** 9202-9203.

16. Remin, M. and Shugar, D. (1972) Conformation of the exocyclic 5'-CH 2'-OH in nucleosides and nucleotides in aqueous solution from specific assignments of the H5' and H5" signals in the NMR spectra. *Biochem Biophys Res Commun*, **48,** 636-642.

17. Santoro, J. and King, G.C. (1992) A constant-time 2D overbodenhausen experiment for inverse correlation of isotopically enriched species. *J Magn Reson*, **97,** 202-207.

18. Vuister, G.W. and Bax, A. (1992) Resolution enhancement and spectral editing of uniformly ^{13}C-enriched proteins by homonuclear broadband ^{13}C decoupling. *J Magn Reson*, **98,** 428-435.

19. Legault, P. (1995) Ph.D. Thesis, University of Colorado at Boulder.

20. 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.

21. Skleñár, V., Miyashiro, H., Zon, G., Miles, H.T. and Bax, A. (1986) Assignment of the ^{31}P and ^{1}H resonances in oligonucleotides by two-dimensional NMR spectroscopy. *FEBS Lett*, **208,** 94-98.

22. Levitt, M.H. (2001) *Spin dynamics : basics of nuclear magnetic resonance*. John Wiley & Sons, Chichester ; New York.

23. Johnson, B.A. and Blevins, R.A. (1994) NMRView - A computer program for the visualization and analysis of NMR data. *J Biomol NMR*, **4,** 603-614.

24. Wijmenga, S.S. and van Buuren, B.N.M. (1998) The use of NMR methods for conformational studies of nucleic acids. *Prog NMR Spectr*, **32,** 287-387.

25. Schwalbe, H., Marino, J.P., King, G.C., Wechselberger, P., Bermel, W. and Griesinger, C. (1994) Determination of a complete set of coupling constants in ^{13}C-labeled oligonucleotides. *J Biomol NMR*, **4,** 631-644.

26. Skleñár, V., Peterson, R.D., Rejante, M.R. and Feigon, J. (1994) Correlation of nucleotide base and sugar protons in a ^{15}N-labeled HIV-1 RNA oligonucleotide by ^{1}H-^{15}N HSQC experiments. *J Biomol NMR*, **4,** 117-122.

CHAPTER 3

NMR STRUCTURE OF THE VS RIBOZYME STEM-LOOP V RNA AND MAGNESIUM-ION BINDING FROM CHEMICAL-SHIFT MAPPING[1]

[1] Dean O. Campbell and Pascale Legault. To be submitted to *Journal of Molecular Biology*.

ABSTRACT

An essential step in the substrate recognition of the *Neurospora* VS ribozyme is the formation of a loop-loop interaction between the terminal loops of stem-loop I and stem-loop V. We have determined the structure of stem-loop V by nuclear magnetic resonance spectroscopy and show that it adopts a U-turn conformation, a common motif found in RNA. Structural comparisons indicate that the U-turn of stem-loop V (SL5) fulfills some but not all of the structural characteristics found in canonical U-turn motifs. Nevertheless, the U-turn conformation in stem-loop V allows the bases involved in the loop I / loop V interaction (G697, A698, and C699) to be accessible, by exposing their Watson-Crick faces to the solvent. The formation of the stem-loop I / stem-loop V interaction is magnesium-dependent. Using chemical shift mapping, we show that magnesium ions interact with the loop of the isolated stem-loop V, and possibly induce a conformational change that may be important for interaction with stem-loop I. This study brings some light on the mechanism of substrate recognition in the VS ribozyme and expands our understanding of the role of U-turn motifs in RNA structure and function.

INTRODUCTION

Like many other ribozymes, the *Neurospora* VS ribozyme provides a useful model system for studying the relationship between RNA structural motifs and their implication for RNA function (1-3). It is now well understood that functional RNAs, including ribozymes, adopt complex three-dimensional structures, which involve elaborate arrangements of mainly helical sub-domains (4,5). It is not always clear however, how individual RNA motifs allow proper

orientation of these various domains, participate in tertiary interactions, and interact with metal ions to provide the unique organization of chemical groups at the active site of ribozymes.

The VS ribozyme originates from the 881-nucleotide Varkud Satellite (VS) RNA, which was initially isolated from the mitochondria of certain isolates of *Neurospora* (6). In these mitochondria, the VS ribozyme is involved in RNA processing as part of its replication cycle (7). Both *in vivo* and *in vitro*, the VS ribozyme is a catalytic RNA capable of site-specific cleavage and ligation reactions (3,6,8,9). Fragments of ~120-180 nt derived from the natural VS RNA undergo cleavage at a specific phosphodiester bond between G620 and A621 (9). The secondary structure of the VS ribozyme containing the minimal contiguous sequence for cleavage consists of six helical domains (I-VI) (Figure 3.1a) (10): stem-loop I forms the substrate domain and stem-loops II-VI comprise the catalytic domain (10). *In vivo*, the VS ribozyme likely acts in *cis*, since stem-loop I is found 5' of the catalytic domain on the same RNA chain. *In vitro,* stem-loop I and the catalytic domains can also be synthesized as separate RNA fragments and the catalytic domain can perform cleavage in *trans* (11). The products of the VS ribozyme cleavage reaction have 5'-OH and 2'3'-cyclic phosphate termini as found for other small catalytic RNAs such as the hammerhead, hairpin, and HDV ribozymes (6).

The VS ribozyme differs from other small catalytic RNAs in its mode of substrate recognition. Most ribozymes recognize their substrate primarily via base-pairing interactions between single-stranded regions of the catalytic domain and the complementary single-stranded substrate. In the case of the *Neurospora* VS ribozyme, the catalytic domain recognizes its hairpin substrate (11), through tertiary interactions (1,12). The best-characterized interaction is a loop-loop interaction between stem-loops I and V, which was identified from mutagenesis and structural probing experiments (1). This interaction, also termed the kissing or pseudoknot

interaction, involves Watson-Crick base pairing between nucleotides G630, U631, and C632 in stem-loop I, and nucleotides C699, A698, and G697 in stem-loop V (see Figure 3.1a) (1). Formation of this tertiary interaction is important for efficient cleavage and ligation by the VS ribozyme (1,3).

It has been shown that the interaction of stem-loop I with stem-loop V causes a conformational change in the stem-loop I substrate (Figure 3.1a) (3,13). Indeed, this conformational change can be triggered not only by the full catalytic domain, but also by an isolated stem-loop V RNA (3,13). Upon binding to stem-loop V, stem-loop I is converted from a "unshifted" to a "shifted" conformation in which nucleotides 623-625 in helix Ib shift their base-pairing partners from nucleotides 634-636 to nucleotides 635-637, leading to a rearrangement in the cleavage site internal loop (3). Stem-loop I RNA mutants that can not undergo this conformational change are not cleaved by the VS ribozyme, whereas stem-loop I mutant substrates that can adopt the shifted conformation are active in the cleavage reaction (13). The loop-loop interaction and the conformational change in stem-loop I are dependent on the presence of magnesium ions (1,3,13).

It has been postulated from mutagenesis (1) and nucleotide analog interference mapping (NAIM) (14,15) experiments that stem-loop V forms a U-turn motif, which facilitates its binding to stem-loop I. U-turns were first found in the anticodon and TψC-loop of tRNA (16). They now constitute a common structural motif found in a large number of RNAs, including the hammerhead ribozyme (4), 23s rRNA (17), U2 snRNA (18), and HIV RNA (19). Some of the structural characteristics of U-turns are: an UNR sequence; a sharp turn in the direction of the RNA backbone after the U; formation of hydrogen bonds between U 2'-OH and R N7 and between U H3 and R 3'-phosphate, and stacking of bases after this U (16,20). The residues U696,

G697, and A698 of stem-loop V match the UNR sequence, and mutational studies of these residues have lead to the original idea that stem-loop V forms a U-turn motif (1). NAIM interference data indicate that substitution of the 2'-OH group of U696 by a 2'-H leads to decreased activity of the VS ribozyme, possibly due to the removal of an important hydrogen bond between U696 2'-OH and A698 N7 in the U-turn of stem-loop V (14,15). Similarly, phosphorothioate substitution of A 698 3'-phosphate leads to decreased activity of the VS ribozyme, possibly due to the removal of an important hydrogen bond between one of the non-bridging oxygen of this phosphate and U696 H3, although direct divalent metal-ion binding proposed at this phosphate could also explain the loss of cleavage activity (15,21).

In this manuscript, we present the solution NMR structure of a stem-loop V RNA and show that its loop adopts a U-turn motif. By comparing the structural characteristics of this U-turn with that of canonical U-turn motifs selected from the structural database, we notice that the stem-loop V RNA was missing a few important structural characteristics of canonical U-turn motifs. Interestingly, the stem-loop V bases that are involved in the loop-loop interaction have their Watson-Crick faces accessible for base pairing with stem-loop I. Using chemical shift mapping, we find that magnesium ions interact with the U-turn of stem-loop V, likely producing a conformational change in this loop. Our finding help to better understand substrate recognition in the VS ribozyme and the role of U-turn motifs in tertiary interactions.

MATERIALS AND METHODS

Sample Preparation

The 17-mer SL5 RNA was synthesized *in vitro* using T7 RNA polymerase, a synthetic oligonucleotide template modified with C2'-methoxyl at the last two 5'-nucleotides of the coding

strand (22,23) (Macromolecular Resources, CO), and unlabeled, ^{15}N-labeled, or ^{13}C/^{15}N-labeled nucleoside triphosphates. T7 RNA polymerase and isotopically-labeled nucleoside triphosphates were produced in-house according to published procedures (24,25). SL5 RNA purification was achieved via 20% denaturing gel-electrophoresis, dephosphorylation at its 5'-end with calf alkaline phosphatase (Roche Molecular Biochemicals, NJ), and DEAE-Sephacel chromatography (Amersham Biosciences, NJ). The RNA sample was concentrated and the buffer was exchanged in Centricon-3 ultrafiltration devices (Millipore, MA) with NMR buffer (10 mM d_{11}-Tris, 50 mM NaCl, 0.2 mM EDTA, 0.05 mM NaN_3, pH 7.0, 90% H_2O and 10% D_2O). For studies in D_2O, the RNA samples were transferred to 99.996 % D_2O through multiple cycles of lyophilisation and resuspension in D_2O. The SL5 RNA concentration of the NMR samples ranged from 0.8 to 3.4 mM as determined from the absorbance at 260 nm using a coefficient of 40 μg/ml per OD unit. Before each set of NMR experiments the RNA sample was heated to 95 ˚C for two minutes and then immediately snap-cooled in iced water.

NMR Spectroscopy

All NMR experiments were conducted at 5 ˚C, 15 ˚C or 25 ˚C on Varian Inova 600 MHz or 800 MHz spectrometers equipped with pulse-field gradient units and actively shielded z gradient probes, either a ^{1}H{^{13}C/^{15}N} triple resonance probe or a ^{1}H{^{15}N-^{31}P} indirect detection probe. Assignments of ^{1}H, ^{13}C, ^{15}N and ^{31}P nuclei in SL5 were obtained from two-dimensional (2D) and three-dimensional (3D) homonuclear and heteronuclear NMR experiments. All non-exchangeable protons and their attached carbons were assigned from the following experiments collected at 25 ˚C in D_2O: 2D ^{1}H-^{13}C CT-HSQC (26,27); 2D ^{1}H-^{15}N MQ-(HC)N(C)H (28); 3D HCCH-COSY (29); 3D HCCH-TOCSY (29); and a 3D ^{13}C-edited HMQC-NOESY (30). The

exchangeable protons and their attached nitrogens were assigned from the following experiments collected at 25 ˚C in H_2O: 2D imino- and amino-optimized 2D ^{1}H-^{15}N HSQC (31); 2D H(NCCC)H for uracil and cytosine residues (32); 2D H(NC)-TOCSY-(C)H for guanosine residues (33), 2D (H)N(C)-TOCSY-(C)H for adenosine residues (34); 2D ^{1}H-^{15}N CPMG-NOESY (35); and a 2D ^{1}H-^{1}H flip-back watergate NOESY (36,37). Non-protonated nitrogens were assigned from a 2D ^{1}H-^{15}N HMQC optimized for transfers via J = 7.0 Hz (38) collected in D_2O. A 2D ^{1}H-^{31}P HETCOR (39) spectrum was also collected in D_2O at 25 ˚C for the assignment of the ^{31}P resonances. Distance restraints were obtained from a 2D ^{1}H-^{1}H flip-back watergate NOESY spectrum collected in H_2O at 25 ˚C with a mixing time of 400 ms (36,37), 2D ^{1}H-^{15}N CPMG-NOESY spectra collected at 5 ˚C and 25 ˚C in H_2O with mixing times of 150 ms (35) and 3D ^{13}C-edited HMQC-NOESY (30) spectra collected in D_2O at 25 ˚C with mixing times of 90 ms and 180 ms. Hydrogen-bonding restraints were obtained from 2D HNN-COSY spectra collected at 25 ˚C (40,41). Restraints for the δ torsion angle were obtained from experiments collected at 25 ˚C in D_2O: a 2D DQF-COSY; a 3D HCCH E.COSY (42), and a 2D ^{1}H-^{13}C CT-TROSY for measurement of CSA-dipolar cross-correlated rate constants (at 600 MHz and 800 MHz) (43). All spectra were processed with the NMRPipe/NMRDraw package (44) and analyzed with NMRView (45). ^{1}H, ^{13}C and ^{15}N chemical shifts were referenced at 25 ˚C to an external standard of DSS at 0.00 ppm (46) and ^{31}P chemical shifts were referenced at 25 ˚C to an external standard of 85% phosphoric acid at 0.00 ppm.

Structure Calculation

The distance restraints from the 2D ^{1}H-^{1}H NOESY and 3D ^{13}C-edited HMQC-NOESY spectra were separated into four ranges: strong (1.8-3.0 Å), medium (1.8-4.1 Å), and weak (1.8-

5.5 Å) based on the intensities of peaks observed at a mixing time of 90 ms; and very weak (1.8-7.0 Å) for signals observed only at mixing times greater than 90 ms. The distance restraints obtained from the 2D ^{1}H-^{15}N CPMG-NOESY were given ranges of either 1.8-5.5 Å or 1.8-7.0 Å based on crosspeak intensities. The distance restraints were calibrated based on the crosspeaks intensities and by comparison with NOE crosspeaks involving protons separated by known intra-residue distances. 2D HNN-COSY spectra (40,41) were collected to detect $^{2}J_{NN}$ couplings across hydrogen bonds in Watson-Crick base pairs. Because of the strong NMR evidence (NOESY and HNN-COSY) for the formation of the A-U and G-C Watson-Crick base pairs in the stem of SL5, canonical distance restraints were employed to define the hydrogen-bonding pattern and planarity of the first five base pairs in the stem of SL5 (residues 690-694 and 702-706). We obtained limited NOE data for the U695-A701 base pair, therefore, in this case, only the U695 N3 to A701 N1 distance was defined (2.82 Å ± 0.1 Å) based on the HNN-COSY data. Sugar pucker conformations were derived from 2D DQF-COSY, 3D HCCH E.COSY (42), and 2D ^{1}H-^{13}C CT-TROSY (43) experiments. All the sugar puckers except for the loop residues 697-700 were set to C3'-endo (δ = 86 +/- 10˚). The γ torsion angle restraints were derived from comparative analyses of NOE data (47). The γ angles for all the residues except the loop residues 697-700 were set to the gauche+ conformation (γ = 60 +/- 20˚).

Three-dimensional structures were calculated with restrained molecular dynamics and simulated annealing in X-PLOR-NIH 2.0.6 (48,49) using a modified version of a previously reported protocol (50). The experimental restraints described in Table 3.1 were used along with restraints to maintain the covalent structure and stereochemistry of the RNA. The CHARMM force field was used with the ribose bond angles modified to prevent flattening of the ring. The same reduced force field was used at all steps except for the first energy minimization. This force

field consists of bond, angle, improper (stereochemical), and repulsive van der Waals energy terms, and NOE and torsion angle pseudoenergy terms. Electrostatic contributions were not included in the force field. Initial structures were generated by randomizing the backbone (α, β, γ, δ, ε, and ζ) and χ angles followed by an energy minimization with a force field that included bond, angle, improper and repulsive van der Waals energy terms. A first cycle (29 psec) of high-temperature dynamics (3000 K) was used to introduce the NOE and torsion angle pseudoenergy terms, gradually increase the weight on the van der Waals energy term, and gradually transform the X-PLOR pseudoenergy NOE term from a soft to a square well. The system was then slowly cooled down (3000 K to 300 K in 6.25 ps) and energy minimized. A second cycle (10 ps) of high-temperature dynamics (3000 K) was then used in which the weight on the torsion angle pseudoenergy term was reduced, slowly increased, reduced again, and then slowly increased up to its final value. The system was subsequently cooled down very slowly (3000 K to 300 K in 12.5 ps) and energy minimized. This second cycle helps improving the percentage of accepted structure. Starting from a set of 100 structures with randomized torsion angles, 61 structures satisfied the experimental restraints (no distance violation > 0.1 Å and no torsion angle violation > 5˚). From these 61 structures, the 10 lowest energy structures were selected for analysis and were used to calculate an average structure that was further minimized against experimental restraints. All structures were visualized with MOLMOL (51) and analyzed with MOLMOL and CURVES (52).

Metal-Ion Binding Studies

Two 1.0 mM samples of $^{13}C/^{15}N$-labeled SL5 in NMR buffer A (10 mM d_{11}-Tris, pH 7.0, 0.02% NaN_3, 50 mM NaCl) were titrated with salts, one with 99.995% $MgCl_2$ (Sigma-Aldrich,

MO) and the other one with NaCl. The titrations were both carried out by adding an increasing amount of a concentrated salt solution directly to the NMR sample. For the $MgCl_2$ titration, the divalent metal concentration was adjusted from 0 mM to 0.25 mM, 0.50 mM, 0.75 mM, 1.0 mM, 2.0 mM, 3.0 mM, 5.0 mM, 10 mM, 20 mM, 30 mM, 40 mM, 50 mM, and 130 mM. For the NaCl titration, the monovalent salt concentration was adjusted from 50 mM to 100 mM, 150 mM, 250 mM, 350 mM, and 500 mM. Chemical shift changes were monitored after each salt addition by collecting 2D ^{1}H-^{13}C CT-HSQC (27) spectra. To quantify the chemical shift changes (Δ_T in ppm $\pm$ 0.05 ppm), the following equation was used: $\Delta_T = [(\Delta_H)^2 + (R*\Delta_C)^2]^{1/2}$, where Δ_H is the ^{1}H chemical shift change, Δ_C is the ^{13}C chemical shift change, and R is a scaling factor calculated as the ratio between the ^{1}H and ^{13}C chemical shift ranges for a particular type of $^{1}H/^{13}C$ pair (53). The change in chemical shift as a function of $MgCl_2$ concentration was plotted for all the $^{1}H/^{13}C$ pairs that had a Δ_T value of > 0.3 ppm at the end of the titration. These binding data were used to calculate apparent K_ds for each member of resolved $^{1}H/^{13}C$ pairs belonging to the loop residues. Assuming fast exchange and a 1:1 stoichiometry, the apparent K_ds were calculated by fitting the following equation: $\Delta_{obs} = (\Delta_T)/(2*[RNA]_T)*\{([Mg]_t + [RNA]_T + K_d) - (([Mg]_t + [RNA]_T + K_d)^2 -(4*[Mg]_t *[RNA]_T))^{1/2}$, where Δ_{obs} is the observed chemical shift change at each $MgCl_2$ concentration, $[Mg]_t$ and $[RNA]_t$ are the total concentration of magnesium ion and RNA, respectively (54,55). All binding curves were fitted using the least-square analysis function of GRACE (http://plasma-gate.weizmann.ac.il/Grace/).

Relaxation Measurements

The ^{13}C $T_{1\rho}$ measurements on the aromatic C6/C8 carbons were performed on a $^{13}C/^{15}N$-labeled SL5 RNA in 10 mM d_{11}-Tris, 50 mM NaCl, 0.2 mM EDTA, 0.05 mM NaN_3, pH 7.0, and

100% D_2O. Measurements were performed at 600 MHz and 25 ˚C with the pulse sequence from Yamazaki, T. *et al.* (56). Nine different spin-lock delays T between 4 ms and 60 ms were used: 4 ms, 8 ms, 12 ms, 16 ms, 24 ms, 32 ms, 40 ms, 48 ms, and 60 ms. For $T < 20$ ms, 20 ms $\leq T < 40$ ms, 40 ms $\leq T < 60$ ms, and T = 60 ms, k = 0, 1, 2, 3 ^{1}H 180˚ pulses, respectively, were applied during the ^{13}C spin-lock interval according to the scheme $[T/(2k) - {}^1H\ 180^\circ - T/(2k)]_{k,k\neq 0}$ in order to prevent cross-correlation effects which can lead to overestimation of ^{13}C $T_{1\rho}$ values (57). The ^{13}C carrier was positioned in the center of the C6/C8 region (139.7 ppm), and a spin-lock power of 3.8 kHz was used. At the largest offset for C6 and C8 resonances, which range from 135.8 ppm to 143.6 ppm, the tip angle of the spin-lock axis was 81˚. Relaxation times were extracted by non-linear curve fitting of the maximum signal intensity from the 2D spectra using GRACE (http://plasma-gate.weizmann.ac.il/Grace/). A 2 parameter fit $\{I(t) = [A_0 * e(-t/T_{1\rho})]\}$ of the experimental decay curve gave low correlation coefficients, likely because of poor signal-to-noise at the longer spin-lock delays. A 3 parameter fit $\{I(t) = [A_0 * e(-t/T_{1\rho})] + A_2\}$ was therefore employed (58), and in all cases the correlation coefficients of the fit were ≥ 0.99. Reported error bars are taken from the uncertainties in the non-linear fit.

RESULTS

Conformation of SL5 RNA

To gain insights into substrate recognition by the VS ribozyme, we synthesized a 17-nucleotide fragment of stem-loop V, termed SL5 (Figure 3.1b), for NMR structural determination. Based on the secondary structure of the VS ribozyme (10), SL5 RNA was designed to form a hairpin structure which retain the natural sequence within its five-member terminal loop and its three-closing base pairs (Figure 3.1b). To ensure that it did indeed form a

hairpin and not a duplex under NMR conditions, SL5 RNA was first analyzed by equilibrium ultracentrifugation. Ultracentrifugation studies indicated that this RNA forms a homogenous sample of unimolecular weight (~6 kD; data not shown) at low concentration (2 μm). This RNA also gave high-quality NMR data at higher concentration (0.8-3.5 mM) consistent with a single conformation, and small aromatic proton linewidths (2.5-5 Hz) were indicative of hairpin formation. We noticed that additional broader signals were present in the spectra when we omitted to heat and snap-cooled the sample prior to collecting the NMR data. It was concluded that the hairpin conformation of SL5 can slowly interconvert to a duplex form over periods of weeks, and we took the necessary precaution to collect the NMR data on the hairpin conformation exclusively.

SL5 Structure Determination

The NMR structure of SL5 was obtained using standard homonuclear and heteronuclear NMR methods and unlabeled, ^{15}N-labeled, and ^{13}C/^{15}N-labeled RNA. The ribose spin systems were assigned unambiguously using 2D ^{1}H-^{13}C CT-HSQC (26,27), 3D HCCH-COSY (29) and 3D HCCH-TOCSY (29) experiments. The base spin systems were assigned unambiguously using 2D ^{1}H-^{13}C CT-HSQC (26,27), 3D HCCH-COSY for pyrimidine residues (29), 2D ^{1}H-^{15}N HSQC (31), 2D H(NCCC)H for pyrimidine residues (32); 2D H(NC)-TOCSY-(C)H for guanosine residues (33), 2D (H)N(C)-TOCSY-(C)H for adenosine residues (34). A 2D ^{1}H-^{15}N MQ-(HC)N(C)H spectrum was collected to establish unambiguous intraresidue ribose-base connectivities. A 2D ^{1}H ^{31}P-HETCOR spectrum (39) helped determine sequential connectivities and specific ^{31}P assignments. In addition to these through-bond experiments which contributed many unambiguous connectivities, 2D ^{1}H-^{1}H flip-back watergate NOESY (36,37), 2D ^{1}H-^{15}N

CPMG-NOESY (35) and 3D ^{13}C-edited HMQC-NOESY spectra provided through-space connectivities which allowed for complete sequential assignment of SL5 RNA. Assignment of ^{1}H, ^{13}C, ^{15}N and ^{31}P nuclei in SL5 are summarized in Table 3.S1.

The combination of the NOESY and the HNN-COSY experiments (40,41) gave direct evidence for base pairing in the stem of SL5. We observed only four imino proton signals in the 1D ^{1}H spectrum of SL5 (Figure 3.2a). Based on NOE connectivities, these imino proton signals were easily assigned to guanine and uridine bases of the four internal base pairs in the stem, C691-G705, G692-C704, A693-U703, G694-C702. In addition, we obtained unambiguous evidence for the formation of these base pairs based on a 2D HNN-COSY spectrum collected to detect $^{2}J_{NN}$ couplings across hydrogen bonds in Watson-Crick base pairs (40) (Figure 3.2a). For Watson-Crick G-C base pairs, this experiment provides correlations between the guanine imino proton (G H1) and its bonded imino nitrogen (G N1) and between G H1 and the imino nitrogen of the paired cytidine (C N3). We obtained these correlations for the C691-G705, G692-C704, and G694-C702 base pairs (Figure 3.2a). For Watson-Crick A-U base pairs, this experiment provides correlations between the uridine imino proton (U H3) and its bonded imino nitrogen (U N3) and between U H3 and the imino nitrogen of the paired adenine (A N1). We observed these correlations for the A693-U703 base pair (Figure 3.2a). One limitation of this HNN-COSY experiment is that $^{2}J_{NN}$ couplings across hydrogen bonds can only be detected if the intervening imino proton is observable. Therefore it did not provide evidence for the U695-A701 base pair predicted from the secondary structure model (10) (Figure 3.1). In order to verify the presence of this base pair, we used a modified version of the HNN-COSY experiment that allows detection of $^{2}J_{NN}$ couplings across hydrogen bonds in A-U base pair via the non-exchangeable adenine H2 (41) (Figure 3.2b). For all adenine residues, this experiment provides intra-base

correlations between the adenine H2 and the adenine N1, N3, and N9 (not shown). For adenines in Watson-Crick A-U base pairs, this experiment also provides a correlation between the adenine H2 and the N3 of the paired uridine. As expected, the A H2 - U N3 correlation was not observed for A698 in the loop, but was observed for the well-defined A693-U703 base pair (Figure 3.2b). In addition, we observed a correlation between A701 H2 and a uridine N3 signal (Figure 3.2b). This N3 chemical shift was unambiguously assigned to U695 N3 based on the U695 N3 – U695 H5 correlation observed in a long-range ^{1}H-^{15}N HMQC (38) (not shown). The modified HNN-COSY spectrum helped confirm the presence of a Watson-Crick U-A base pair between U695 and A701 in the absence of a detectable imino proton for U695. However, planarity or geometry constraints typical of Watson-Crick U-A base pairs were not enforced for these residues in the structure calculation. Instead a single distance constraint was defined between U695 N3 and A701 N1 based on the HNN-COSY data.

In addition to showing base pairing in the stem, the NMR data indicate that the residues in the stem (G690-U695 and A701-C706) adopt a typical of A-form helix. Intra- and inter-residue NOE patterns are characteristics of A-form helical geometry (47). We analyzed three NMR experiments to determine sugar pucker conformation: a 2D DQF-COSY for measurement of $^{3}J_{H1'-H2'}$ coupling constants; a 3D HCCH E. COSY for measurement of $^{3}J_{H1'-H2'}$ and $^{3}J_{H3'-H4'}$ coupling constants (42); and a 2D ^{1}H-^{13}C CT-TROSY for measurement of CSA-dipolar cross-correlated rate constants (43). All three experiments indicate that residues in the stem (G690-U695 and A701-C706) as well as residue U696 adopt a C3'-endo sugar pucker conformation. Comparative analyses of NOE data, also indicate that all residues in the stem (G690-U695 and A701-C706) as well as residue U696 have their γ torsion angles in the gauche+ conformation (59). The combination of torsion angle data (not shown) and NOE data indicate that the A-form

helix topology extends to U696 on the 5'-strand of the RNA. For residues G697-U700, the available NMR data indicated that theirs sugars exit as a mixture of rapidly interconverting C3'-endo and C2'-endo conformers (not shown). Therefore, γ and δ torsion angles of only residues G690-U696 and A701-C706 were constrained for structure calculation.

NOEs which are not characteristic of A-form geometry were observed in the loop of SL5 and represent important features of this loop, including the stacking pattern and the position of phosphate backbone reversal (Figures 3.3 and 3.4). We observed NOEs between the H2' of U696 and the following protons: A698 H8, G697 H8, C699 H6, and C699 H5 (Figure 3.3a) and between the H1' of U696 and the following protons: A698 H8, A701 H2, G697 H8, C699 H6, and C699 H5 (Figure 3.3b). These NOEs from U696 to residues G697, A698, C699, and A701 indicated that there is a sharp turn in the loop after U696, which positions U696 in a central location with respect to these subsequent loop residues. Another indication of this turn is the absence of base-to-base NOEs between U696 and G697 (Figure 3.4), which suggests that these bases are not stacked on each other. The sequential ribose-to-base and base-to-base NOEs continue from G697 through to C699 in a similar fashion to what is observed for the stem, indicating that G697 stacks on A698, which stacks on C699. However, at U700 there is a disruption of this pattern. Non-sequential NOEs between the ribose of C699 and the base of A701, combined with the absence of base-to-base NOEs between C699 and U700 suggests that the base of U700 is somewhat extruded from the loop (Figure 3.4). In addition, an NOE between the H4' of A698 and the H1' of U700 provide evidence that the ribose of U700 also adopts an unusual position within this loop.

A schematic of the observed NOEs summarizes the distance information used for structure calculation and provides an initial representation of the SL5 structure (Figure 3.4). In

addition to NOE-derived distance constraints, we used data from the HNN-COSY experiments and torsion angle constraints (see Materials and Methods) as input for structure calculation. Three-dimensional structures were calculated using restrained molecular dynamics and simulated annealing. The ten lowest energy structure that satisfied our experimental restraints (no distance violation > 0.1 Å and no torsion angle violation > 5˚) were retained for analysis. An average structure was calculated from these low energy structures and further minimized against experimental restraints. The structural statistics (Table 3.1) indicate that the structure of the SL5 RNA, including the terminal loop, is well defined by the NMR data. The superposition of the 10 lowest energy structures on the minimized average structure (Figure 3.5a) also illustrates the overall good precision for the SL5 RNA structure.

The SL5 Loop Forms a Loose U-Turn Motif

As predicted from the secondary structure model of Beatie *et. al.*, the SL5 RNA forms a hairpin characterized by a six-base-pairs stem and and a five-member terminal loop (Figure 3.5). The SL5 stem adopts a typical A-form helix. The U695 imino proton was not observed in our NMR data and only a few distance restraints were used to constrain the U695-A701 base pair in the structure calculation. However, we find that this base pair adopts a canonical Watson-Crick U-A base pair that is stacked unto G694-C702 and is part of the A-form helix. On the 3'-strand base stacking is continuous in the stem only (from C706 to A701), whereas on the 5'-strand base stacking continues beyond the stem residues (from G690 to U695) up to U696, the first member of the loop. U696 is however not involved in any base-pairing interactions with U700 or other loop residues. Directly after U696 there is a sharp turn in the direction of the backbone of SL5 (Figure 3.5). The next three residues after the turn (G697-C699) are all stacked on each other,

and their Watson-Crick faces are exposed to the solvent. The final loop residue U700 has its base completely extruded from the bases of the other residues in the loop and in the stem. As observed in the minimized average structure (Figure 3.5b), the base of U700 is stacked unto the 5'-phosphate group of C699. With the exclusion of the base of U700 from the interior of the loop, it seemed possible for C699 to stack unto A701 and form a base-pairing interaction with U696. However, this is not the case in all the ten calculated structures. Instead, the base of C699 is oriented at an angle and points towards the base of A701, and C699 is too far to form any base-pairing interaction with U696.

The overall geometry of the first four residues of the SL5 loop (U696-C699) resembles that of a U-turn motif. The characteristics of a canonical U-turn motif are: a UNR sequence (U = uracil, N = any base, R = purine), a non-Watson-Crick base pair involving the nucleotide 5' of the UNR sequence, a sharp turn in the backbone after the U, stacking of the bases 3' of the U, stacking of the U base with the 5'-phosphate group of the R, and two hydrogen-bonding interactions involving the U residue (16,20,60). These hydrogen-bonding interactions are between the U 2'-OH and the R N7 and between the U H3 and the 3' phosphate group of the R. The U-turn characteristics of SL5 are summarized in Table 3.2. In SL5, the UNR sequence is represented by U696, G697 and A698. As opposed to the canonical U-turn, the nucleotide U695 5' of the UNR sequence forms a Watson-Crick U-A base pair with A701. We observed a sharp turn in the backbone after U696, which is characterized by an unusual value of the α torsion angle of G697 (116 $\pm$ 7 Å in the 10 lowest energy structures). The three bases immediately after the turn (G697, A698, and C699) are stacked on each other. However, the base of U696 does not stack with the phosphate group of A698. As for the hydrogen-bonding interactions that are known to stabilize the U-turn, only one of these hydrogen bonds is present in SL5. The distance

between the U696 2'-OH and A698 N7 is 2.77 Å in the minimized average structure, which is indicative of a hydrogen bond. However, U696 H3 and the 5'-phosphate group of C699 are 8.65 Å apart in the minimized average structure, and are clearly not within hydrogen-bonding distance of each other. We have compared the heavy atom distances that are characteristic of U-turn motifs measured in SL5 with those of a few selected canonical U-turns in other RNA structures (Table 3.3). The U O2' to R N7 distance of SL5 is close to the range observed for other U-turn motifs, whereas the U N3 to R 3'P is much longer than the range observed for these U-turn motifs. In summary, SL5 forms a U-turn motif, but lacks some of the structural characteristics found in canonical U-turns. We are referring to the U-turn of SL5 as a loose U-turn motif to emphasize the contrast between its structure and that of the more compact backbone fold adopted by canonical U-turn motifs.

Effect of Metal Ions on the SL5 RNA Structure

The catalytic activity of the *Neurospora* VS ribozyme requires the presence of metal ions. Divalent and monovalent salts, such as $MgCl_2$, $MnCl_2$, $CaCl_2$, NH_4OAC, and Li_2SO_4 support catalysis to various degrees, with magnesium being the most likely catalytically-active biological cation (61,62). It has been shown that magnesium ions are necessary for formation of the loop-loop interaction between stem-loops I and V (1,13). In stem-loop V, the 5'-phosphate of C699 has been implicated in direct divalent metal-ion coordination from managanese rescue of phosphorothioate inhibition (21). We therefore used chemical-shift mapping to investigate the effect of $MgCl_2$ on the structure of SL5, and to possibly identify Mg^{2+}-binding site(s). A $MgCl_2$ titration was performed under the conditions used for structure determination by several additions of concentrated $MgCl_2$ to achieved final $MgCl_2$ concentrations ranging from 0-130

mM. After each $MgCl_2$ addition a 2D ^{1}H-^{13}C CT-HSQC (26,27) spectrum was collected to monitor chemical shift changes. Almost all the peaks observed in the ^{1}H-^{13}C CT-HSQC spectrum were affected by addition of $MgCl_2$ as illustrated for the H6-C6/H8-C8 region of the spectra (Figure 3.6a). Changes in ^{1}H and ^{13}C chemical shifts attain their mid-point around 2-3 mM $MgCl_2$ and have reached a plateau by the last point of titration (130 mM $MgCl_2$). Total changes in ^{1}H and ^{13}C chemical shifts (Δ_T) between 0 mM and 130 mM $MgCl_2$ were calculated for all the well-resolved peaks in the ^{1}H-^{13}C CT-HSQC spectrum (Figure 3.6b). Significant changes (Δ_T > 0.2 ppm) are observed in the SL5 loop (residues U696-U700), its closing base-pair (U695-A701) and a few stem residues (G690, G694, and C702).

In order to ascertain that the chemical shift changes were not simply a result of increased ionic strength (*I*), we compared the total changes in chemical shifts obtained at 130 mM $MgCl_2$ in NMR buffer that contained 50 mM NaCl (I = 0.44 M) from those obtained in NMR buffer supplemented with 450 mM NaCl (I = 0.50 M). Although chemical shift changes were observed in the presence of 500 mM NaCl (Supplementary Figure 3.S1), the magnitude of these changes was usually smaller, particularly for loop residues, than what we observed with 130 mM $MgCl_2$ in the presence of 50 mM NaCl (Figure 3.6b). These results indicate that the observed chemical shift changes may be partially due to ionic strength effects, but that specific magnesium-ion binding must be inferred to explain the larger chemical shift changes observed with magnesium ions under similar ionic conditions.

The largest Mg^{2+}-dependent chemical shift changes (Δ_T > 0.3 ppm) were observed mainly for SL5 loop residues and these changes were mapped onto the structure of SL5 (Figure 3.6c). All residues in the loop are affected by $MgCl_2$, and it is not possible to precisely map one or more magnesium binding sites using these chemical shift changes. Furthermore, the chemical

shift changes in the loop may not only be due to the presence of one or more local charges, but may also result from conformational changes in the SL5 loop. Apparent K_ds for magnesium ions were calculated based on Mg^{2+}-dependent ^{1}H and ^{13}C chemical shift changes for the 34 resolved signals of the SL5 loop residues (U696 to U700). We obtained an average apparent K_d ($\pm \sigma$) of 2.4 ± 0.5 mM for magnesium-ion binding to the SL5 loop.

Evidence of Internal Motion in the SL5 Loop from ^{13}C Relaxation

Measurement of ^{13}C $T_{1\rho}$ relaxation times has proven useful to probe conformational dynamics in RNA (63). It allows for identification of dynamics on picosecond to nanosecond timescale when $T_{1\rho}$ values are larger than those of rigid residues and for identification of dynamics on the microsecond to millisecond timescale when $T_{1\rho}$ values are smaller that those of rigid residues. In order to identify unusual motion in the SL5 RNA, particularly in its U-turn loop, aromatic ^{13}C (C6/C8) $T_{1\rho}$ relaxation times were determined for all resolved resonances in the ^{1}H-^{13}C correlation spectrum of SL5 (Figure 3.7). The measured $T_{1\rho}$ values of all C6 and C8 carbons in the SL5 stem, excluding those of terminal residues, vary from 18.3 ms to 26.9 ms. The $T_{1\rho}$ values for a particular carbon type (in this case C6 or C8) should be identical for all nucleotides that are rigid with respect to the molecular frame. For SL5, the C6/C8 $T_{1\rho}$ values within the stem are similar but not identical within experimental error, indicating the presence of some internal motions within the helical stem region. Also, no difference was observed within experimental error between the $T_{1\rho}$ values of C6 and C8 residues. These ^{13}C relaxation data indicate that the stem of SL5 RNA does not form a fully rigid structure, but may be best described by a semi-rigid body given the fairly narrow range of C6/C8 $T_{1\rho}$ values and the high-quality of the NMR data for the stem residues. The molecular details of the undergoing dynamics

in the stem are unclear, but the dynamic processes are possibly related to base-pair opening/closing within this small RNA with a marginally stable secondary structure (see Conformation of SL5 RNA).

The C6/C8 $T_{1\rho}$ values of most loop residues (U696, G697, U700, and A701) fall within the range observed for the stem, indicating that no unusual dynamics characterize these residues. Indeed, the uniformity of the observed C6/C8 $T_{1\rho}$ relaxation times throughout the stem and part of the loop is consistent with a well-ordered RNA structure where the internal motions of most residues are essentially similar. Interestingly, a significant difference is observed between the C6/C8 $T_{1\rho}$ values of residues A698 (32.5 ± 0.5 ms) and C699 (35.1 ± 1.2 ms) in the loop and those of the stem residues (Figure 3.7). These larger C6/C8 $T_{1\rho}$ values indicate that these bases are undergoing internal motion on the pico- to nanosecond timescale. It is important to note that A698 and C699 have been implicated in two of the three base pairs which stabilize the loop-loop interaction with stem-loop I (1), and that their unusual dynamics may be important for this tertiary interaction.

DISCUSSION

Stem-loop V performs at least two important functions in the cleavage mechanism of the VS ribozyme. It participates in the recognition of stem-loop I substrate via a tertiary loop-loop interaction and activates this substrate for catalysis. The stem-loop I / stem-loop V interaction involves the formation of Watson-Crick base pairs between the stem-loop I residues G630-C632, and the stem-loop V residues G697-C699. Formation of this loop-loop interaction is essential for optimal catalytic activity of the VS ribozyme. Our laboratory has initiated a structural study to better understand this RNA-RNA interaction, the effect of divalent metal ions, and

conformational changes involved in formation of this interaction. Here, we described the structure of SL5 RNA and started to investigate the effect of magnesium ion on this RNA.

We showed that the loop of stem-loop V forms a U-turn motif in the absence of divalent metal ions as previously proposed by Collins and coworkers (1). Comparison with structures of canonical U-turn motifs revealed that the SL5 loop lacks three of the seven structural characteristics commonly found in canonical U-turn motifs. Whereas canonical U-turn loop motifs are closed by a non-canonical base pair, the loop of SL5 is closed by a Watson-Crick U-A base pair. In addition, the characteristic stacking of the U base and R 5'-phosphate group and formation of a hydrogen bond between U H3 and R 3'-phosphate group are not found in the U-turn of SL5. We concluded that SL5 forms a loose U-turn motif by comparison with the more compact backbone fold of canonical U-turn motifs.

Why isn't SL5 forming a more compact fold with all the structural characteristics of the canonical U-turns? We propose that this could be due to 1) the intrinsic nature of the SL5 loop sequence and 2) the buffer conditions under which the structure was determined. As will be discussed below, it is possible that increased ionic stength or addition of divalent metal ions could help stabilize a more compact U-turn fold. However, canonical U-turn motifs can form in the absence of divalent metal ions and at relatively low ionic strength (19). Why not SL5? First, it is important to understand that the UNR sequence is a pre-requisite, and not a guaranty for formation of the U-turn motif. This point is illustrated by looking at two other RNA loops with similar UGA sequences (Figure 3.8a and 3.8b). In the NMR structure of the eukaryotic 16S-like rRNA (64), a UGAA tetraloop is flanked by a Watson-Crick G-C base pair (Figure 3.8a), and doesn't form a U-turn motif. Instead, in this $U_5G_6A_7A_8$ tetraloop, U5 and G6 are stacked on each other and on the 5'-strand, there is a turn in the backbone after G_6 and the two following A's (A7

and A8) are stacked on each other and on the 3'-strand (Figure 3.8a). In addition, U_5 and A_8 form a non-canonical ribose-base interaction (64). In the NMR structure of the 690 loop from the *E. coli* 16S rRNA (65), a UGAA tetraloop forms a canonical U-turn motif (Figure 3.8b). In this $G_{691}U_{692}G_{693}A_{694}A_{695}A_{696}$ loop, residues U692, G693, and A694 represent the UNR sequence, with the sharp turn in the backbone after U692 and stacking of residues G693, A694, and A695 after the turn. (Table 3.3 and Figure 3.8b). The loop is closed by a non-canonical sheared G-A base pair between nucleotides G691 and A696. How can we explain why this 690 loop foms a U-turn motif and the 16S-like rRNA UGAA tetraloop does not? Of course formation of a loop with reversal of strand orientation represents a topological problem by itself, which results in added constraints to the phosphodiester backbone. The size of the loop, stacking interactions within and outside the loop, the geometry of the closing-base pair, and the presence of modified residues or metal ions can all play a role in defining the most stable loop fold (64,66). In the case of the 690 loop and the 16S-like UGAA tetraloop, the loop sequence is the same, therefore the size of the loop is the same, and ionic conditions for structure determination were also comparable. This UGAA loop may not form a U-turn motif, most likely because the backbone is restricted by the Watson-Crick closing base pair and the adjacent helix, as well as by stacking interactions dictated by the closing base pair (64). We therefore propose that in stem-loop V, the UGACU pentaloop structure allows for formation of a U-turn motif, but that the backbone of this loop closed by a Watson-Crick base pair is too constrained to allow formation of a more compact backbone fold under the conditions used for structural determination (Figure 3.5).

The closing Watson-Crick base pairs found in the UGAA loop of 16S-like rRNA and in the VS ribozyme stem-loop V may hinder formation of the canonical U-turn motif, and this may explain why in general U-turn loop motifs are closed by non-canonical base pairs (20). We have

found only two examples in which a consensus UNR sequence forms a U-turn motif closed by a Watson-Crick base pair. In the 30S ribosomal subunit from *Thermus thermophilus* (*T. thermophilus*), the UAAC RNA loop of helix H26a is closed by a C-G Watson-Crick base pair and forms a U-turn motif (67) (not shown). In the 50S subunit from *Hamolarcula marismortui* (*H. marismortui*), the UCAC RNA loop of helix H25 is also closed by a C-G Watson-Crick base pair and also forms a U-turn motif (68) (Figure 3.8c). In this $C_{252}U_{253}C_{254}A_{255}C_{256}G_{257}$ loop, the UNR sequence corresponds to residues U253, C254, and A255, with the sharp turn in the backbone after U253 and stacking of residues C254, A255, and C256 after the turn. (Figure 3.8c). Similarly to the SL5 RNA loop, the UAAC loop of 30S helix H26a and the UCAC loop of 50S helix H25 form U-turn motifs, but fail to meet all the criteria of canonical U-turns (Table 3.3). The only two missing structural characteristics in these two later cases, is the absence of the non-canonical closing base pair and of a hydrogen-bonding interaction between U H3 and R 3'-phosphate group. The U253 N3 - A255 3'-P distance in the UCAC loop of 50S helix H25 is 6.79 Å and the corresponding distance in the UAAC loop of 30S helix H26a is 6.59 Å, more than 2 Å longer than what we measured for canonical U-turn motifs (Table 3.3). Furthermore, in both cases, the relative orientation of the hydrogen-bond donor and acceptor is not compatible with hydrogen-bond formation (Figure 3.8c). For SL5 loop, the U696 N3 - A698 3'-P distance and their relative orientation are also incompatible with hydrogen bonding (Figure 3.5 and Table 3.3). It is interesting to note that in Watson-Crick base pairs, the C1'-C1' distance of the paired residues is larger (10.4 Å for A-U base pairs and 10.7 Å for G-C base pairs) (69) than for many non-canonical base pairs, such as sheared G-A base pairs (8.47 Å for the sheared G691-A696 in Figure 3.8b). This shorter C1'-C1' distance of the closing base pair may better accomodate the backbone constraints of the U-turn motif. Conversely, the larger C1'-C1' distance of Watson-

Crick base pairs may be incompatible with formation of the hydrogen bond between the U H3 and R 3'-phosphate group, particularly for tetraloops when this hydrogen bond involves loop residues directly adjacent to the Watson-Crick closing base pair.

One important function of U-turns is to facilitate long-range tertiary interactions by exposing to the solvent the Watson-Crick faces of the bases located 3' after the turn. In stem-loop V, the three bases after the turn, G697, A698, and C699 are stacked on each other and have their Watson-Crick faces exposed to the solvent. In order to verify that the SL5 structure is compatible with formation of the loop-loop interaction between stem-loop I residues G630-C632 and stem-loop V residues G697-C699 (1), we superposed the heavy atoms of nucleotides G697-C699 with those of an equivalent trinucleotide forming an ideal A-form helix (Figure 3.9). We obtained an RMSD of 1.7 Å for this superposition. Although the loop I - loop V interaction does not necessarily form an ideal trinucleotide A-form helix, this superposition indicates that only small displacements of the A698 and C699 bases would be necessary for formation of the modeled interaction (Figure 3.9). These small base motions could be easily achievable, given the fast internal motion of the A698 and C699 bases (Figure 3.7) and the availability of space. It is also likely that magnesium ion binding brings the necessary conformational changes in stem-loop V, since magnesium ions stabilize the tertiary interaction with stem-loop I (13).

We used chemical shift mapping to analyze the effect of magnesium ions on the SL5 structure. Changes in chemical shifts were observed for most nuclei observed and particularly for those in the loop (Figure 3.6). The lack of confined chemical shift changes prevented us from localizing magnesium ions onto the NMR structure of SL5. Instead, results from the chemical shift mapping indicated that multiple magnesium ions are interacting with SL5 or that addition of magnesium ions brings a conformational change in the structure of SL5, or both. The largest

changes in chemical shifts were upfield shifts of C699H5 and C699H6 and their attached carbons. These changes in chemical shifts could result from changes in the electronic environment caused by a nearby positively charged magnesium ion (70). Indeed, phosphorothioate substitution interference and manganese rescue experiments have indicated that the nearby C699 5'-phosphate group is coordinated to a magnesium ion in the VS ribozyme (14). These large upfield shifts could also result from bringing H5/C5 and H6/C6 atoms of C699 closer to the center of the ring current of another base. For example, magnesium ion binding could lead to improved base stacking of C699 with A698 or formation of a stacking interaction between C699 and A701. The large upfield shift also observed for A701 H8 and C8 (Figure 3.7b) is consistent with magnesium ion-dependent formation of base stacking interactions between C699 and A701. These conformational changes could help formation of a more compact SL5 loop structure and stabilization of structural features of the canonical U-turn motif not present in the magnesium-ion free structure of SL5. Structure determination of SL5 in the presence of magnesium ion are currently underway to further understand the effects of magnesium ions on the structure of stem-loop V.

ACKNOWLEDGEMENTS

We would like to thank J. G. Omichinski for discussions, J. Brewer for equilibrium ultracentrifugation, J. Boisbouvier and J.-P. Simorre for the CT-TROSY pulse sequence, C. Hoogstraten for the HCCH-E.COSY pulse sequence, and A. Majumdar for the HNN-COSY pulse sequences. This work was supported by an NSF Career Award to P.L.

FIGURES and TABLES

Table 3.1. Structural statistics.

NOE Restraints	527
Number of NOE-derived distance restraints	498
From standard NOESY experiments	422
Inter-nucleotide	177
Intra-nucleotide	245
From 2D ^{1}H-^{15}N CMPG-NOESY	76
Hydrogen-bond restraints	29
Dihedral Restraints	26
Sugar Pucker (δ)	13
Backbone (γ)	13
Total Number of Restraints	553
RMSD from Experimental Restraints	
NOE (Å) (none > 0.1 Å)	0.0026 +/- 0.0003
Dihedral (˚) (none > 5˚)	0.027 + 0.015
RMSD from Idealized Geometry	
Bonds (Å)	0.003929 +/- 0.000018
Angles (˚)	0.957890 +/- 0.000907
Impropers (˚)	0.360240 +/- 0.000880
Heavy-Atom RMSDs to the Minimized Average Structure (Å)	
Overall (residues 691-705)	0.90 +/- 0.14
Stem (residues 691-695 and 701-705)	0.46 +/- 0.07
Loop (residues 696-700)	0.53 +/- 0.05

Table 3.2. Characteristics of canonical U-turn structures and their occurrence in SL5.

U-Turn Characteristics	Occurence in SL5
UNR sequence	Yes (U696, G697, A698)
Non-canonical flanking base pair	No (Flanking base pair is the Watson-Crick U696-A701 base pair)
Sharp turn in backbone	Yes (α angle in 10 lowest energy structures: 116 ± 7 Å)
Stacking of bases immediately after turn	Yes (stacking of G697, A698, and C699)
Stacking of U base and R 5'-phosphate group	No
H-bond between U 2'-OH and R N7	Yes (U696 O2' - A698 N7 distance range in 10 lowest energy structures: 2.91 - 3.93 Å)
H-bond between U H3 and R 3'-phosphate group	No (U696 N3 - A698 3'-P distance range in 10 lowest energy structures: 8.19 - 9.18 Å)

Table 3.3. Distances for the hydrogen bonds found in U-turns measured for SL5 and selected canonical U-turns from the PDB.

RNA	PDB Entry	U O2' - R N7 (Å)	U N3 - R 3'-P (Å)
SL5[a]		3.43	8.65
Loop 690 of eukaryotic 16 S-like rRNA[b] (65)	1FHK	2.75	4.42
HIV-1 A rich loop[b] (19)	1BVJ	2.86	4.21
U2 snRNA[a] (18)	2U2A	3.04	4.19
Hammerhead ribozyme[c] (71)	300D	3.07	4.39
TψC anticodon loop of $tRNA^{Phe}$ (yeast)[c] (72)	1EHZ	2.37	4.17

[a] Mimimized average structure
[b] Best representative conformer in the ensemble of NMR structures
[c] X-ray structure

Figure 3.1. Stem-loop V RNA within the VS ribozyme. **a)** Sequence and secondary structure of the *Neurospora* VS ribozyme. The cleavage site is indicated by the arrowhead. The interaction between stem-loops I and V is indicated by a dashed line and residues involved in this interaction are circled. Upon interaction with stem-loop V, stem-loop I (subdivided in Ia and Ib) undergoes a structural change from an inactive to an active conformation. **b)** Sequence and secondary structure of the stem-loop V (SL5) RNA used for NMR study. Wild-type and mutant nucleotides are represented by upper and lower cases, respectively.

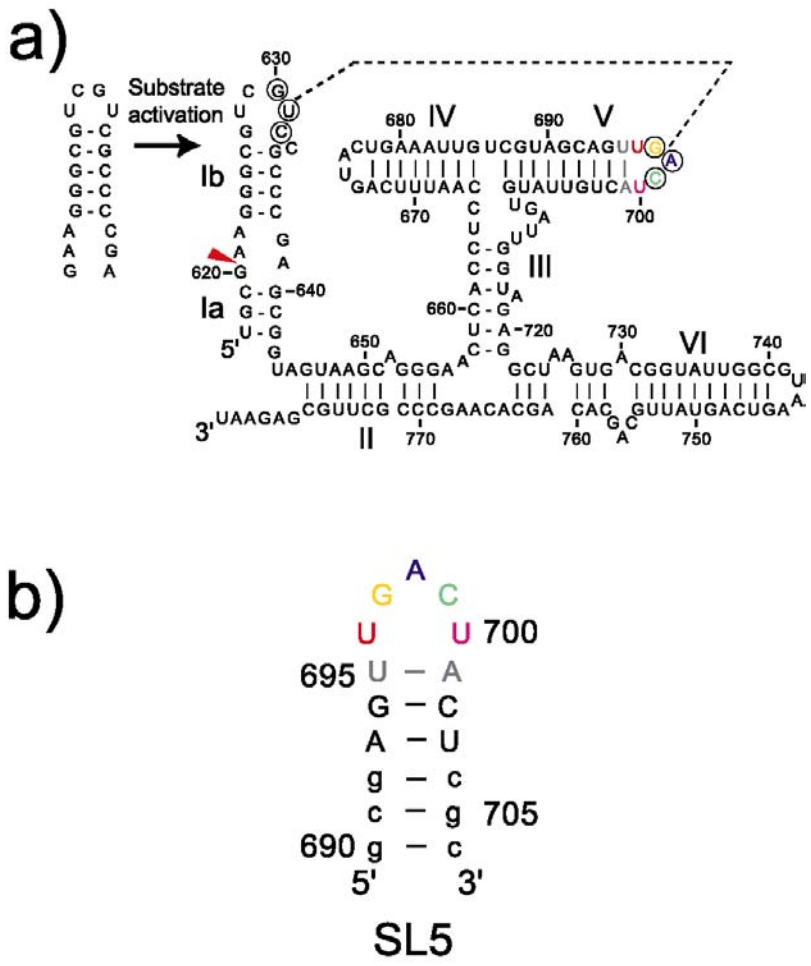


Figure 3.2. Watson-Crick base pairs in SL5. a) 2D spectrum of an HNN-COSY experiment used to detect $^{2}J_{NN}$ couplings across hydrogen bonds in Watson-Crick base pairs. The $^{2}J_{NN}$ coupling between the N1 of the purines and the N3 of the pyrimidines of paired residues is detected via the intervening imino proton (inset). The 1D imino proton spectrum is shown above the 2D HNN-COSY spectrum. b) 2D spectrum of a modified 2D HNN-COSY experiment used to detect $^{2}J_{NN}$ couplings across hydrogen bonds in Watson-Crick A-U base pairs. In this case, the $^{2}J_{NN}$ coupling between the N1 of the adenine and the N3 of the paired uridine is detected via the non-exchangeable adenine H2 (inset). The spectra in a) and b) were collected at 25 ˚C on a 600 MHz NMR spectrometer, with SL5 samples containing 90% D_2O:10% H_2O and 100% D_2O, respectively.

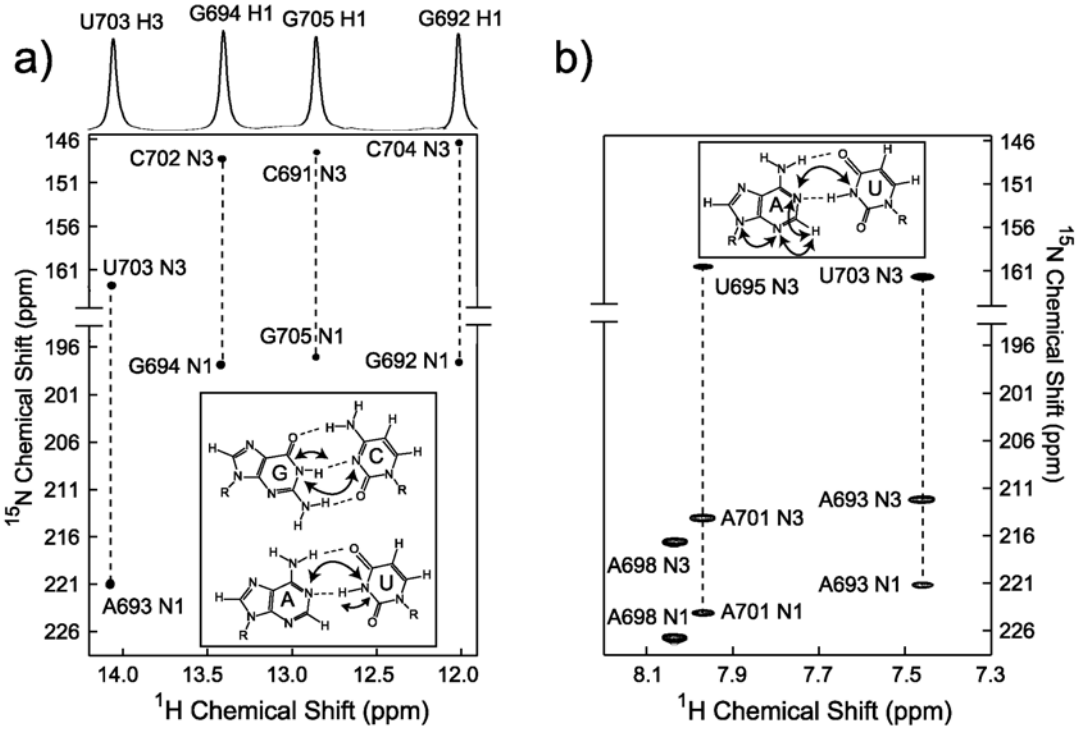


Figure 3.3. Regions from 3D ^{13}C-edited HMQC-NOESY spectra showing characteristics NOEs of the SL5 loop. **a)** Region of the spectrum collected with a 90 ms mixing time and taken at the ^{13}C frequency of U696 C2' (75.3 ppm). **b)** Region of the spectrum collected with a 180 ms mixing time and taken at the ^{13}C frequency of U696 C1' (93.6 ppm). The spectra in a) and b) were collected in 100% D_2O at 25 ˚C on a 600 MHz NMR spectrometer.

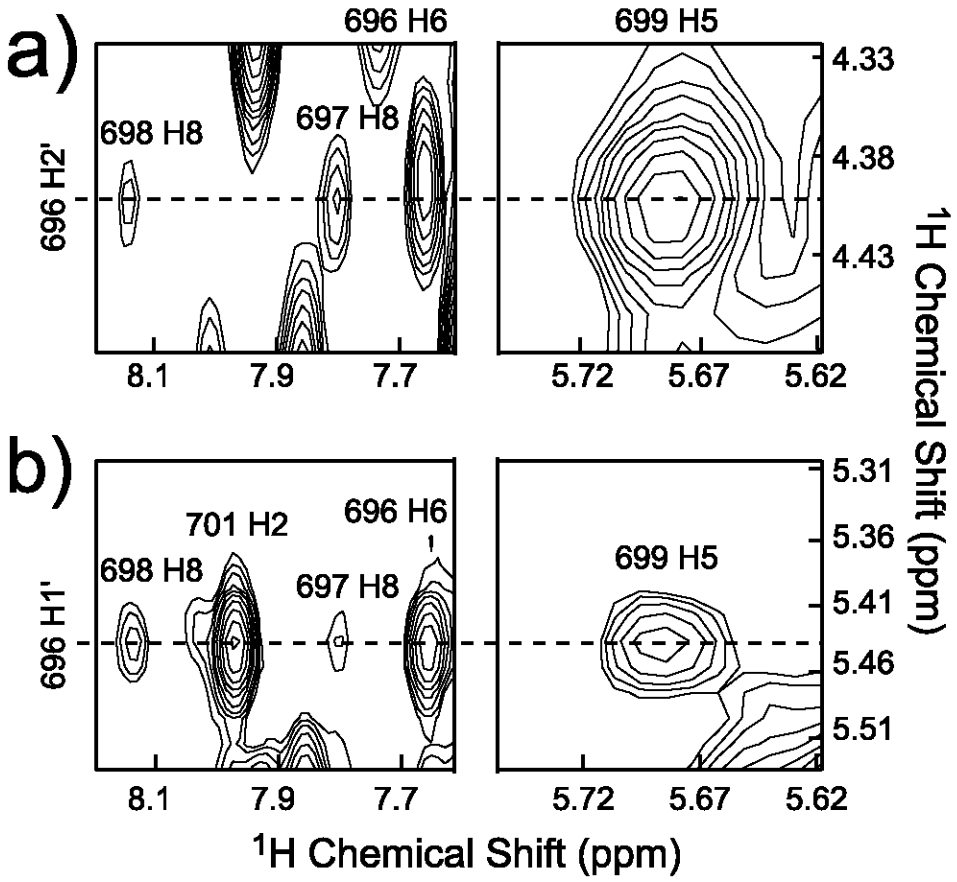


Figure 3.4. Summary diagram of the inter-residue NOEs used in the structure calculation of SL5. Dashed lines represent one or more inter-residue NOE(s) observed between two protons as defined in the legend.

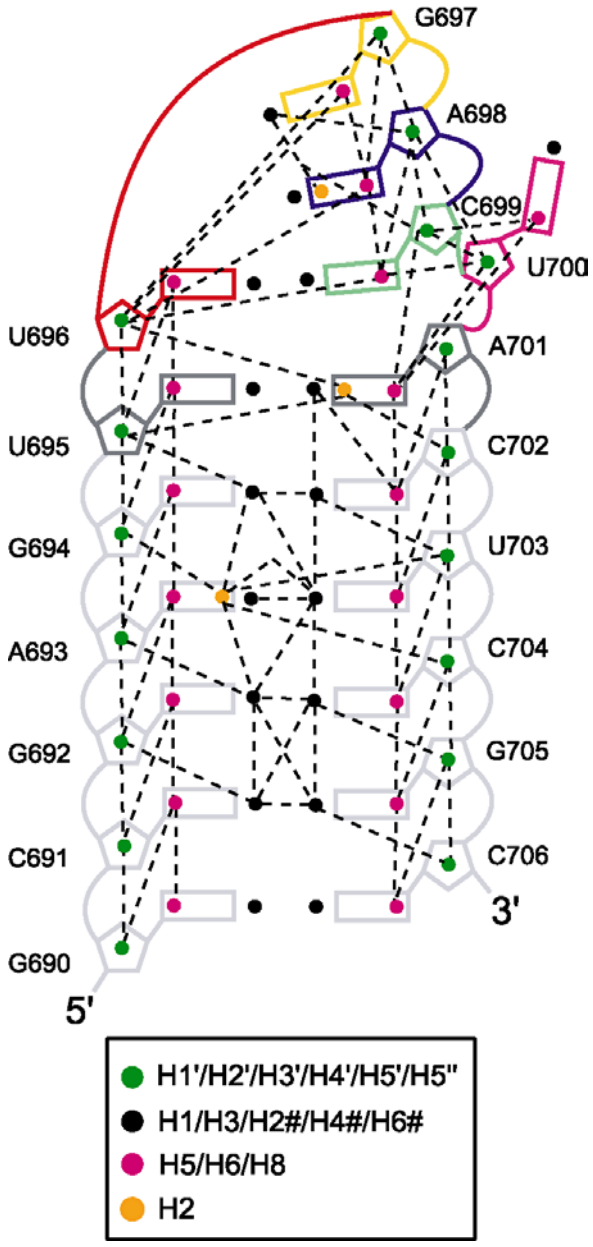


Figure 3.5. The loop of SL5 forms a U-turn motif. a) Stereo view of the superposition of the 10 lowest energy structures of SL5 on the minimized average structure. The superposition was obtained by minimization of pairwise heavy atom RMSD of each lowest energy structure to the minimized average structure (residues 691-705). b) Stereo view of the minimized average structure of the SL5 loop showing some characteristics of the U-turn motif: the sharp turn in the backbone between U696 and G697; the short distance (2.8 Å) between the U696 2'-OH and A698 N7 (dashed line); and the stacking of the three bases following the turn. For simplicity, only heavy atoms are shown in a) and b).

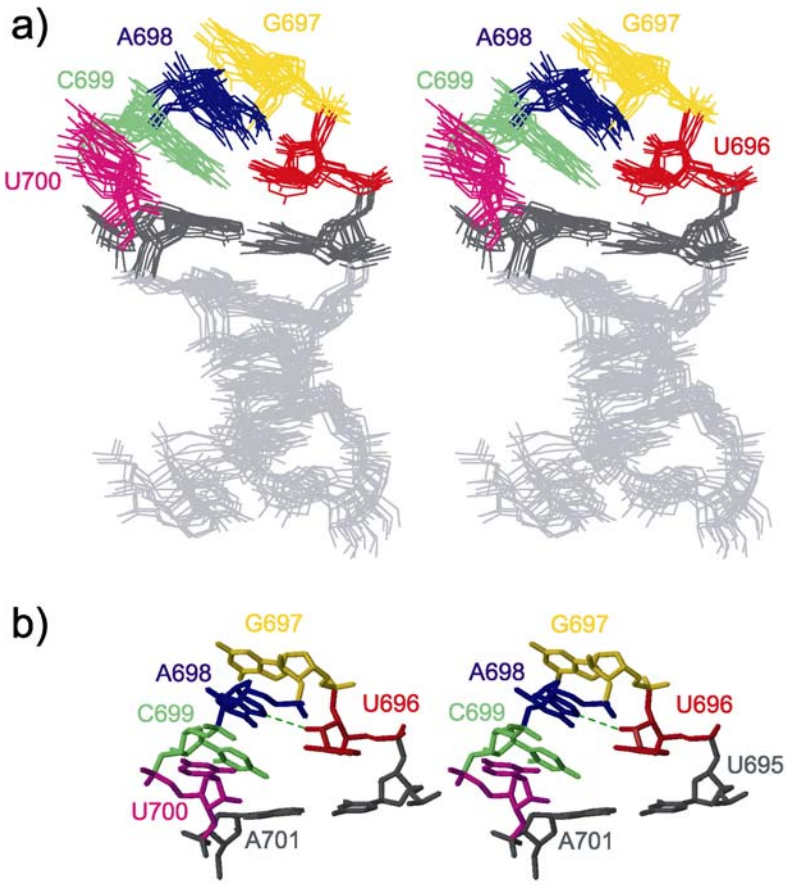


Figure 3.6. Effect of magnesium ions on SL5 RNA. a) Overlay of the H6-C6/H8-C8 region of ^{1}H-^{13}C CT-HSQC spectra of SL5 at selected points of the $MgCl_2$ titration. Spectra are shown for 0 mM (black), 0.5 mM (purple), 1.0 mM (green), 2.0 mM (blue), 3.0 mM (yellow), 5.0 mM (pale violet), 10 mM (pale blue), 30 mM (orange), 50 mM (grey), and 130 mM (red) $MgCl_2$. Peak labels are adjacent to the last point of the $MgCl_2$ titration. These spectra were collected in 100% D_2O at 25 ˚C on a 600 MHz NMR spectrometer. b) Chemical shift changes (Δ_T values) after addition of 130 mM $MgCl_2$ for the resolved peaks from the ^{1}H-^{13}C CT-HSQC spectra. The largest shifts occur in the loop, indicating that magnesium ion(s) may bind to the loop of SL5 (see text). c) Summary of the magnesium ion binding data on the minimized average structure of the SL5 loop. Significant chemical shift changes after addition of 130 mM $MgCl_2$ ($\Delta_T > 0.3$ ppm) were mapped as black sphere on the corresponding hydrogen atom (U696 H2'; G697 H1'; A698 H2', H3', H4', and H8; C699 H1', H5 and H6; U700 H1').

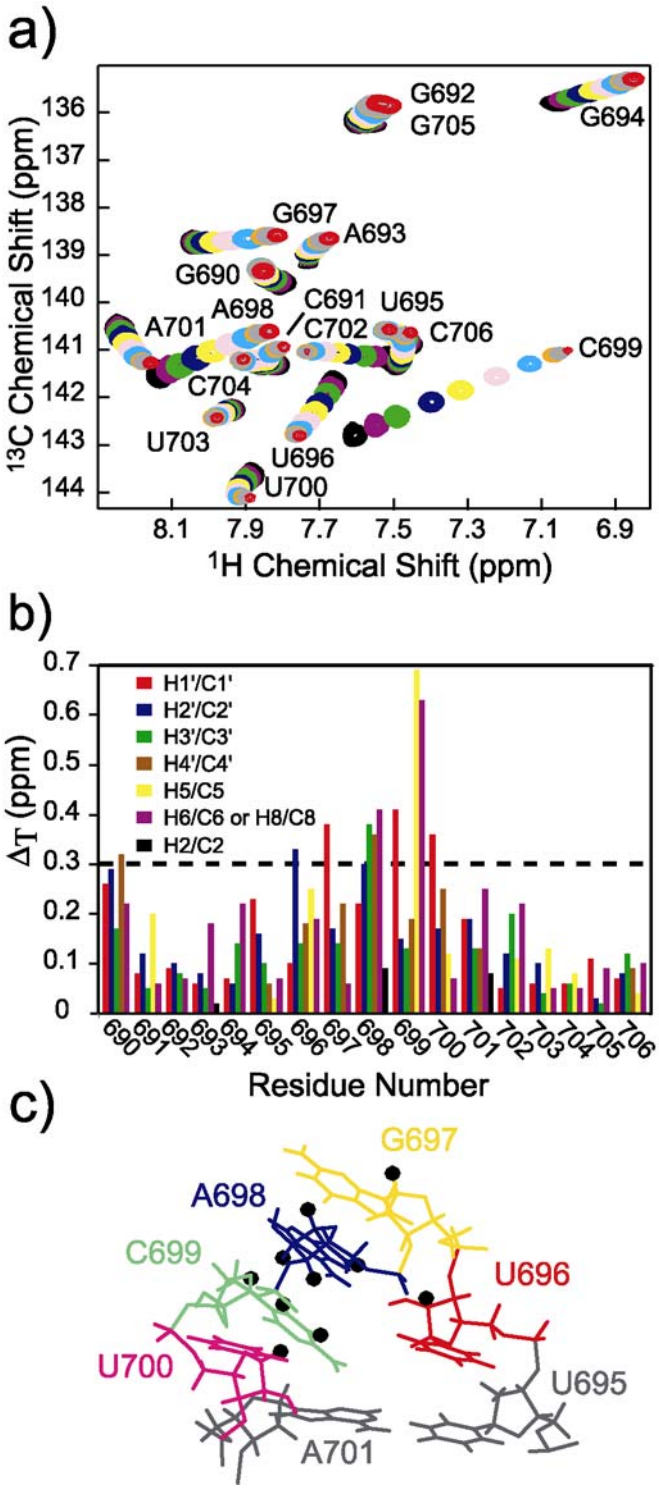

a)
G692
G705
G694
G697
A693
G690
A698
C691
U695
A701
C702
C706
C699
C704
U703
U696
U700
136
137
138
139
140
141
142
143
144
8.1
7.9
7.7
7.5
7.3
7.1
6.9
13C Chemical Shift (ppm)
1H Chemical Shift (ppm)
b)
H1'/C1'
H2'/C2'
H3'/C3'
H4'/C4'
H5/C5
H6/C6 or H8/C8
H2/C2
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
ΔT (ppm)
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
Residue Number
c)
G697
A698
C699
U696
U700
U695
A701

Figure 3.7. Histogram of $T_{1\rho}$ values for purine C8 and pyrimidine C6 base resonances in SL5 RNA. The values for G692 and G705 are not reported because of spectral overlap, which prevented accurate analysis. Open bars and filled bars indicate purine and pyrimidine residues, respectively. The horizontal dashed lines indicate the range of values observed in helical regions, excluding the terminal base pairs.

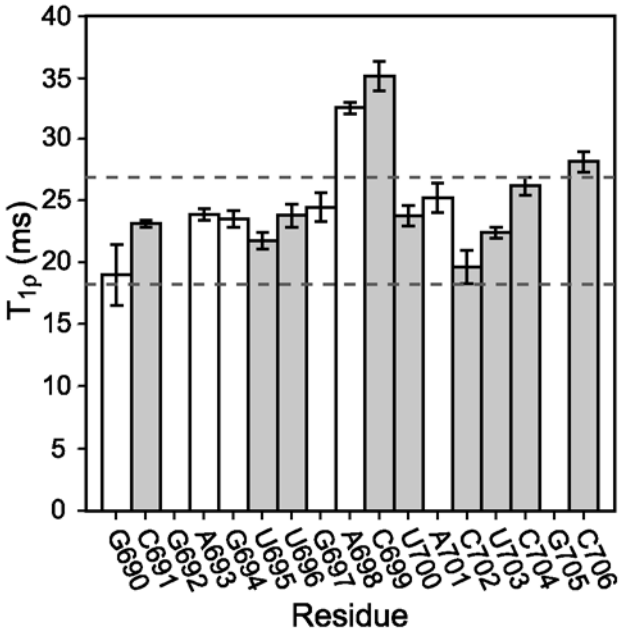


Figure 3.8. Other RNA loop structures with a UGA sequence. a) Representative NMR structure of the UGAA tetraloop from the eukaryotic 16 S-like rRNA (PDB entry 1AFX ((64)). This loop does not form a stable U-turn motif. b) Representative NMR structure of the 690 loop from the E. coli 16 S rRNA (PDB entry 1FHK (65)). This RNA loop forms a canonical U-turn motif with residues U692, G693, and A694 representing the UNR sequence. c) Terminal loop of helix 15 from the 2.4 Å X-ray structure of the large ribosomal subunit of *Haloarcula marismortui.* This RNA loop forms a U-turn motif with residues U253, C254, and A255 representing the UNR sequence. Like for SL5 RNA, this U-turn is closed by a Watson-Crick base pair (C252-G257) and does not share all the structural charateristics of canonical U-turn motifs.

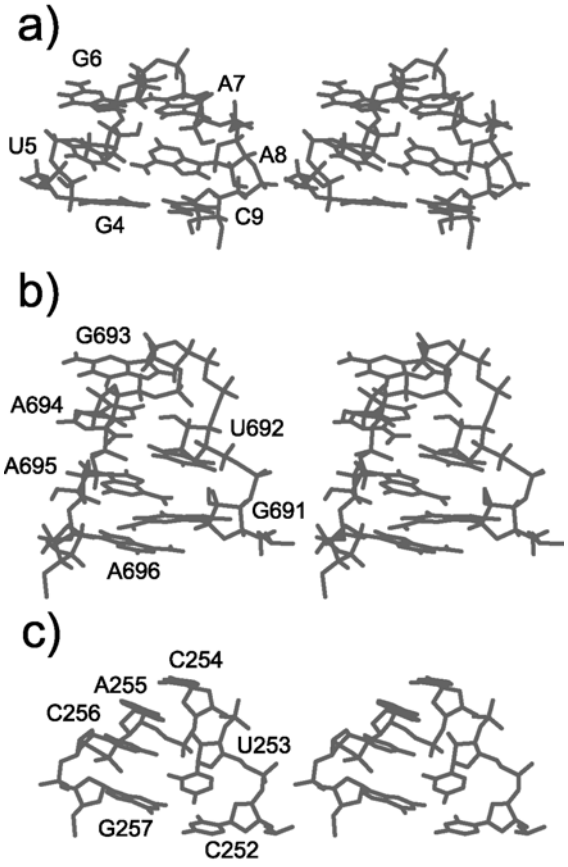
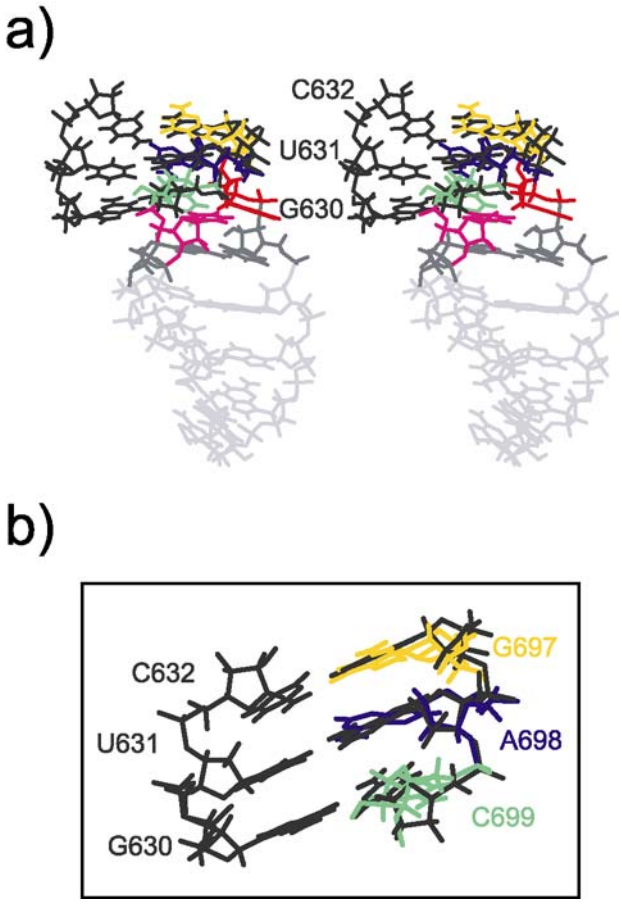


Figure 3.9. Model of the loop-loop interaction between stem-loop I and stem-loop V. a) Stereo view of the minimized average structure of SL5 (color coded as in Fig. 1b) superposed to the structure of a three base-paired A-form helix (in dark grey) formed between a 5'-GAC-3' and a 5'-CUG-3' sequence. The superposition was obtained by minimizing the heavy atom RMSD between the bases of the 5'-GAC-3' sequences (1.66 Å). b) Larger view of the modeled loop I - loop V interaction from the superposition shown in a).



REFERENCES

1. Rastogi, T., Beattie, T.L., Olive, J.E. and Collins, R.A. (1996) A long-range pseudoknot is required for activity of the *Neurospora* VS ribozyme. *EMBO J*, **15,** 2820-2825.

2. Sood, V.D. and Collins, R.A. (2001) Functional equivalence of the uridine turn and the hairpin as building blocks of tertiary structure in the *Neurospora* VS ribozyme. *J Mol Biol*, **313,** 1013-1019.

3. Andersen, A.A. and Collins, R.A. (2000) Rearrangement of a stable RNA secondary structure during VS ribozyme catalysis. *Mol Cell*, **5,** 469-478.

4. Pley, H.W., Flaherty, K.M. and McKay, D.B. (1994) Three-dimensional structure of a hammerhead ribozyme. *Nature*, **372,** 68-74.

5. 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.

6. Saville, B.J. and Collins, R.A. (1990) A site-specific self-cleavage reaction performed by a novel RNA in *Neurospora* mitochondria. *Cell*, **61,** 685-696.

7. Kennell, J.C., Saville, B.J., Mohr, S., Kuiper, M.T., Sabourin, J.R., Collins, R.A. and Lambowitz, A.M. (1995) The VS catalytic RNA replicates by reverse transcription as a satellite of a retroplasmid. *Genes Dev*, **9,** 294-303.

8. Saville, B.J. and Collins, R.A. (1991) RNA-mediated ligation of self-cleavage products of a *Neurospora* mitochondrial plasmid transcript. *Proc Natl Acad Sci U S A*, **88,** 8826-8830.

9. Guo, H.C., De Abreu, D.M., Tillier, E.R., Saville, B.J., Olive, J.E. and Collins, R.A. (1993) Nucleotide sequence requirements for self-cleavage of *Neurospora* VS RNA. *J Mol Biol*, **232,** 351-361.

10. Beattie, T.L., Olive, J.E. and Collins, R.A. (1995) A secondary-structure model for the self-cleaving region of *Neurospora* VS RNA. *Proc Natl Acad Sci U S A*, **92,** 4686-4690.

11. Guo, H.C. and Collins, R.A. (1995) Efficient trans-cleavage of a stem-loop RNA substrate by a ribozyme derived from *Neurospora* VS RNA. *EMBO J*, **14,** 368-376.

12. Hiley, S.L. and Collins, R.A. (2001) Rapid formation of a solvent-inaccessible core in the *Neurospora* Varkud satellite ribozyme. *EMBO J*, **20,** 5461-5469.

13. Andersen, A.A. and Collins, R.A. (2001) Intramolecular secondary structure rearrangement by the kissing interaction of the *Neurospora* VS ribozyme. *Proc Natl Acad Sci U S A*, **98,** 7730-7735.

14. Sood, V.D., Yekta, S. and Collins, R.A. (2002) The contribution of 2'-hydroxyls to the cleavage activity of the *Neurospora* VS ribozyme. *Nucleic Acids Res*, **30,** 1132-1138.

15. Jones, F.D. and Strobel, S.A. (2003) Ionization of a critical adenosine residue in the neurospora Varkud Satellite ribozyme active site. *Biochemistry*, **42,** 4265-4276.

16. Quigley, G.J. and Rich, A. (1976) Structural domains of transfer RNA molecules. *Science*, **194,** 796-806.

17. Lebars, I., Yoshizawa, S., Stenholm, A.R., Guittet, E., Douthwaite, S. and Fourmy, D. (2003) Structure of 23S rRNA hairpin 35 and its interaction with the tylosin-resistance methyltransferase RlmAII. *EMBO J*, **22,** 183-192.

18. Stallings, S.C. and Moore, P.B. (1997) The structure of an essential splicing element: stem loop IIa from yeast U2 snRNA. *Structure*, **5,** 1173-1185.

19. Puglisi, E.V. and Puglisi, J.D. (1998) HIV-1 A-rich RNA loop mimics the tRNA anticodon structure. *Nature Struct Biol*, **5,** 1033-1036.

20. Gutell, R.R., Cannone, J.J., Konings, D. and Gautheret, D. (2000) Predicting U-turns in ribosomal RNA with comparative sequence analysis. *J Mol Biol*, **300,** 791-803.

21. Sood, V.D., Beattie, T.L. and Collins, R.A. (1998) Identification of phosphate groups involved in metal binding and tertiary interactions in the core of the *Neurospora* VS ribozyme. *J Mol Biol*, **282,** 741-750.

22. Milligan, J.F., Groebe, D.R., Witherell, G.W. and Uhlenbeck, O.C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res*, **15,** 8783-8798.

23. Kao, C., Zheng, M. and Rudisser, S. (1999) A simple and efficient method to reduce nontemplated nucleotide addition at the 3' terminus of RNAs transcribed by T7 RNA polymerase. *RNA*, **5,** 1268-1272.

24. Butler, E.T. and Chamberlin, M.J. (1982) Bacteriophage SP6-specific RNA polymerase. I. Isolation and characterization of the enzyme. *J Biol Chem*, **257,** 5772-5778.

25. Nikonowicz, E.P., Sirr, A., Legault, P., Jucker, F.M., Baer, L.M. and Pardi, A. (1992) Preparation of ^{13}C and ^{15}N labelled RNAs for heteronuclear multi-dimensional NMR studies. *Nucleic Acids Res*, **20,** 4507-4513.

26. Vuister, G.W. and Bax, A. (1992) Resolution enhancement and spectral editing of uniformly ^{13}C-enriched proteins by homonuclear broadband ^{13}C decoupling. *J Magn Reson*, **98,** 428-435.

27. Santoro, J. and King, G.C. (1992) A constant-time 2D overbodenhausen experiment for inverse correlation of isotopically enriched species. *J Magn Reson*, **97,** 202-207.

28. Skleñár, V., Dieckmann, T., Butcher, S.E. and Feigon, J. (1998) Optimization of triple-resonance HCN experiments for application to larger RNA oligonucleotides. *J Magn Reson*, **130,** 119-124.

29. Pardi, A. and Nikonowicz, E.P. (1992) Simple procedure for resonance assignment of the sugar protons in ^{13}C-labeled RNAs. *J Am Chem Soc*, **114,** 9202-9203.

30. Ikura, M., Kay, L.E., Tschudin, R. and Bax, A. (1990) Three-dimensional NOESY-HMQC spectroscopy of a ^{13}C-labeled protein. *J Magn Reson*, **86,** 204-209.

31. Kay, L.E., Keifer, P. and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J Am Chem Soc*, **114,** 10663-10665.

32. Simorre, J.P., Zimmermann, G.R., Pardi, A., Farmer, B.T. and Mueller, L. (1995) Triple resonance HNCCCH experiments for correlating exchangeable and nonexchangeable cytidine and uridine base protons in RNA. *J Biomol NMR*, **6,** 427-432.

33. Simorre, J.P., Zimmermann, G.R., Mueller, L. and Pardi, A. (1996) Correlation of the guanosine exchangeable and nonexchangeable base protons in C-13-(IN)-N-15-labeled RNA with an HNC-TOCSY-CH experiment. *J Biomol NMR*, **7,** 153-156.

34. Simorre, J.P., Zimmermann, G.R., Mueller, L. and Pardi, A. (1996) Triple-resonance experiments for assignment of adenine base resonances in C-13/N-15-labeled RNA. *J Am Chem Soc*, **118,** 5316-5317.

35. Mueller, L., Legault, P. and Pardi, A. (1995) Improved RNA structure determination by detection of NOE contacts to exchange-broadened amino protons. *J Am Chem Soc*, **117,** 11043-11048.

36. Piotto, M., Saudek, V. and Skleñár, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR*, **2,** 661-665.

37. Grzesiek, S. and Bax, A. (1993) The importance of not saturating water in protein NMR. Application to sensitivity enhancement of NOE measurements. *J Am Chem Soc*, **115,** 12593-12594.

38. Legault, P. (1995) Ph.D. Thesis, University of Colorado at Boulder.

39. Skleñár, V., Miyashiro, H., Zon, G., Miles, H.T. and Bax, A. (1986) Assignment of the ^{31}P and ^{1}H resonances in oligonucleotides by two-dimensional NMR spectroscopy. *FEBS Lett*, **208,** 94-98.

40. Dingley, A.J. and Grzesiek, S. (1998) Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide (2)J(NN) couplings. *J Am Chem Soc*, **120,** 8293-8297.

41. 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.

42. Schwalbe, H., Marino, J.P., King, G.C., Wechselberger, P., Bermel, W. and Griesinger, C. (1994) Determination of a complete set of coupling constants in ^{13}C-labeled oligonucleotides. *J Biomol NMR*, **4,** 631-644.

43. Boisbouvier, J., Brutscher, B., Pardi, A., Marion, D. and Simorre, J.P. (2000) NMR determination of sugar puckers in nucleic acids from CSA-dipolar cross-correlated relaxation. *J Am Chem Soc*, **122,** 6779-6780.

44. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR*, **6,** 277-293.

45. Johnson, B.A. and Blevins, R.A. (1994) NMRView - A computer program for the visualization and analysis of NMR data. *J Biomol NMR*, **4,** 603-614.

46. Wishart, D.S., Bigam, C.G., Yao, J., Dyson, H.J., Oldfield, E., Markley, J.L. and Sykes, B.D. (1995) ^{1}H, ^{13}C, ^{15}N Chemical shift referencing in biomolecular NMR. *J Biomol NMR*, **6,** 135-140.

47. Roberts, G.C.K. (1993) *NMR of macromolecules : a practical approach*. IRL Press at Oxford University Press, Oxford England ; New York.

48. Brunger, A.T. (1992) *X-PLOR : version 3.1 : a system for x-ray crystallography and NMR*. Yale University Press, New Haven, Conn.

49. Schwieters, C.D., Kuszewski, J.J., Tjandra, N. and Clore, G.M. (2003) The Xplor-NIH NMR molecular structure determination package. *J Magn Reson*, **160,** 66-74.

50. Varani, G., Aboul-ela, F. and Allain, F.H.-T. (1996) NMR investigation of RNA structure. *Prog NMR Spectr*, **29,** 51-127.

51. Koradi, R., Billeter, M. and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graphics*, **14,** 51-55.

52. Lavery, R. and Sklenar, H.J. (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J Biomol Struct Dyn*, **6,** 63-91.

53. Farmer, B.T.I., Constantine, K.L., Goldfarb, V., Friedrichs, M.S., Wittekind, M., Yanchunas, J.J., Robertson, J.G. and Mueller, L. (1996) Localizing the NADP+ binding site on the MUrB enzyme by NMR. *Nature Struct Biol*, **3,** 995-997.

54. Gonzalez, R.L., Jr. and Tinoco, I., Jr. (1999) Solution structure and thermodynamics of a divalent metal ion binding site in an RNA pseudoknot. *J Mol Biol*, **289,** 1267-1282.

55. Maderia, M., Horton, T.E. and DeRose, V.J. (2000) Metal interactions with a GAAA RNA tetraloop characterized by (31)P NMR and phosphorothioate substitutions. *Biochemistry*, **39,** 8193-8200.

56. Yamazaki, T., Muhandiram, R. and Kay, L.E. (1994) NMR Experiments for the Measurement of Carbon Relaxation Properties in Highly Enriched, Uniformly 13C, 15N-Labeled Proteins: Application to 13C-alpha Carbons. *J Am Chem Soc* **116,** 8266-8278.

57. Korzhnev, D.M., Skrynnikov, N.R., Millet, O., Torchia, D.A. and Kay, L.E. (2002) An NMR experiment for the accurate measurement of heteronuclear spin-lock relaxation rates. *J Am Chem Soc*, **124,** 10743-10753.

58. Viles, J.H., Duggan, B.M., Zaborowski, E., Schwarzinger, S., Huntley, J.J., Kroon, G.J., Dyson, H.J. and Wright, P.E. (2001) Potential bias in NMR relaxation data introduced by peak intensity analysis and curve fitting methods. *J Biomol NMR*, **21,** 1-9.

59. Wijmenga, S.S. and van Buuren, B.N.M. (1998) The use of NMR methods for conformational studies of nucleic acids. *Prog NMR Spectr*, **32,** 287-387.

60. Ashraf, S.S., Ansari, G., Guenther, R., Sochacka, E., Malkiewicz, A. and Agris, P.F. (1999) The uridine in "U-turn": contributions to tRNA-ribosomal binding. *RNA*, **5,** 503-511.

61. Collins, R.A. and Olive, J.E. (1993) Reaction conditions and kinetics of self-cleavage of a ribozyme derived from *Neurospora* VS RNA. *Biochemistry*, **32,** 2795-2799.

62. Murray, J.B., Seyhan, A.A., Walter, N.G., Burke, J.M. and Scott, W.G. (1998) The hammerhead, hairpin and VS ribozymes are catalytically proficient in monovalent cations alone. *Chem Biol*, **5,** 587-595.

63. Hoogstraten, C.G., Wank, J.R. and Pardi, A. (2000) Active site dynamics in the lead-dependent ribozyme. *Biochemistry*, **39,** 9951-9958.

64. Butcher, S.E., Dieckmann, T. and Feigon, J. (1997) Solution structure of the conserved 16 S-like ribosomal RNA UGAA tetraloop. *J Mol Biol*, **268,** 348-358.

65. Morosyuk, S.V., Cunningham, P.R. and SantaLucia, J., Jr. (2001) Structure and function of the conserved 690 hairpin in Escherichia coli 16 S ribosomal RNA. II. NMR solution structure. *J Mol Biol*, **307,** 197-211.

66. Cabello-Villegas, J., Tworowska, I. and Nikonowicz, E.P. (2004) Metal ion stabilization of the U-turn of the A37 N6-dimethylallyl-modified anticodon stem-loop of Escherichia coli tRNA(Phe). *Biochemistry*, **43,** 55-66.

67. Carter, A.P., Clemons, W.M., Brodersen, D.E., Morgan-Warren, R.J., Wimberly, B.T. and Ramakrishnan, V. (2000) Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature*, **407,** 340-348.

68. 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 A resolution. *Science*, **289,** 905-920.

69. Saenger, W. (1984) *Principles of nucleic acid structure*. Springer-Verlag, New York.

70. Feigon, J., Butcher, S.E., Finger, L.D. and Hud, N.V. (2001) Solution nuclear magnetic resonance probing of cation binding sites on nucleic acids. *Methods Enzymol*, **338,** 400-420.

71. Scott, W.G., Murray, J.B., Arnold, J.R.P., Stoddard, B.L. and Klug, A. (1996) Capturing the structure of a catalytic RNA intermediate: The hammerhead ribozyme. *Science*, **274,** 2065-2069.

72. Shi, H.J. and Moore, P.B. (2000) The crystal structure of yeast phenylalanine tRNA at 1.93 angstrom resolution: A classic structure revisited. *RNA-A Publication of the RNA Society*, **6,** 1091-1105.

SUPPLEMENTARY MATERIAL

Figure 3.S1. Effect of NaCl on SL5 RNA. **a)** Overlay of the H6-C6/H8-C8 region of 2D ^{1}H-^{13}C CT-HSQC spectra of SL5 at the various points of the NaCl titration. Spectra are shown for 50 mM (black), 100 mM (red), 150 mM (green), 250 mM (blue), 350 mM (yellow), and 500 mM (pink). Peak labels are adjacent to the last point of the NaCl titration. As a reference, the last point of the $MgCl_2$ titration (Figure 3.6a; 130 mM $MgCl_2$) is also shown (pale blue). These spectra were collected in 100% D_2O at 25 ˚C on a 600 MHz NMR spectrometer. **b)** Chemical shift changes (Δ_T values) after addition of 130 mM $MgCl_2$ (pale blue) and 450 mM NaCl (pink; 500 mM NaCl total) for the resolved peaks of the H6-C6/H8-C8 region of the ^{1}H-^{13}C CT-HSQC spectrum.

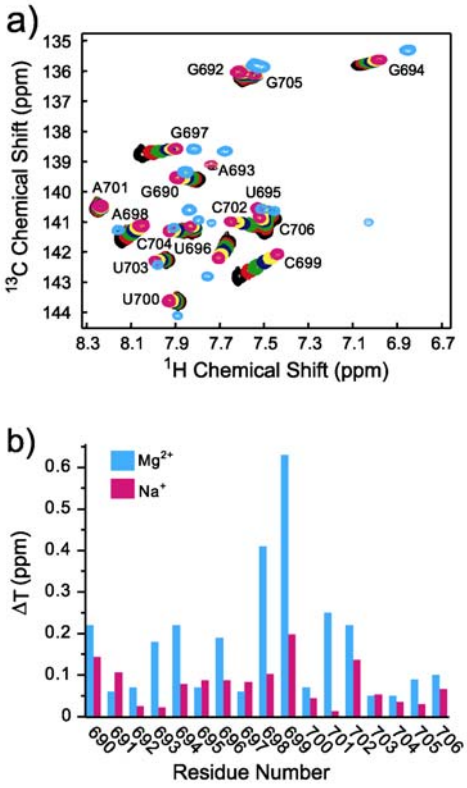


Table 3.S1. Resonance assignment of SL5.

NON-EXCHANGEABLE PROTON CHEMICAL SHIFTS (ppm) (± 0.01 ppm)										
	H1'	**H2'**	**H3'**	**H4'**	**H5'**	**H5''**	**H2**	**H5**	**H6**	**H8**
G690	5.70	4.75	4.53	4.31	4.00	3.88				8.00
C691	5.64	4.65	4.59	4.46	4.55	4.18		5.27	7.81	
G692	5.74	4.63	4.62	4.51	4.50	4.15				7.55
A693	5.93	4.68	4.64	4.50	4.58	4.14	7.46			7.72
G694	5.59	4.40	4.32	4.44	4.43	4.02				7.06
U695	5.58	4.38	4.33	4.39	4.46	4.05		5.09	7.45	
U696	5.47	4.40	4.50	4.34	4.32	4.07		5.60	7.66	
G697	5.43	4.62	4.60	4.26	4.19	4.02				7.80
A698	5.91	4.58	4.78	4.42	4.24	4.02	8.04			8.14
C699	5.85	4.31	4.63	4.40	4.31	4.12		5.69	7.62	
U700	5.79	4.54	4.56	4.51	4.29	4.23		5.84	7.88	
A701	5.92	4.73	4.63	4.54	4.47	4.23	7.97			8.24
C702	5.45	4.32	4.36	4.46	4.48	4.10		5.31	7.51	
U703	5.58	4.51	4.56	4.44	4.58	4.11		5.39	7.94	
C704	5.56	4.48	4.54	4.46	4.54	4.13		5.66	7.86	
G705	5.68	4.46	4.57	4.43	4.53	4.09				7.59
C706	5.70	3.98	4.12	4.17	4.50	4.04		5.23	7.49	
CARBON and PHOSPHORUS CHEMICAL SHIFTS (± 0.2 ppm for carbon chemical shifts)										
	C1'	**C2'**	**C3'**	**C4'**	**C5'**	**C2**	**C5**	**C6**	**C8**	**5'-P[b]**
G690	93.0	75.0	74.0	84.8	62.3				138.7	
C691	93.8	75.5	72.5	82.1	65.0		97.3	141.3		-0.52
G692	92.9	75.5	73.1	82.1	65.9				136.2	0.11
A693	92.8	75.8	72.7	82.0	65.0	153.3			139.2	-0.18
G694	92.8	75.5	72.9	82.2	65.8				135.8	-0.11
U695	93.2	75.8	72.9	83.2	64.9		103.6	140.9		-0.64
U696	93.6	75.4	73.5	83.2	65.6		104.6	141.6		-0.15
G697	91.1	75.0	75.4	84.1	66.3				139.6	0.37
A698	91.0	76.5	76.0	84.7	66.8	155.2			141.6	0.05
C699	91.7	76.2	76.2	84.3	67.0		98.7	142.8		-0.39
U700	92.0	75.3	76.0	84.5	67.8		104.9	143.6		0.11
A701	92.5	75.4	73.8	82.9	66.1	154.1			140.5	-0.05
C702	94.3	75.5	72.7	82.3	65.6		97.5	141.2		-0.38
U703	93.8	75.4	72.3	82.2	64.4		103.2	97.9		-0.76
C704	93.8	75.4	72.6	81.9	65.0		97.9	141.3		-0.37
G705	93.0	75.4	72.5	82.1	65.1				136.2	-0.26
C706	92.8	77.5	69.5	83.3	65.1		97.5	141.2		-0.51

BASE EXCHANGEABLE PROTON CHEMICAL SHIFTS (ppm)								
	H1[a]	**H3[a]**	**H41[b]**	**H42[b]**	**H21[a]**	**H22[a]**	**H61[a]**	**H62[a]**
G690								
C691			8.49	6.59				
G692	12.04							
A693								
G694	13.47				8.51	6.29		
U695								
U696								
G697					6.19	6.19		
A698							6.56	6.56
C699			6.79	6.79				
U700								
A701							7.67	6.65
C702			8.47	6.94				
U703		14.13						
C704			8.29	6.77				
G705	12.90							
C706			8.28	6.85				
BASE NITROGEN CHEMICAL SHIFTS (ppm)								
	N1[c]	**N2[d]**	**N3[e]**	**N4[d]**	**N6[d]**	**N7[e]**	**N9[e]**	
G690						231.9	169.2	
C691	150.9		147.5	97.8				
G692	146.0	73.9				235.3	169.1	
A693	221.3		212.3		83.6	229.8	170.5	
G694	148.1	74.5				234.9	169.2	
U695	145.8		160.6					
U696	145.6		157.1					
G697		72.8				235.3	168.9	
A698	226.8		216.7		80.3	230.3	169.3	
C699	151.1							
U700	144.9		156.7					
A701	224.3		214.2		82.6	230.1	170.5	
C702	151.1		148.3	98.0				
U703	146.7		162.7					
C704	150.6		146.4	97.5				
G705	147.2	74.3				234.9	168.9	
C706	152.3			97.5				

a. ± 0.02 ppm; b. ± 0.04 ppm; c. ± 0.4 ppm for purines, and ± 0.1 ppm for pyrimidines; d. ± 0.2 ppm ; e. ± 0.4 ppm.

CHAPTER 4

NMR STRUCTURE OF THE VS RIBOZYME STEM-LOOP V IN THE PRESENCE OF MAGNESIUM IONS AND LOCALIZATION OF DIVALENT METAL ION BINDING SITES[1]

[1] Dean O. Campbell and Pascale Legault. To be submitted to *Journal of Molecular Biology*.

ABSTRACT

The terminal loop of stem-loop V of the *Neurospora* VS ribozyme is involved in a loop-loop interaction that is important for the catalytic activity of the ribozyme. Mutagenesis and site-specific substitution have suggested that this interaction is facilitated by the presence of a U-turn in the loop of stem-loop V. The solution structure of the terminal loop of stem-loop V (SL5) in the absence of magnesium was solved by NMR spectroscopy. The loop of SL5 contains a loose U-turn motif that did not contain all the characteristics of the canonical U-turn motif, including one of the two characteristic hydrogen bonds. Here, we present the high-resolution solution NMR structure of SL5 in the presence of magnesium ions and show that magnesium ions induce a conformational change in the loop of stem-loop V. This conformational change results in the U-turn motif of SL5 having more characteristics of a canonical U-turn motif in the presence than in the absence of magnesium ions. We used paramagnetic line broadening from manganese ions to localize four divalent metal ion binding sites in the loop of stem-loop V. Three of these divalent metal ion binding sites are specific to the U-turn motifs and have been found in structures of other U-turn motifs.

INTRODUCTION

The *Neurospora* VS ribozyme is a small catalytic RNA that is dependent on metal ions for its folding and activity (1). The primary activity of the ribozyme is autolytic cleavage of the VS RNA to produce 5'-OH and 2'3'-cyclic phosphate termini similarly to other small catalytic RNAs such as the hammerhead, hairpin and hepatitis delta virus (2-4). The VS ribozyme is also capable of *trans* cleavage and under the right conditions can perform ligation reactions (3,5,6). The secondary structure of the VS ribozyme is shown in Figure 4.1a. The catalytic domain of the

ribozyme is represented by stem-loops II-VI, whereas stem-loop I represents the substrate domain. Cleavage occurs in the internal loop found in stem-loop I. The active site is thought to be in the A730 loop of stem-loop VI (7,8). It has been suggested that in order for cleavage to occur, stem-loop I docks into a cleft formed between stem-loops II and VI (9).

The substrate recognition of the VS ribozyme involves tertiary interactions that form between the substrate and catalytic domains (10). A critical step in the substrate recognition of the VS ribozyme is the formation of an important loop-loop interaction between residues 630-632 in stem-loop I and residue 697-699 in stem-loop V (11). This interaction, which depends on the presence of divalent metal ions, leads to a conformational change in the structure of stem-loop I (Figure 4.1a). Such a conformational change in stem-loop I can be induced by the addition of stem-loop V and magnesium ions only (12). The conformational change is characterized by a shift in the 3' strand of stem-loop I, in which nucleotides 623-625 in stem-loop I shift their base-paring partners from nucleotides 634-636 to nucleotides 635-637 (Figure 4.1a), extruding C634 from the stem and leading to a structural rearrangement in the active site internal loop. It has been demonstrated that this conformational change is essential for activation of the stem-loop I RNA (11). Indeed, stem-loop I RNA mutants that can not undergo this conformational change are not cleaved by the VS ribozyme, whereas stem-loop I mutant substrates that can adopt the shifted conformation are active in the cleavage reaction. NMR structures of both the unshifted (inactive) and shifted (active) conformations have been determined by NMR spectroscopy (13,14). These structures indicate that the conformational change in stem-loop I leads to the formation of a metal ion binding site and a tertiary interaction motif in the cleavage site internal loop of the shifted conformation (14), both of which may be important for catalysis.

The catalytic activity of the VS ribozyme has a strict requirement for the presence of metal ions for activity with magnesium being the most preferred divalent metal ion (1). Other ions such as calcium, manganese and certain monovalent ions support catalysis, whereas cobalt hexammine supports the correct folding of the ribozyme, but not catalysis (1,3,15,16). Phosphorothioate substitution interference and manganese rescue experiments have been used to define sites of metal ion interaction (Figure 4.1a) (17). Using these methods, metal ion coordination sites were found in the terminal loops of stem-loops I and V, which involve the 5'-phosphate of U631 and C699, two nucleotides that directly participate in the magnesium-ion dependent loop I - loop V interaction (Figure 4.1a). There are two other sites of metal ion coordination in the A730 loop, which is the loop that contains the proposed active site (Figure 4.1a) (7,17).

The terminal loop of stem-loop V was postulated to contain a U-turn motif based on its sequence, as well as mutagenesis and site-specific substitution experiments that disrupted its ability to form a U-turn (18,19). U-turn motifs were first identified in the anticodon and TψC-loop of tRNA (20). They have been identified in several other RNAs such as the hammerhead, HIV, 23S rRNA, and the U2 snRNA (21-24). One important function of the U-turn motif is to make the bases immediately following the turn available for tertiary interactions, which may be intramolecular or intermolecular (25,26).

Residues U696-C699 of stem-loop V form a U-turn consensus sequence and substitution of U696 with any other nucleotides reduced catalytic activity (18). A characteristic interaction of U-turn motifs is the formation of a hydrogen bond between the U 2'-OH and the R N7. Substitution of the 2'-OH group by 2'H of residue U696 reduces catalytic activity of the VS ribozyme (19). A three-dimensional structure of stem-loop V in the absence of magnesium ions,

$SL5^{free}$, was determined by NMR spectroscopy (27). It was found that loop residues 695-701 form a loose U-turn motif (27). Indeed, the U-turn motif of $SL5^{free}$ does not meet all the criteria of canonical U-turn motifs (25,27,28). Whereas canonical U-turn loop motifs are closed by a non-canonical base pair, the loop of SL5 is closed by formation of a Watson-Crick U-A base pair. In addition, the characteristic stacking of the U base and R 5'-phosphate group and formation of a hydrogen bond between the H3 of residue U (U696 H3) and 3'-phosphate group of residue R (A698 3'P) are not found in the U-turn of SL5. Nevertheless, the three bases following the turn (G697, A698, and C699) are all stacked on each other and their hydrogen bonding face is oriented to allow base-pairing interactions with the complementary residues on stem-loop I (27).

We have previously titrated magnesium ions to $SL5^{free}$ in order to understand the effect of magnesium ions on the SL5 structure. Addition of saturating amount of Mg^{2+} affected most NMR-active nuclei in SL5, with the largest changes being in the loop region. We have verified that these chemical changes were not simply a result of increased ionic strength. However, since the effect of magnesium ions are not localized to a few atoms, it was not possible to map the magnesium ion binding sites from these changes in chemical shifts. The large changes observed also indicated that addition of magnesium ions trigger a conformational change in the SL5 loop.

In this report, we used NMR spectroscopy to further investigate the effect of Mg^{2+} on the solution structure of SL5. We present the three-dimensional structure of SL5 in the presence of magnesium ions, $SL5^{Mg}$, and show that the U-turn motif of $SL5^{Mg}$ forms a compact U-turn motif with all the characteristics of canonical U-turn motifs, except for its Watson-Crick closing base pair. NMR methods were also employed to localize divalent metal ion binding sites in $SL5^{free}$. Cobalt hexammine was used as a probe for magnesium ion binding by localizing the cobalt

hexammine binding sites through intermolecular NOEs. Manganese ions were also used as a probe for magnesium ion binding, through the paramagnetic line broadening of resonances that are close to the manganese ions. Structure modeling was used to position manganese ions in the loop of $SL5^{Mg}$.

MATERIALS AND METHODS

Sample Preparation

Unlabeled, ^{15}N-labeled, and $^{13}C/^{15}N$-labeled SL5 RNA was synthesized and purified as previously described (27).The prepared RNA samples ranged in concentration from 0.8 to 2.0 mM as determined from absorbance at 260 nm using a coefficient of 40 µg/ml per OD unit. The sample was exchanged in either NMR buffer A (10 mM d_{11}-Tris, 50 mM NaCl, 0.05 mM NaN_3, pH 7.0) or NMR buffer B (10 mM d_{11}-Tris, 0.05 mM NaN_3, pH 7.0), which contained 90% H_2O and 10% D_2O. For studies in D_2O, the RNA samples were transferred to 99.996% D_2O through multiple cycles of lyophilisation and resuspension in D_2O. Before addition of any metal ions, the RNA sample was heated to 95 ºC for two minutes and then immediately snap-cooled in iced water. The magnesium-ion bound form of SL5, $SL5^{Mg}$, was prepared by the addition of 99.995% $MgCl_2$ (Sigma-Aldrich, MO) to the NMR samples: 40 mM $MgCl_2$ to the RNA sample in NMR buffer A and either 20 mM or 46 mM $MgCl_2$ to a RNA sample in NMR buffer B.

NMR Spectroscopy

All NMR experiments were conducted at 25 ˚C on a Varian Inova 600 MHz spectrometer equipped with a pulse-field gradient unit and an actively shielded z gradient probe, either a $^1H\{^{13}C/^{15}N\}$ triple resonance probe or a $^1H\{^{15}N\text{-}^{31}P\}$ indirect detection probe. Assignment of the

^{1}H, ^{13}C, ^{15}N and ^{31}P nuclei in $SL5^{Mg}$ were obtained by first starting with an assigned spectrum of $SL5^{free}$ and then tracking the peaks during a $MgCl_2$ titration. Resolved peaks and confirmation of the assignments were obtained from two-dimensional (2D) and three-dimensional (3D) homonuclear and heteronuclear NMR experiments. All non-exchangeable protons and their attached carbons were assigned from the following experiments collected at 25 ºC in D_2O on a SL5 sample in NMR buffer A supplemented with 40 mM $MgCl_2$: 2D ^{1}H-^{13}C CT-HSQC (29,30); 3D HCCH-COSY (31); 3D HCCH-TOCSY (31); and 3D ^{13}C-edited HMQC-NOESY (32). The exchangeable protons and the nitrogens were assigned at 25 ºC in H_2O on a SL5 sample in NMR buffer B with 20 mM $MgCl_2$ using the following experiments: imino- and amino-optimized 2D ^{1}H-^{15}N HSQC (33) and long-range ^{1}H-^{15}N HSQC optimized for detection of adenine N1 and N3 and purine N7 and N9 (34). A 2D ^{1}H-^{31}P HETCOR (35) spectrum collected in D_2O at 25 ºC on a $SL5^{Mg}$ sample in NMR buffer A supplemented with 40 mM $MgCl_2$ was used for the assignment of the ^{31}P resonances. Distance restraints were obtained from a 2D ^{1}H-^{15}N CPMG-NOESY (36) spectrum collected at 25 ˚C in H_2O with a mixing time of 150 ms and a 3D ^{13}C-edited HMQC-NOESY (32) spectra collected in D_2O at 25 ˚C with mixing times of 75 ms and 150 ms. The 2D ^{1}H-^{15}N CPMG-NOESY and the 3D ^{13}C-edited HMQC-NOESY spectra were collected on a $SL5^{Mg}$ sample in NMR buffer A supplemented with 40 mM $MgCl_2$. Hydrogen-bonding restraints were obtained from 2D HNN-COSY spectra (37,38),collected on two $SL5^{Mg}$ samples. One HNN-COSY spectrum, which detect $^{2}J_{NN}$ couplings via the non-exchangeable H2, was collected on a sample in NMR buffer B supplemented with 20 mM $MgCl_2$ and was used for defining the U696 N3 to A700 N1 distance restraints (38). The other HNN-COSY spectrum, which detects $^{2}J_{NN}$ couplings via the exchangeable imino proton, was collected on a sample in NMR buffer A supplemented with 40 mM $MgCl_2$ and was used to define all other distance restraints from this

dataset. Restraints for the δ torsion angle were obtained from a 3D HCCH E.COSY (39) collected at 25 ºC in D_2O on a $SL5^{Mg}$ sample in NMR buffer A supplemented with 40 mM $MgCl_2$. All spectra were processed with the NMRPipe/NMRDraw package (40) and analyzed with NMRView (41). ^{1}H, ^{13}C and ^{15}N chemical shifts were referenced at 25 ˚C to an external standard of DSS at 0.00 ppm (42) and ^{31}P chemical shifts were referenced at 25 ºC to an external standard of 85% phosphoric acid at 0.00 ppm.

Structure Calculation

The distance restraints from the 3D ^{13}C-edited HMQC-NOESY spectra were separated into four ranges: strong (1.8-3.0 Å), medium (1.8-4.1 Å), and weak (1.8-5.5 Å) based on the intensities of peaks observed at a mixing time of 75 ms; and very weak (1.8-7.0 Å) for signals observed only at the 150 ms mixing time. The distance restraints obtained from the 2D ^{1}H-^{15}N CPMG-NOESY were given ranges of either 1.8-5.5 Å or 1.8-7.0 Å based on crosspeak intensities. The distance restraints were calibrated based on the crosspeaks intensities and by comparison with NOE crosspeaks involving protons separated by known intra-residue distances. 2D HNN-COSY spectra (37,38) were collected to detect $^{2}J_{NN}$ couplings across hydrogen bonds in Watson-Crick base pairs. Because of the strong NMR evidence (from NOESY and HNN-COSY spectra) for the formation of the A-U and G-C Watson-Crick base pairs in the stem of $SL5^{Mg}$, canonical distance restraints were employed to define the hydrogen-bonding pattern and planarity of the first five base pairs in the stem of $SL5^{Mg}$ (residues 690-694 and 702-706). We obtained limited NOE data for the U695-A701 base pair, therefore, in this case, only the U695 N3 to A701 N1 distance was defined (2.82 Å +/- 0.1 Å) based on the HNN-COSY data. Sugar pucker conformations were derived from a 3D HCCH E.COSY (39) experiment. All the sugar

puckers except for the loop residues 696-700 were set to C3'-endo (δ = 86 +/- 10˚). For residues U696-U700, the $^{3}J_{H1'-H2'}$ values derived from the HCCH E.COSY represent the average values for C3'-endo and C2'-endo conformers, indicating that their ribose puckers are in equilibrium between these two conformations. The γ torsion angle restraints were derived from comparative analyses of NOE data (43). The γ angles for all the residues, except the loop residues 696-700, were set to the gauche+ conformation (γ = 60 +/- 20˚).

Three-dimensional structures were calculated with restrained molecular dynamics and simulated annealing in X-PLOR-NIH 2.0.6 (44,45) using a previously described protocol (27). Starting from a set of 75 linear structures with randomized torsion angles, 69 structures satisfied the experimental restraints (no distance violation > 0.1 Å and no torsion angle violation > 5˚). From these 69 structures, the 10 lowest energy structures were selected for further analysis. These 10 structures were used to calculate an average structure that was further minimized against experimental restraints.

Metal-Ion Binding Studies

A 1.0 mM sample of $^{13}C/^{15}N$-labeled SL5 in NMR buffer B (10 mM d_{11}-Tris, 0.02% NaN_3, pH 7.0) was titrated with $MgCl_2$. The titration was carried out by directly adding varying amounts of concentrated 99.995% $MgCl_2$.(Sigma-Aldrich, MO) to the RNA sample. The divalent metal ion concentration was adjusted from 0 mM to 0.25 mM, 0.50 mM, 0.75 mM, 1.0 mM, 2 mM, 3 mM, 5 mM, 7.5 mM, 15 mM, 36 mM, and 46 mM. Chemical shift changes were monitored after each salt addition by collecting 2D $^{1}H-^{13}C$ CT-HSQC spectra (30).

A 1.0 mM sample of $^{13}C/^{15}N$-labeled SL5 in NMR buffer A (10 mM d_{11}-Tris, 50 mM NaCl, 0.02% NaN_3, pH 7.0) was titrated with $Co(NH_3)_6Cl_3$. The titration was carried out by

directly adding varying amounts of concentrated $Co(NH_3)_6Cl_3$ (Sigma-Aldrich, MO) to the RNA sample. The trivalent metal ion concentration was adjusted from 0 mM to 0.25 mM, 0.50 mM, 0.75 mM, 1.0 mM, 2.0 mM, 5.0 mM, and 6.0 mM. Chemical shift changes were monitored after each salt addition by collecting 2D ^{1}H-^{13}C CT-HSQC spectra (30). NOEs between $Co(NH_3)_6^{3+}$ and SL5 were obtained by collecting two ^{13}C-edited HMQC-NOESY spectra with mixing times of 90 ms and 180 ms on a 1.0 mM sample of SL5 in NMR buffer A supplemented with 4 mM $Co(NH_3)_6Cl_3$.

A 1.0 mM sample of $^{13}C/^{15}N$-labeled SL5 in NMR buffer B with 46 mM $MgCl_2$ and 100% D_2O was titrated with $MnCl_2$. The titration was carried out by directly adding varying amounts of concentrated 99.99% $MnCl_2$ (Sigma-Aldrich, MO) to the RNA sample. The Mn^{2+} concentration for this titration was adjusted from 0 μM to 10 μM, 20 μM, 40 μM, and 80 μM $MnCl_2$. The paramagnetic effect of the Mn^{2+} on non-exchangeable protons and their attached carbons was detected by 2D ^{1}H-^{13}C CT-HSQC spectra. A second $MnCl_2$ titration was carried out as described above except that a 1.0 mM sample of ^{15}N-labeled SL5 in NMR buffer B with 20 mM $MgCl_2$ and 90% H_2O: 10% D2O was used. In this case the paramagnetic effect of Mn^{2+} on exchangeable protons and nitrogens was detected by 1D ^{1}H and imino- and amino-optimized 2D ^{1}H-^{15}N HSQC spectra. A third $MnCl_2$ titration was carried out as described above except that a 2.0 mM sample of ^{15}N-labeled SL5 in NMR buffer B with 20 mM $MgCl_2$ and 100% D_2O was used. In this case the paramagnetic effect of Mn^{2+} on non-protonated nitrogens was detected by long-range 2D ^{1}H-^{15}N HSQC spectra optimized for detection of adenine N1 and N3 and purine N7 and N9 (J = 21 Hz) (34). A fourth titration was carried out as described above except that a 1.4 mM sample of unlabeled SL5 in NMR buffer B with 46 mM $MgCl_2$ and 100% D_2O was

used. With this titration, the paramagnetic effect of Mn^{2+} on the phosphate backbone of $SL5^{Mg}$ was detected by 1D ^{31}P and 2D ^{1}H-^{31}P HETCOR spectra.

Structure modeling of $Mn(H_2O)_6^{2+}$ sites

Structures of a complex between $SL5^{Mg}$ and $Mn(H_2O)_6^{2+}$ were obtained by repeating the structure calculation of $SL5^{Mg}$ in the presence of $Mn(H_2O)_6^{2+}$. Peaks from the $MnCl_2$ titration that were broadened to baseline at 20 μM $MnCl_2$ indicated that the nuclei or the pairs of nuclei giving rise to these signals were close in space. Broadened peaks were found in the 1D ^{31}P, the 2D ^{1}H-^{13}C CT-HSQC (29,30), the 2D ^{1}H-^{15}N HSQC spectra (33,34), and the nuclei involved are ^{31}P, directly bonded ^{1}H-^{13}C pairs, directly bonded imino and amino ^{1}H-^{15}N pairs, as well as H2-N3, H2-N1, H8-N7, and H8-N9 pairs. We therefore defined distance constraints between the nuclei or pairs of nuclei that had their peaks broadened to baseline and manganese ions. Exceptionally, for the H8-N7 and H8-N9 pairs, distance constraints were only defined to the nitrogen atoms. For all these distance constraints derived from Mn^{2+} broadening, we used bounds of 1.8-7.0 Å between the nuclei involved and the Mn^{2+} center of the $Mn(H_2O)_6^{2+}$ complex. Only distance restraints defining $Mn(H_2O)_6^{2+}$-binding sites in the loop (residues U695-A702) were modeled. Four $Mn(H_2O)_6^{2+}$-binding sites were defined (see Results). To help localize these $Mn(H_2O)_6^{2+}$, repulsive restraints (> 7.0 Å) were used between these four $Mn(H_2O)_6^{2+}$ and nuclei with resonance that showed little or no specific line broadening at 80 μM $MnCl_2$. For the structure calculation, we used the same protocol used for NMR structure determination of $SL5^{free}$ and $SL5^{Mg}$ (27). However, after randomization of the backbone angles and before the first cycle of high temperature dynamics, four $Mn(H_2O)_6^{2+}$ were placed linearly at the 3'-end of the SL5 RNA. The coordinates for the $Mn(H_2O)_6^{2+}$ complex were obtained by substituting the Mg atom

in a $Mg(H_2O)_6^{2+}$ complex with a Mn atom. The $Mg(H_2O)_6^{2+}$ coordinates and parameters and the Mn parameters were all obtained from HIC-Up (46).

Starting from a set of 50 structures, 46 structures satisfied the experimental restraints (no distance violation > 0.1 Å and no torsion angle violation > 5º). From these 46, the 10 lowest energy structures were selected for further analysis and were used to calculate an average structure that was further minimized against experimental restraints. All structures were visualized and analyzed with MOLMOL (47).

RESULTS

Resonance Assignment of $SL5^{Mg}$

We have previously determined the structure of SL5 in the absence of Mg^{2+} ($SL5^{free}$) and used chemical shift mapping to study the effect of Mg^{2+} on this RNA (27). The addition of Mg^{2+} to $SL5^{free}$ caused chemical shift changes in many of the observed resonances, with larger and widespread chemical shift changes for the residues that form the U-turn motif of $SL5^{free}$ (27). Because Mg^{2+} likely cause a conformational change in SL5, we decided to determine the NMR structure of SL5 in the presence of Mg^{2+} ($SL5^{Mg}$) to better understand the effect of Mg^{2+} on the SL5 structure. The buffer conditions for structure determination of $SL5^{Mg}$ are the same as those chosen for $SL5^{free}$ (NMR buffer A: 10 mM d_{11}-Tris, 50 mM NaCl, 0.02% NaN_3, pH 7.0), except that 40 mM $MgCl_2$ was added to the sample. This $MgCl_2$ concentration caused 90% of the total Mg^{2+}-induced chemical shift changes in the loop resonances of $SL5^{free}$; addition of higher concentrations of $MgCl_2$ were not found practical for structure determination of $SL5^{Mg}$ because they lead to broader linewidths and poorer signal-to-noise ratio (not shown). At a later stage in this study, we discovered that essentially identical spectra of SL5 RNA were obtained in NMR

buffer B supplemented with 20 mM $MgCl_2$ (Figure 4.S1b). This low ionic strength buffer condition leads to improved spectral linewidth and sensitivity; it was used for assignments of exchangeable ^1H and nitrogens and for obtaining one distance constraint from the HNN-COSY experiment which used H2 detection (see below). All other constraints for structural calculation of $SL5^{Mg}$ were obtained from experiments collected in NMR buffer A supplemented with 40 mM $MgCl_2$.

Resonance assignment of $SL5^{Mg}$ was substantially facilitated because the complete resonance assignment of $SL5^{free}$ was available (27). By carefully tracking the ^1H-^{13}C CT-HSQC peaks during the Mg^{2+} titration (27) the assignment of many ^1H and ^{13}C resonances could be obtained for $SL5^{Mg}$. Of course, this method was suitable only for peaks that remain resolved throughout the titration; it allowed complete assignment of the base non-exchangeable protons and carbons, but provided only a few assignments for the ribose protons and carbons. The ribose protons and carbons were assigned from 3D HCCH-COSY (31) and 3D HCCH-TOCSY (31) spectra. Exchangeable protons and their nitrogens were assigned by tracking the ^1H and ^{15}N resonances during the Mg^{2+} titration using imino- and amino-optimized 2D ^1H-^{15}N HSQC (33) as well as long-range 2D ^1H-^{15}N HSQC spectra optimized for detection of adenine N1 and N3 and purine N7 and N9 (34). NOE connectivities from a 2D ^1H-^{15}N CPMG-NOESY (36) was also used to confirm and complete these ^1H and ^{15}N assignments. Partial ^{31}P assignments were obtained from a 2D ^1H-^{31}P HETCOR. Sequential assignments were obtained from a ^{13}C-edited HMQC-NOESY with a mixing time of 150 ms. A list of the resonance assignment for ^1H, ^{13}C, ^{15}N and ^{31}P nuclei in $SL5^{Mg}$ is provided in Table 4.S1.

SL5^{Mg} Structure Determination

The NMR structure of SL5^{Mg} was obtained using standard homonuclear and heteronuclear NMR methods and unlabeled, ^{15}N-labeled, and ^{13}C/^{15}N-labeled RNA. As for SL5^{free}, evidence for base pairing in the stem was obtained from 1D ^{1}H, 2D ^{1}H-^{15}N CPMG-NOESY and the 2D HNN-COSY experiments (37,38). Evidence for formation of the four internal base pairs C691-G705, G692-C704, A693-U703, G694-C702 was obtained from NOE connectivities typical of A-form RNA and from $^{2}J_{NN}$ couplings across hydrogen bonds in Watson-Crick base pairs (37,38). The standard HNN-COSY experiment which detect $^{2}J_{NN}$ couplings via the intervening imino proton was not useful for characterizing the U695-A701 base pair because the U695 imino proton was not observed. We used a modified version of the HNN-COSY experiment which allows detection of $^{2}J_{NN}$ couplings across hydrogen bonds in A-U base pair via the non-exchangeable adenine H2 (37,38). Using the common buffer conditions used for structure determination (NMR buffer A + 40 mM $MgCl_2$), no evidence of hydrogen bonding between U695 N3 and A701 N1 was obtained, because of the low sensitivity of the data. In order to improve on the spectral sensitivity, the same experiments was recorded on a 2.0 mM SL5 RNA in NMR buffer B and 20 mM $MgCl_2$. One advantage of these low ionic strength conditions is that they improved spectral linewidth and sensitivity and allowed for detection of the two-bond U696 N3 to A700 N1 J coupling in the modified HNN-COSY experiment (38). For the structure calculation, planarity and geometry constraints typical of Watson-Crick base pairs were enforced for all the base pairs in SL5^{Mg} except for the U695-U701 base pair. Instead a single distance constraint was defined between U695 N3 and A701 N1 based on the HNN-COSY data.

As was found for SL5^{free}, the residues in the stem (690-695 and 701-706) of SL5^{Mg} showed spectral characteristics typical of A-form helices (48). Intra- and inter-residue NOEs

from the 3D ^{13}C-edited HMQC-NOESY spectra for the stem residues are characteristics of A-form helical geometry and $^{3}J_{H1'-H2'}$ derived from a 3D HCCH E. COSY (39) indicated that all stem residues adopt a C3'-endo sugar pucker conformation.

In the structure of SL5^{free} there is a sharp turn in the backbone after residue U696. The residues after the turn G697, A698, and C699 all have their bases stacked on each other and residue U700 has its base extruded from the loop (27). The loop NOEs indicated differences in the structure between SL5^{free} and SL5^{Mg}. In both SL5^{free} and SL5^{Mg}, the presence of NOEs from the ribose of U696 to G697 H8, A698 H8 and C699 H5 indicated that there is a turn in the loop after U696. However, there is an increase in the calibrated intensities of the NOEs between U696 and residues A698 and C699 in SL5^{Mg} compare to SL5^{free}. The stronger calibrated intensities indicate that U696 is closer to A698 and C699 in SL5^{Mg}. Two other observations specific to the spectra of SL5^{Mg} are the increase in the calibrated intensity of the NOEs between the ribose of C699 and base of A701, and the presence of an NOE between C699 H5 and A701 H2. There is no stacking between the bases of residues C699 and A701 in SL5^{free}, however these NOE data indicate stacking between these residues in SL5^{Mg}. As for SL5^{free}, NOE data indicate that in SL5^{Mg} the base of U700 does not stack with other bases in the loop.

Three-dimensional structures were calculated using the distance and torsion restraints described in the Materials and Methods. The ten lowest energy structure from the group of 69 that satisfied our experimental restraints (no distance violation > 0.1 Å and no torsion angle violation > 5˚) were retained for analysis. An average structure was calculated from these ten low energy structures and minimized against experimental restraints (Figure 4.2a). The structural statistics (Table 4.1) indicate that the structure of SL5^{Mg}, including the terminal loop, is well defined by the NMR data.

SL5^{Mg} Forms a Canonical U-Turn Motif

As we have found for SL5^{free}, SL5^{Mg} also forms a hairpin with an A-form stem and a five-membered loop, which adopts a U-turn motif (Figure 4.2). SL5^{free} and SL5^{Mg} share many of the structural characteristics found in U-turn motifs (Figures 4.2 and Table 4.2). In both RNAs, U696, G697, and A698 form the consensus UNR sequence (U = uracil, N = any base, R = purine) of U-turn motifs. They both display the sharp turn after U696 that allows backbone reversal, and stacking of G697, A698, and C699 bases after the turn. In SL5^{free}, the α torsion angle at the turning phosphate (G697 5'P) is in a staggered conformation with values of 116 ± 7º for the ten lowest energy structures (Table 4.2). However in SL5^{Mg}, the values for these angles was 167 ± 34º indicating that the α torsion angle at the turning phosphate was in the trans conformation. The short distance between U696 O2' and A698 N7 in SL5^{free} and SL5^{Mg} indicate that they also both form one of the two characteristic hydrogen bonds found in canonical U-turn motifs. In SL5^{Mg} the distance between the U696 O2' and the A698 N7 is on average 0.7 Å shorter than the same distance in SL5^{free} (Table 4.2). SL5^{free} and SL5^{Mg} also share one characteristic that is not observed in canonical U-turns; their flanking base pair is the Watson-Crick U695-A701 base pair, whereas canonical U-turns contain non-canonical flanking base pairs. In SL5^{free}, there is no stacking between U696 and 5'-phosphate of A698. However, this stacking interaction found in canonical U-turns is also present in SL5^{Mg}. In SL5^{free}, there is no evidence for hydrogen bonding between the U696 N3 and 3'-phosphate group of A698. However, the U696 N3 to A698 N7 distance in the 10 lowest structures of SL5^{Mg} is compatible with this hydrogen-bonding interaction, which is characteristic of canonical U-turn motifs. In summary, if we exclude the flanking base pair, all the structural characteristics of the canonical U-turn motifs are present in SL5^{Mg}, whereas they were not in SL5^{free}. As a result the U-turn

backbone fold of $SL5^{Mg}$ appears more compact than what we observed previously for $SL5^{free}$ (Figure 4.3).

A superposition of the heavy atoms in the loops (residues U696-A701) of $SL5^{free}$ and $SL5^{Mg}$ illustrates the subtle but important conformational change that occurs in the loop as a result of magnesium-ion binding (Figure 4.3). The heavy atom RMSD for this superposition is 2.77 Å. The conformational change that occurs between $SL5^{free}$ and $SL5^{Mg}$ involves primarily the formation of a stacking interaction between the bases of C699 and A701. In $SL5^{Mg}$, the stacking of C699 unto A701 brings C699 closer to U696 by an average distance of 1.6 Å. Two hydrogen-bonding interactions between U696 and C699 that were not observed in $SL5^{free}$ are now found in $SL5^{Mg}$. There are: 1) the one already mentioned between U696 N3 and the C699 5'-phosphate group and 2) another one between U696 O2 and an amino proton of C699 (distance of 3.36 Å in the average structure). Another significant difference between $SL5^{free}$ and $SL5^{Mg}$ is in the position of U700. In both cases, the base of U700 is completely extruded from the other bases of the loop. However, whereas the U700 base is found on the major groove stacked on the 5'-phosphate of C699 in $SL5^{free}$, it is found on the minor groove face of the loop in $SL5^{Mg}$ (Figure 4.3). This difference in the base orientation of U700 may be a result of the large conformational change in the phosphate backbone between C699 and A701. In $SL5^{Mg}$, the U700 ribose is also extruded from the loop and is located between the riboses of A701 and A702, whereas in $SL5^{free}$, this ribose was found in the usual register between the riboses of C699 and A701. In the $SL5^{Mg}$ conformation, the ribose and base of U700 are free for recognition and are found on the same groove side as the hydrogen-bonding face of G697, A698 and C699 which are known to form base-pairing interactions with stem-loop I.

Elucidation of Metal Ion-Binding Sites in SL5^{Mg}

Previous attempts to use a titration of Mg^{2+} and chemical shift mapping to elucidate binding-site information lead to the idea that SL5 undergoes a conformational change upon metal-ion binding (27). Here we have shown that Mg^{2+}-dependent chemical shift changes are at least partly due to conformational changes in the SL5 loop. By inspecting the structure of SL5^{Mg} (Figure 4.2a) some of the chemical shift changes in the magnesium titration may be explained. As stated earlier, the stacking between C699 and A701 would lead to an upfield shift in the chemical shifts of the base protons in both residues. This is the case as C699 H6 is shifted upfield by 0.55 ppm and A701 H8 is shifted upfield by 0.11 ppm (Figure 4.S1). To get a clearer understanding of which residues may be involved in direct magnesium-ion interaction in SL5^{Mg} two other indirect methods were employed to try to identify sites of magnesium interaction; NOEs to cobalt hexammine and paramagnetic line-broadening caused by Mn^{2+}. Both methods involved direct detection of other metal cations that may bind to similar sites and in a similar manner as magnesium ions (49-51).

NOEs to Cobalt Hexammine

Cobalt hexammine ions are known to fold the ribozyme into the correct conformation but they do not support catalysis (15). From the titration shown in Figure 4.S1a, it can be seen that the chemical shift changes upon addition of cobalt hexammine ions are very similar to those observed upon addition of magnesium ions for most of the peaks shown. Even though the peaks do not have the exact same chemical shift changes they are affected to a similar degree in terms of direction of the shift being the same in both the ^{1}H and ^{13}C dimensions. This may indicate that cobalt hexammine ions may induce the same conformational change in SL5 but may not bind to

all the same sites or in the exact same manner as magnesium ions. However, since the conformational change is the similar it justifies using $Co(NH_3)_6^{3+}$, as a probe for magnesium ion binding.

NOEs to cobalt hexammine ions were obtained by collecting two 3D ^{13}C-edited HMQC-NOESY spectra with mixing times of 90 and 180 ms in the presence of cobalt hexammine ions (4 mM $Co(NH_3)_6^{3+}$ in NMR buffer A). The 18 protons of the $Co(NH_3)_6^{3+}$ resonated at a single frequency of 3.60 ppm, showing that the cobalt hexammine ion is free to rotate in the bound conformation and is in fast exchange between the bound and unbound states (52). This means that on the timescale of the NMR experiment, a single cobalt hexammine ion is free to interact with more than one binding site. This behavior has been seen for other RNAs that have been shown to bind $Co(NH_3)_6^{3+}$ by NMR spectroscopy (36,53,54). Protons that were resolved and showed an NOE crosspeak to cobalt hexammine ions were mapped unto the structure of the $SL5^{Mg}$ loop (Figure 4.5a) and are listed in Supplementary Table 4.S2. The $Co(NH_3)_6^{3+}$ ions interact with the major groove of $SL5^{Mg}$ and NOEs are observed in the loop with 695 H5 and H6 and 696 H5. The exact number of binding sites was difficult to discern, as it not clear whether the binding sites in the major groove are discrete sites or caused by the movement of ions along the major groove. We did not attempt to localize the binding sites for $Co(NH_3)_6^{3+}$ using this data. There are two residues in the loop that gave NOEs to $Co(NH_3)_6^{3+}$, U695 and U696 (Figure 4.5a). Whether these residues represent one binding site or two is not immediately clear from the data. The turn in the loop occurs right after residue U696 and the fact that the cobalt hexammine ions bind so close to the turn suggests that it may play a role in stabilizing the turn. The binding of $Co(NH_3)_6^{3+}$ gives an indication of where hexahydrated magnesium ions may bind, but may not unravel all Mg^{2+}-binding sites. $Co(NH_3)_6^{3+}$ does not give up its NH_3 ligands when it binds, it

only exhibits an outer-sphere mechanism of binding (50). Magnesium ions, on the other hand, can exhibit several binding modes in aqueous solution including outer and inner sphere mechanisms when binding to RNA (55).

Mn^{2+}-Induced Paramagnetic Line-Broadening

Mn^{2+} is a paramagnetic ion that specifically broadens the linewidth of any nuclei that is close to it (within ≈ 10 Å) (49). This paramagnetic line broadening is a through-space effect, and is proportional to r^{-6}, where r is the distance between the Mn^{2+} and the observed nuclei (56). It takes only micromolar concentrations of Mn^{2+} to show the effect, and in very high concentrations of Mn^{2+} all the resonances will be affected due to non-specific binding. Mn^{2+} was used as a probe for magnesium-ion binding by titrating $SL5^{Mg}$ with micromolar amounts of $MnCl_2$. The experiments employed to detect resonance line broadening are listed in the Materials and Methods. Manganese ions like magnesium ions can exhibit both outer-sphere and inner-sphere binding modes. The added advantage of using Mn^{2+} is that it allowed for other nuclei apart from ^1H to be investigated. This allowed for the observation of line broadening of potential hydrogen-bond donors and acceptors in the $SL5^{Mg}$ loop. The effect on the phosphorous was particularly useful as it allowed for investigation of the backbone of the loop. It should be noted that since the titration was carried out in the presence of near saturating amounts of $MgCl_2$, the Mn^{2+} did not affect the chemical shifts of the atoms, only their linewidths. Supplementary Table 4.S3 lists the atoms which resonances are broadened by the manganese ions (at a concentration of 20 μM) in $SL5^{Mg}$. Almost every atom that gave an NOE to $Co(NH_3)_6^{3+}$ are also affected by Mn^{2+}. Mn^{2+} however affects more sites in the loop of $SL5^{Mg}$ than $Co(NH_3)_6^{3+}$. Indeed, every residue in the loop except for U700 was affected by Mn^{2+} at a concentration of 20 μM. Indeed,

Figure 4.5b shows all the nuclei in the loop that were broadened to baseline at 20 μM Mn^{2+} during the manganese ion titration.

Structure Modeling of $SL5^{Mg}$ with $Mn(H_2O)_6^{2+}$

In order to get a model of a metal-ion bound complex of the loop of $SL5^{Mg}$ it was decided to do a structure calculation using $Mn(H_2O)_6^{2+}$ as a model for magnesium-ion binding. To determine the number of manganese ions to include in the model an initial analysis of the distances between the affected atoms was done. The affected atoms of U696, being located at one edge of the loop, constitute one site. The N7 atoms of G697 and A698 are on the minor groove side of the loop, whereas the other affected atoms are located on the major groove side. These atoms constitute another site. The 5' phosphate of C699 belongs to a third independent site. The amino group of C702 is located close to the N7 of A701 and the H5 of U695, and these atoms may share the same site. So based on separation alone, we inferred four ion-binding sites in the loop. The structure calculation of $SL5^{Mg}$ was repeated with the inclusion of four $Mn(H_2O)_6^{2+}$. Restraints between the Mn center of the $Mn(H_2O)_6^{2+}$ and the $SL5^{Mg}$ loop atoms that were specifically broadened to baseline in the presence of 20 μM $MnCl_2$ were used (Supplementary Table 4.S2). In the first round of the structure calculation, Site 1 was constrained to U696 H6 and C6; Site 2 was constrained to G697 N7 and N9, and 698 N7 and N9; Site 3 was constrained to C699 5' phosphate; and Site 4 was constrained to 702 N4, H41 and H42. The rest of the affected loop atoms (U695 H5 and C5, U696 H5 and C5, G697 5'P, 698 5'P, and A701 N7 and N9) were restrained to any of the four manganese ions. These restraints were conservatively set to 1.8-7.0 Å. In order to better position the ions within the binding sites, repulsive restraints (> 7.0 Å) were

added for those atoms that did not experience any line broadening in the presence of 80 μM $MnCl_2$.

After an initial round of structure calculation it was found that U695 H5 and C5 had a preference for the manganese at Site 3, whereas the phosphates of G697 and A698, and U696 H5 and C5 had a preference for the manganese at Site 1, and A701 N7 and N9 had a preference for Site 4. These atoms were restrained to these sites and the structure calculation was redone.

The superposition of the loops of the ten lowest energy structures and minimized average structure of the loop of $SL5^{Mg}$ modeled with the four hexahydrated manganese ions are shown in Figure 4.6a. The heavy atom RMSD for the overlay is 0.65 Å for residues G694 to C702. If the four manganese ions are included, the RMSD increases to 0.983 Å. The RMSD for an overlay of residues 694-702 of $SL5^{Mg}$ in the presence and absence of Mn^{2+} was 0.703 Å. To determine which atoms coordinated the manganese ions, potential ligands that were within 4.0 Å of the oxygen atoms of $Mn(H_2O)_6^{2+}$ were determined (Table 4.3). Three of the manganese ions are directly coordinated to the UNR sequence of the U-turn motif. The manganese ion at Site 1 seems to be coordinated to a pocket formed by the 5'-phosphate groups of U696, G697 and A698. Site 2 is coordinated to the N7 atoms of G697 and A698, whereas Site 3 seems to be coordinated to the O4 of U696 and U695, as well as, the non-bridging oxygens of the 5'-phosphate of C699. Site 4 seems to be coordinated to the non-bridging oxygen of the 5'-phosphate and O3'of U700 and the N7 of A701 (Table 4.3).

DISCUSSION

Mg^{2+} Stabilizes the U-Turn Motif of Stem-Loop V in its Canonical Form

In the absence of any magnesium ions or any other multivalent ion, the terminal loop of SL5 forms a loose U-turn motif (Paper 1). The addition of magnesium ions to the buffer causes a conformational change in the loop of SL5 resulting in the formation of a canonical U-turn. The presence of a Watson-Crick base pair 5' of the U-turn greatly reduces the flexibility of the backbone in the loop. The A-form conformation of the helix adjacent to the loop puts a structural constraint on the loop that prevents other RNAs with UNR sequences to form U-turns (57). In canonical U-turns, the constraint on the loop is removed by having a non-Watson Crick base pair that stops the A-form character of the helix from constraining the loop (25). Another way to remove this constraint is to have a very large loop with some of the residues unpaired as seen in the 23S rRNA (22). This makes it easier for the transition back to the A-form conformation by the backbone. The formation of a canonical U-turn requires a very sharp turn in the RNA backbone and the stacking of the 5'-phosphate of R on the base of U. This conformation results in the clustering of three phosphate groups at the turn in the backbone. This repulsive effect can be overcome in the absence of metal ions by the formation of one or two hydrogen bonding interaction involving the U residue preceding the turn and by the stacking of the residues after the turn. The more critical of the two hydrogen bonds occurs between the U N3 and the R+1 5'-phosphate. This is the defining hydrogen bonding interaction for U-turns (25).

In the structure of $SL5^{Mg}$ (Figure 4.2) the phosphates of U696, G697 and A698 are all grouped together right at the turn in the backbone. In the absence of magnesium ions, $SL5^{free}$ was unable to form all the stabilizing interactions that would mitigate the repulsive effect of the phosphates because of the constraints placed on the backbone. It was more energetically

favorable to relax the U-turn by preventing the clustering of the three phosphates. Hence, the phosphate of A698 was not stacked unto the base of U696. This prevented the formation of the defining hydrogen bond between U696 N3 and C699 5'-phosphate. The base of residue U700 was also extruded from the loop and was not involved in any base-pairing interactions. By extruding this residue the backbone of $SL5^{free}$ was given the possibility, to rejoin the A-form conformation at the base of the loop. In the structure of $SL5^{Mg}$ we see that the binding of magnesium ions to the loop causes a conformational change in the U-turn of $SL5^{Mg}$ adopting a compact conformation with more canonical features. The magnesium ions helps to overcome the constraints placed on the loop by the adjacent Watson-Crick base pair by mitigating the repulsive effects of the phosphate clustering.

Manganese ions have been used to model the binding of magnesium ions to the loop of $SL5^{Mg}$. There are four sites of manganese-ion coordination to the loop of $SL5^{Mg}$, with three $Mn(H_2O)_6^{2+}$ binding specifically to the residues directly involved in the U-turn. The repulsive effect of the phosphate clustering of the 5'-phosphates of U696, G697 and A698 is countered by the presence of a manganese ion that binds in the pocket created by these three phosphate groups (Figure 4.6, Site 1). A second manganese-ion binding site is to the N7 atoms of G697 and A698 (Figure 4.6, Site 2). Binding to this site may be a consequence of the N7 groups being exposed to solvent. A third manganese ion binds to the phosphate involved in the defining hydrogen bond between U696 H3 and the 5'-phosphate of C699 (Figure 4.6, Site 3). The final site is Site 4 and it involves the coordination of the U700 5'-phosphate and the N7 of A701. The manganese ions may help to stabilize the turn in the backbone that occurs after U700. Site 3 is an important site in the context of the full ribozyme, as disruption of the magnesium ion binding ability of this phosphate group by phosphorothioate interference results in a reduction of the catalytic activity

of the ribozyme (17). Manganese ion coordination to the 5' phosphate of C699 was also shown through manganese rescue of the phosphorothioate interference (17). The manganese ion at Site 3 is also coordinated to U696 O4 and may play a role in bringing the groups involved in the hydrogen bond together. The manganese rescue of the phosphorothioate interference did not identify the phosphate cluster at Site 1 presented here. A possible reason for this is the fact that the interference reaction affects only one phosphate at a time. Since this site involves three phosphates, if one phosphate was affected the magnesium ion can still coordinate to the two other phosphates. The ribose and the base of U700 are extruded from the loop, possibly to remove any steric conflict that might have arisen from C699 and U696 moving closer together and the binding of the magnesium ion at Site 3. In $SL5^{free}$, the base of U700 was stacked on the 5'-phosphate of C699; in $SL5^{Mg}$ this stacking interaction is not present to allow the formation of the hydrogen bond between this phosphate group and U696 H3 and for the coordination of this phosphate to magnesium. The disadvantage of using Mn^{2+} to model magnesium ion binding is that the exact hydration number of the Mn^{2+} is unknown, as Mn^{2+} can exhibit both inner-sphere and outer-sphere binding mechanisms. The Mn^{2+} results were compared to the results of using $Co(NH_3)_6^{3+}$, which only exhibit an outer-sphere mechanism of metal ion binding. The $Co(NH_3)_6^{3+}$ only gave NOEs to the base protons of residues U695 and U696 in the loop of SL5 (Figure 4.5a). These residues are located at Sites 3 and 1 respectively. One interpretation of this result is that Mn^{2+}, and by extension Mg^{2+}, binds to these sites in the hexahydrated form, and to the other sites with a lower coordination number than six.

We found several instances in which magnesium ions are coordinated directly to U-turns, with the same sites found in $SL5^{Mg}$, although the degree of occupancy varied. There was one instance in which all three sites were occupied at the same time. In the crystal structure of the

30S ribosomal unit of *Thermus Thermophilus* complexed with messenger RNA fragments and the anticodon loop of $tRNA^{Phe}$ (58), there is a U-turn at position U14 to A16 of the 16S rRNA fragment. The sequence of the U-turn is UGA and the loop is closed by a U-U mismatch. All three sites of manganese ion coordination seen in the U-turn of $SL5^{Mg}$ (Sites 1 to 3) model are occupied by magnesium ions in the 16S rRNA fragment (Figure 4.7a). In the 16S rRNA structure the distance from the magnesium ion to the N7 atoms after the U-turn is greater than 4.3 Å for the G15 N7 and greater than 5.0 Å for the A16 N7. These distances are large enough for the interaction to be mediated by water molecules suggesting that the magnesium may be hydrated. In the 2.2 Å crystal structure of $tRNA^{Ile}$ complexed with isoleucyl-transfer RNA (tRNA) synthetase and mupirocin, the U-turn found in the $tRNA^{Ile}$ TψC loop has two hexahydrated magnesium bound to it (59). Both these sites correspond to Sites 1 and 3 in $SL5^{Mg}$ (Figure 4.7b). The TψC sequence has been modified to UUC in this crystal structure. The UNR sequence for this $tRNA^{Ile}$ U-turn is UCA, therefore there is no N7 position at the N residue for the site analogous to $SL5^{Mg}$ Site 2 to occur. This is the case for all U-turns in the TψC loop of tRNAs, which means that Site 2 is not expected in the TψC loop of tRNAs. The fact that in the $tRNA^{Ile}$ TψC loop U-turn, the magnesium ions are hexahydrated will give a better insight into what groups coordinate the magnesium ions. At the site equivalent to Site 1 of $SL5^{Mg}$ the magnesium ion seems to be coordinated to the three-phosphate cluster. The second hexahydrated magnesium ion equivalent to Site 3 of $SL5^{Mg}$ has an oxygen atom positioned equidistant from the 5'-phosphate of A58 and the O4 group of U55 (Figure 4.7b). It can be concluded that these two sites equivalent to Sites 1 and 3 in $SL5^{Mg}$ may play an important role in stabilizing U-turn motifs in RNA.

How Mg^{2+} May Facilitate the Binding of Stem-Loop V to Stem-Loop I

An important step in the cleavage reaction pathway for the *Neurospora* VS ribozyme is the formation of the loop-loop interaction between stem loops I and V. This interaction involves the formation of Watson-Crick base pairs between the stem-loop I residues 630-632, and the SL5 residues G697-C699. This means that the SL5 loop residues G697-C699 need to be accessible for formation of this interaction. This is the function of the U-turn in SL5, as all three residues have their Watson-Crick faces exposed to the solvent and are in a position to form this interaction. In previous work when we superposed $SL5^{free}$ against an idealized A-form helical model of the loop-loop interaction the fit was not ideal. The heavy-atom RMSD for residues G697-C699 was 1.66 Å and the base with the lowest RMSD for the fit was C699. When $SL5^{Mg}$ is superposed unto the idealized A-form helix the RMSD for the three loop nucleotides was an improved 1.3 Å (Figure 4.8). The bases are aligned better in $SL5^{Mg}$, with all the hydrogen bond donors and acceptors in the model aligned very well (Figure 4.8). It can be seen that the stacking of C699 unto A701 allowed for better positioning of the base of C699 to form a hydrogen bond with G632. The terminal loop of stem-loop I has been postulated to contain a U-turn motif based on sequence, NAIM analysis, and site-specific substitution (18,19,60). Residues U628-G630 in stem-loop I form a UNR consensus sequence. Replacing U628 with any other residue severely inhibits the ribozyme (18). Removal of the 2'-OH of U628 also inhibits the ribozyme (19,60). The terminal loop of stem-loop I is homologous to stem-loop V in that there is a Watson-Crick base pair 5' to the U-turn motif. Phosphorothioate interference and manganese rescue shows that there is a site of manganese interaction at the phosphate of 3' of G630 (17). This is homologous to site 2 in $SL5^{Mg}$. A structure of wild-type stem-loop I was solved by NMR, but the terminal loop was disordered (13). The structure was solved in the absence of magnesium ions so it is

possible that it was unable to adopt the U-turn conformation. Here we have shown that indeed magnesium ions are important for formation of the compact U-turn motif in stem-loop V and likely have a similar role in stem-loop I. We propose that one of the roles of magnesium ions in the VS ribozyme is the formation of the proper U-turn conformation in stem-loop V, and possibly stem-loop I, which allows these loops to form a stable loop-loop interaction. The importance of U-turn motifs in both terminal loops of stem-loop I and V was illustrated through a mutagenic study of the loop-loop interaction (18). Substitutions that disrupted the U-turn motif in either stem-loop I or V reduced the catalytic activity of the ribozyme.

We have shown that magnesium ions bind to the terminal loop of SL5 causing a conformational change in the loop. This conformational change causes the U-turn in SL5 to adopt a more canonical conformation. This is achieved by the binding of four magnesium ions in the SL5 loop, three sites in the UNR sequence that alleviate the constraints placed on the backbone by the A-form helical conformation of the stem. These three sites have been seen in other U-turn structures and may play a similar role in the other U-turn structures. The more compact U-turn structure in the presence of magnesium ions is likely to be necessary for optimal binding to stem-loop I.

ACKNOWLEDGEMENTS

We would like to thank C. Hoogstraten for the HCCH-E.COSY pulse sequence, and A. Majumdar for the HNN-COSY pulse sequences. This work was supported by an NSF Career Award to P.L.

FIGURES and TABLES

Table 4.1. Statistics from the structure calculation of $SL5^{Mg}$.

NOE Restraints	408
Number of NOE-derived distance restraints	379
From standard NOESY experiments	340
Inter-nucleotide	124
Intra-nucleotide	216
From 2D ^{1}H-^{15}N CMPG-NOESY	39
Hydrogen-bond restraints	29
Dihedral Restraints	24
Sugar Pucker (δ)	12
Backbone (γ)	12
Total Number of Restraints	432
RMSD from Experimental Restraints	
NOE (Å) (none > 0.1 Å)	0.005 +/- 0.001
Dihedral (º) (none > 5º)	0.046 +/- 0.011
RMSD from Idealized Geometry	
Bonds (Å)	0.003959 +/- 0.000017
Angles (º)	0.960090 +/- 0.000573
Impropers (º)	0.362660 +/- 0.001690
Heavy-Atom RMSDs to the Minimized Average Structure (Å)	
Overall (residues 691-705)	1.00 +/- 0.23
Stem (residues 691-695 and 701-705)	0.51 +/- 0.20
Loop (residues 696-700)	0.72 +/- 0.18

Table 4.2. Comparison of the canonical U-turn characteristics (25,28) between $SL5^{free}$ and $SL5^{Mg}$

U-turn Characteristics	**$SL5^{free}$**	**$SL5^{Mg}$**
UNR sequence	Yes (U696, G697, A698)	Yes (U696, G697, A698)
Non-canonical flanking base pair	No (Flanking base pair is the Watson-Crick U696-A701 base pair)	No (Flanking base pair is the Watson-Crick U696-A701 base pair)
Sharp turn in backbone	Yes (α angles are 116 ± 7 Å in 10 lowest energy structures)	Yes (α angles are 167 ± 34 Å in 10 lowest energy structures)
Stacking of bases immediately after turn	Yes (stacking of G697, A698, and C699)	Yes (stacking of G697, A698, and C699)
Stacking of U base and R 5'-phosphate group	No	Yes
H-bond between U 2'-OH and R N7	Yes (U696 O2' - A698 N7 distance range in 10 lowest energy structures: 2.91 - 3.93 Å)	Yes (U696 O2' - A698 N7 distance range in 10 lowest energy structures: 2.45 - 2.95 Å)
H-bond between U H3 and R 3'-phosphate group	No (U696 N3 - A698 3'-P distance range in 10 lowest energy structures: 8.19 - 9.18 Å)	Yes (U696 N3 - A698 3'-P distance range in 10 lowest energy structures: 4.06 - 5.25 Å)

Table 4.3. List of possible coordination sites for the $Mn(H_2O)_6^{2+}$ ligands from the modeling of $Mn(H_2O)_6^{2+}$ to the loop of $SL5^{Mg}$.

$Mn(H_2O)_6^{2+}$ Site	**Coordinating Atoms[a]**
1	696 5'-OP1, 5'-OP2, O5' 697 5'-OP1, O3' 698 5'-OP1, 5'-OP2
2	696 O2' 697 N7 698 N7
3	695 O4 696 O4, 699 5'-OP1, 5'-OP2
4	700 5'-OP1, O3' 701 N7

[a] The coordinating atoms were obtained by selecting all hydrogen bond acceptors that were within 4.0 Å of a $Mn(H_2O)^{2+}$ oxygen in the minimized average structure.

Figure 4.1. Secondary structure and proposed Mg^{2+} binding sites in the VS ribozyme and in SL5. **a)** Sequence and secondary structure of the *Neurospora* VS ribozyme. The cleavage site is indicated by the red arrowhead. The interaction between stem-loops I and V is indicated by a dashed line and residues involved in this interaction are circled. Upon interaction with stem-loop V, stem-loop I (subdivided in Ia and Ib) undergoes a structural change from an inactive to an active conformation. **b)** Sequence and secondary structure of the stem-loop V (SL5) RNA used for NMR study. Wild-type and mutant nucleotides are represented by upper and lower cases, respectively. In **a)** and **b)** the blue arrowheads indicate proposed sites of magnesium ion binding from phosphorothioate interference and manganese rescue experiments (17).

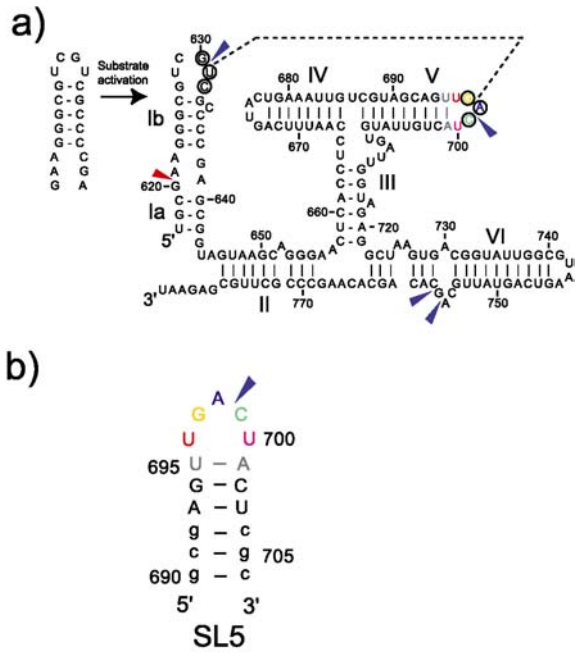


Figure 4.2. NMR structure of $SL5^{Mg}$. **a)** Stereo view of the superposition of the ten lowest energy structures on the minimized average structure of $SL5^{Mg}$. The superposition was obtained by minimization of pairwise heavy atom RMSD of each lowest energy structure to the minimized average structure (residues 691-705). **b)** Minimized average structure of the loop of $SL5^{Mg}$ showing the two characteristic hydrogen bonds found in U-turn motifs: 1) U696 H3 to A698 3'-phosphate; and 2) U696 2'-OH to A698 N7. For simplicity, only heavy atoms are shown in **a)** and **b)**.

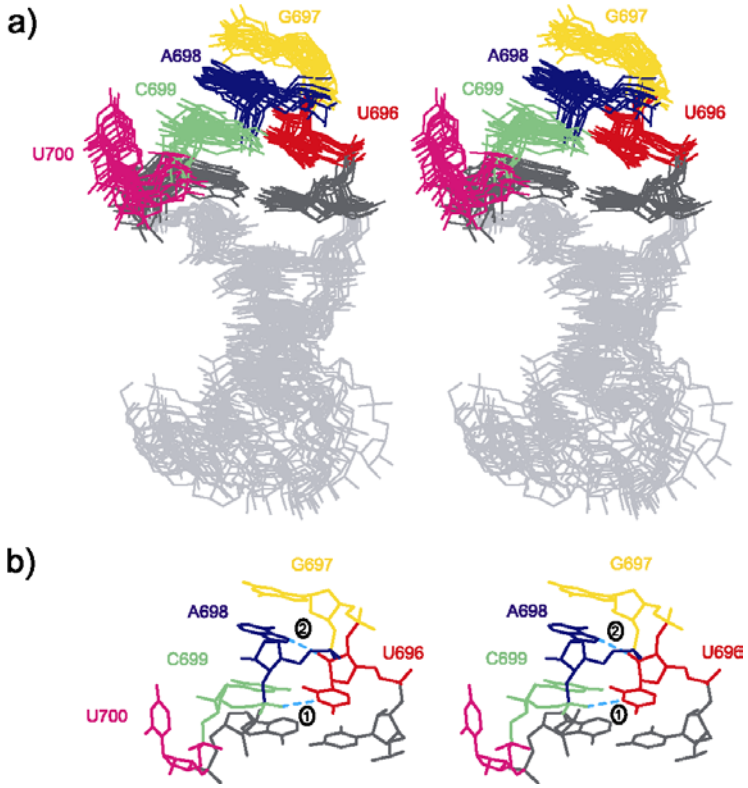


Figure 4.3. Conformational change in the loop of SL5 upon magnesium-ion binding. Superposition of the terminal loops of $SL5^{free}$ (in pale grey) (27) and $SL5^{Mg}$ (in dark grey). The superposition is of the heavy atoms of residues U695 to A701. The heavy atom RMSD for the superposition is 2.77 Å.

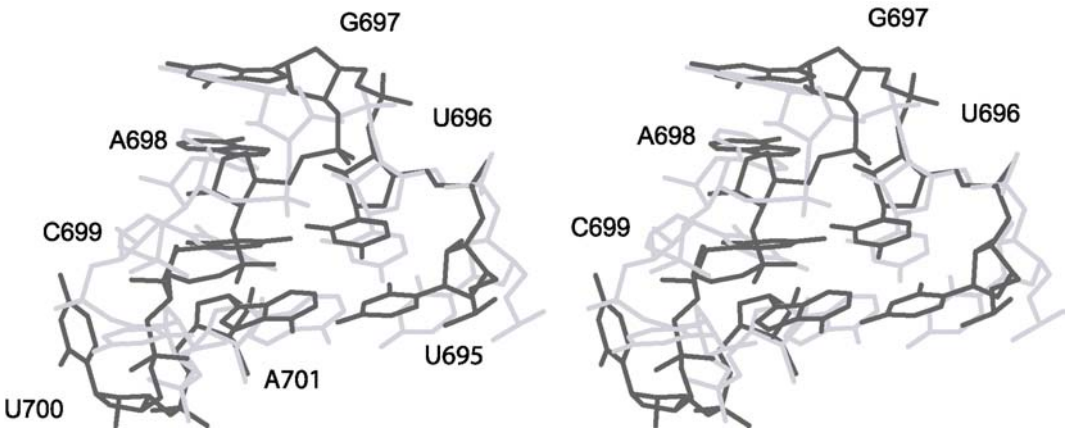


Figure 4.4. Paramagnetic effect of manganese ions on $SL5^{Mg}$. H6-C6/H8-C8 region of 2D ^{1}H-^{13}C CT-HSQC spectra of $SL5^{Mg}$ in the presence of **a)** 0 μM and **b)** 20 μM $MnCl_2$. These spectra were collected in 100% D_2O at 25 ˚C on a 600 MHz NMR spectrometer. Peaks that disappear below the baseline level are indicated by an empty circle.

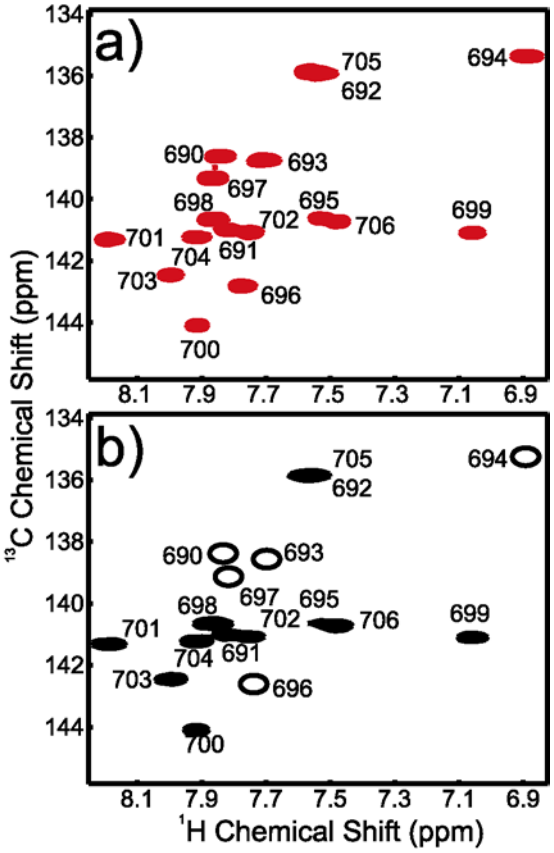


Figure 4.5. The interaction of $Co(NH_3)_6^{3+}$ and Mn^{2+} with the loop of SL5. **a)** Protons that give a NOE to $Co(NH_3)_6^{3+}$ were mapped as red spheres on the corresponding atoms of the minimized average structure of $SL5^{Mg}$ (U695 H5, U695 H6, and U696 H5). **b)** Nuclei that gave rise to resonances broadened below baseline in the presence of 20 μM Mn^{2+} were mapped as blue spheres on the corresponding atoms of the minimized average structure of $SL5^{Mg}$ (U695 C5 and H5; U696 C5 and H5; U696 C6 and H6; G697 5'-P; G697 C8 and H8; G697 N7; A698 5'-P; A698 N7; C699 5'-P; and 701 N7)

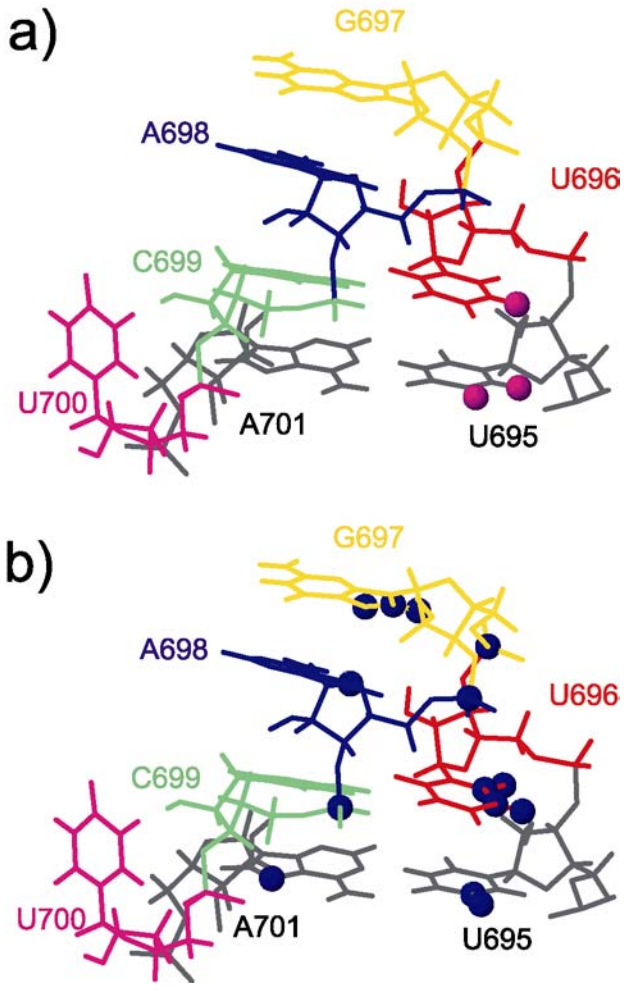
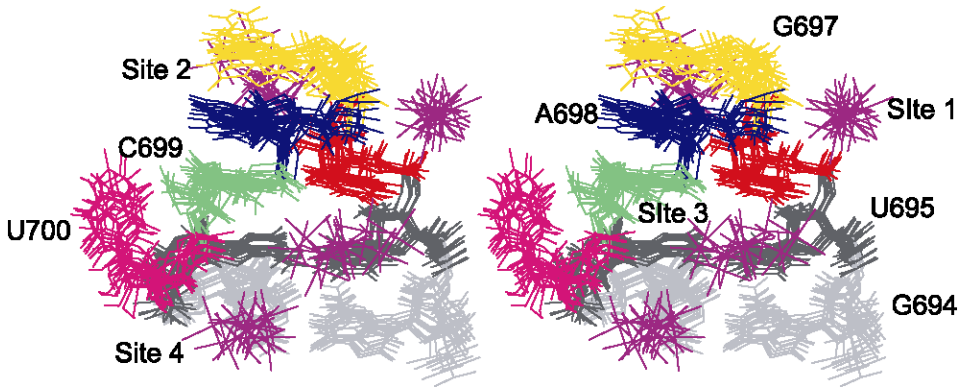


Figure 4.6. Localization of divalent metal-ion binding sites in $SL5^{Mg}$. a) Superposition of the 10 lowest energy structure $SL5^{Mg}$ modeled with four $Mn(H_2O)_6^{2+}$ complexes. The superposition was obtained by minimizing the RMSD of the heavy atoms of residues 694-702. b) Minimized average structure of $SL5^{Mg}$ modeled with four $Mn(H_2O)_6^{2+}$ complexes. In a) and b) only residues G694 to C702 are shown.

a)



b)

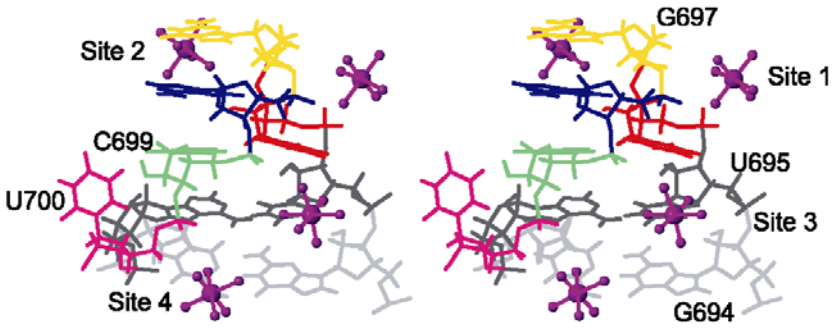


Figure 4.7. Examples of Mg^{2+} binding to U-turns. **a)** The U-turn motif found in the anticodon loop of $tRNA^{Phe}$ from the 3.11 Å resolution crystal structure of the 30S ribosomal unit complexed with messenger RNA fragments and $tRNA^{Phe}$ (PDB entry 1IBL) (58). Three magnesium ions were found associated with this U-turn motif, which are anologous to sites 1, 2, and 3 found in $SL5^{Mg}$. **b)** The U-turn motif found in the $tRNA^{Ile}$ TψC loop from the 2.2 Å resolution crystal structure of $tRNA^{Ile}$ complexed with isoleucyl-transfer RNA (tRNA) synthetase and mupirocin (PDB entry 1FFY) (59). The magnesium ions complexed to this U-turn are hexahydrated, and they are found in sites analogous to sites 1 and sites 3 in $SL5^{Mg}$.

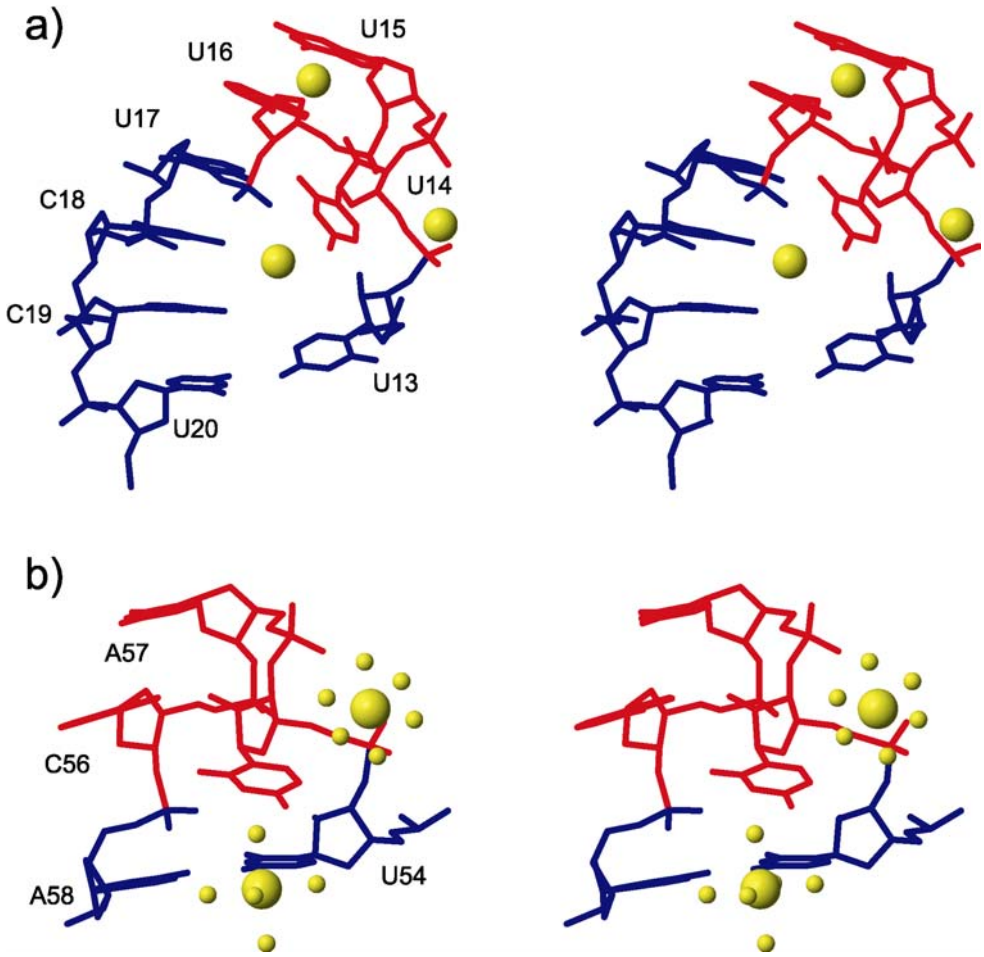
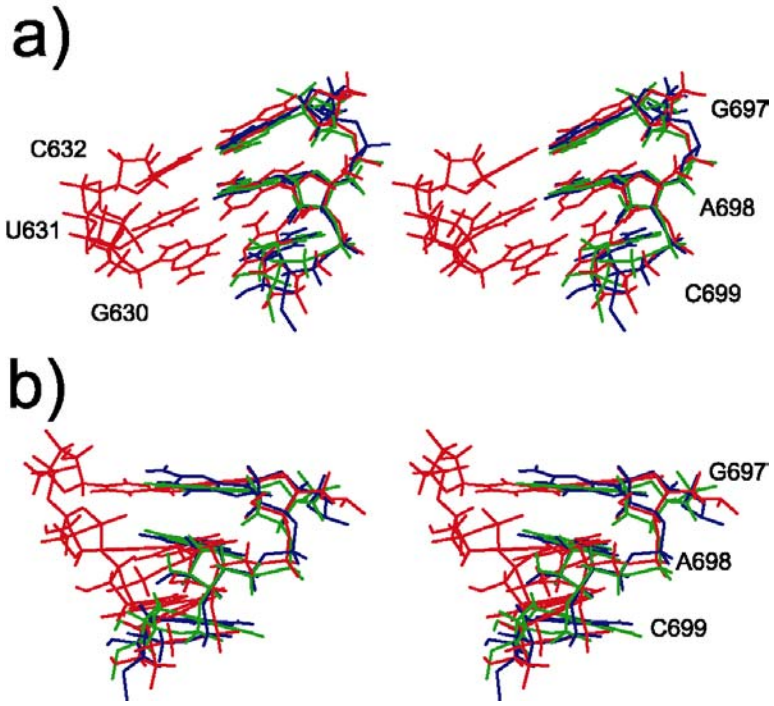


Figure 4.8. Model of base-pairing interactions between stem-loop I and stem-loop V. **a)** Stereo view of the minimized average structure of $SL5^{free}$ (green) and $SL5^{Mg}$ (blue) superposed to the structure of a three base-paired A-form helix (red) formed between a 5'-GAC-3' and a 5'-CUG-3' sequence. The superposition was obtained by pairwise minimization of the heavy atom RMSD between the bases of the 5'-GAC-3' sequences (1.66 Å for $SL5^{free}$ and 1.30 for $SL5^{Mg}$). Only residues G697, A698, and C699 are shown for $SL5^{free}$ and $SL5^{Mg}$. **b)** Same as **a)** but shown at a different angle.



REFERENCES

1. Collins, R.A. and Olive, J.E. (1993) Reaction conditions and kinetics of self-cleavage of a ribozyme derived from *Neurospora* VS RNA. *Biochemistry*, **32,** 2795-2799.

2. Saville, B.J. and Collins, R.A. (1990) A site-specific self-cleavage reaction performed by a novel RNA in *Neurospora* mitochondria. *Cell*, **61,** 685-696.

3. Collins, R.A. (2002) The *Neurospora* Varkud satellite ribozyme. *Biochem Soc Trans*, **30,** 1122-1126.

4. Lilley, D.M. (1999) Structure, folding and catalysis of the small nucleolytic ribozymes. *Curr Opin Struct Biol*, **9,** 330-338.

5. Guo, H.C. and Collins, R.A. (1995) Efficient trans-cleavage of a stem-loop RNA substrate by a ribozyme derived from *Neurospora* VS RNA. *EMBO J*, **14,** 368-376.

6. Jones, F.D., Ryder, S.P. and Strobel, S.A. (2001) An efficient ligation reaction promoted by a Varkud Satellite ribozyme with extended 5'- and 3'-termini. *Nucleic Acids Res*, **29,** 5115-5120.

7. Lafontaine, D.A., Wilson, T.J., Norman, D.G. and Lilley, D.M. (2001) The A730 loop is an important component of the active site of the VS ribozyme. *J Mol Biol*, **312,** 663-674.

8. Hiley, S.L., Sood, V.D., Fan, J. and Collins, R.A. (2002) 4-thio-U cross-linking identifies the active site of the VS ribozyme. *EMBO J*, **21,** 4691-4698.

9. Lafontaine, D.A., Norman, D.G. and Lilley, D.M. (2002) Folding and catalysis by the VS ribozyme. *Biochimie*, **84,** 889-896.

10. Lilley, D.M. (2004) The Varkud satellite ribozyme. *RNA*, **10,** 151-158.

11. Andersen, A.A. and Collins, R.A. (2000) Rearrangement of a stable RNA secondary structure during VS ribozyme catalysis. *Mol Cell*, **5,** 469-478.

12. Andersen, A.A. and Collins, R.A. (2001) Intramolecular secondary structure rearrangement by the kissing interaction of the *Neurospora* VS ribozyme. *Proc Natl Acad Sci U S A*, **98,** 7730-7735.

13. Flinders, J. and Dieckmann, T. (2001) A pH controlled conformational switch in the cleavage site of the VS ribozyme substrate RNA. *J Mol Biol*, **308,** 665-679.

14. Hoffmann, B., Mitchell, G.T., Gendron, P., Major, F., Andersen, A.A., Collins, R.A. and Legault, P. (2003) NMR structure of the active conformation of the Varkud satellite ribozyme cleavage site. *Proc Natl Acad Sci U S A*, **100,** 7003-7008.

15. Maguire, J.L. and Collins, R.A. (2001) Effects of cobalt hexammine on folding and self-cleavage of the *Neurospora* VS ribozyme. *J Mol Biol*, **309,** 45-56.

16. Murray, J.B., Seyhan, A.A., Walter, N.G., Burke, J.M. and Scott, W.G. (1998) The hammerhead, hairpin and VS ribozymes are catalytically proficient in monovalent cations alone. *Chem Biol*, **5,** 587-595.

17. Sood, V.D., Beattie, T.L. and Collins, R.A. (1998) Identification of phosphate groups involved in metal binding and tertiary interactions in the core of the *Neurospora* VS ribozyme. *J Mol Biol*, **282,** 741-750.

18. Rastogi, T., Beattie, T.L., Olive, J.E. and Collins, R.A. (1996) A long-range pseudoknot is required for activity of the *Neurospora* VS ribozyme. *EMBO J*, **15,** 2820-2825.

19. Sood, V.D., Yekta, S. and Collins, R.A. (2002) The contribution of 2'-hydroxyls to the cleavage activity of the *Neurospora* VS ribozyme. *Nucleic Acids Res*, **30,** 1132-1138.

20. Quigley, G.J. and Rich, A. (1976) Structural domains of transfer RNA molecules. *Science*, **194,** 796-806.

21. Doudna, J.A. (1995) Hammerhead ribozyme structure: U-turn for RNA structural biology. *Structure*, **3,** 747-750.

22. Lebars, I., Yoshizawa, S., Stenholm, A.R., Guittet, E., Douthwaite, S. and Fourmy, D. (2003) Structure of 23S rRNA hairpin 35 and its interaction with the tylosin-resistance methyltransferase RlmAII. *EMBO J*, **22,** 183-192.

23. Stallings, S.C. and Moore, P.B. (1997) The structure of an essential splicing element: stem loop IIa from yeast U2 snRNA. *Structure*, **5,** 1173-1185.

24. Puglisi, E.V. and Puglisi, J.D. (1998) HIV-1 A-rich RNA loop mimics the tRNA anticodon structure. *Nature Struct Biol*, **5,** 1033-1036.

25. Gutell, R.R., Cannone, J.J., Konings, D. and Gautheret, D. (2000) Predicting U-turns in ribosomal RNA with comparative sequence analysis. *J Mol Biol*, **300,** 791-803.

26. Franch, T. and Gerdes, K. (2000) U-turns and regulatory RNAs. *Curr Opin Microbiol*, **3,** 159-164.

27. Campbell, D.O. and Legault, P. (2004) NMR structure of the VS ribozyme stem-loop V RNA and magnesium-ion binding from chemical-shift mapping. *Journal of Molecular Biology***,** Submitted.

28. Ashraf, S.S., Ansari, G., Guenther, R., Sochacka, E., Malkiewicz, A. and Agris, P.F. (1999) The uridine in "U-turn": contributions to tRNA-ribosomal binding. *RNA*, **5,** 503-511.

29. Vuister, G.W. and Bax, A. (1992) Resolution enhancement and spectral editing of uniformly ^{13}C-enriched proteins by homonuclear broadband ^{13}C decoupling. *J Magn Reson*, **98,** 428-435.

30. Santoro, J. and King, G.C. (1992) A constant-time 2D overbodenhausen experiment for inverse correlation of isotopically enriched species. *J Magn Reson*, **97,** 202-207.

31. Pardi, A. and Nikonowicz, E.P. (1992) Simple procedure for resonance assignment of the sugar protons in ^{13}C-labeled RNAs. *J Am Chem Soc*, **114,** 9202-9203.

32. Ikura, M., Kay, L.E., Tschudin, R. and Bax, A. (1990) Three-dimensional NOESY-HMQC spectroscopy of a ^{13}C-labeled protein. *J Magn Reson*, **86,** 204-209.

33. Kay, L.E., Keifer, P. and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J Am Chem Soc*, **114,** 10663-10665.

34. Skleñár, V., Peterson, R.D., Rejante, M.R. and Feigon, J. (1994) Correlation of nucleotide base and sugar protons in a ^{15}N-labeled HIV-1 RNA oligonucleotide by ^{1}H-^{15}N HSQC experiments. *J Biomol NMR*, **4,** 117-122.

35. Skleñár, V., Miyashiro, H., Zon, G., Miles, H.T. and Bax, A. (1986) Assignment of the ^{31}P and ^{1}H resonances in oligonucleotides by two-dimensional NMR spectroscopy. *FEBS Lett*, **208,** 94-98.

36. Mueller, L., Legault, P. and Pardi, A. (1995) Improved RNA structure determination by detection of NOE contacts to exchange-broadened amino protons. *J Am Chem Soc*, **117,** 11043-11048.

37. Dingley, A.J. and Grzesiek, S. (1998) Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide (2)J(NN) couplings. *J Am Chem Soc*, **120,** 8293-8297.

38. 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.

39. Schwalbe, H., Marino, J.P., King, G.C., Wechselberger, P., Bermel, W. and Griesinger, C. (1994) Determination of a complete set of coupling constants in ^{13}C-labeled oligonucleotides. *J Biomol NMR*, **4,** 631-644.

40. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) *J Biomol NMR*, **6,** 277-293.

41. Johnson, B.A. and Blevins, R.A. (1994) NMRView - A computer program for the visualization and analysis of NMR data. *J Biomol NMR*, **4,** 603-614.

42. Wishart, D.S., Bigam, C.G., Yao, J., Dyson, H.J., Oldfield, E., Markley, J.L. and Sykes, B.D. (1995) ^{1}H, ^{13}C, ^{15}N Chemical shift referencing in biomolecular NMR. *J Biomol NMR*, **6,** 135-140.

43. Roberts, G.C.K. (1993) *NMR of macromolecules : a practical approach*. IRL Press at Oxford University Press, Oxford England ; New York.

44. Brunger, A.T. (1992) *X-PLOR : version 3.1 : a system for x-ray crystallography and NMR*. Yale University Press, New Haven, Conn.

45. Schwieters, C.D., Kuszewski, J.J., Tjandra, N. and Clore, G.M. (2003) The Xplor-NIH NMR molecular structure determination package. *J Magn Reson*, **160,** 66-74.

46. Kleywegt, G.J. and Jones, T.A. (1998) Databases in protein crystallography. *Acta Crystallogr D Biol Crystallogr*, **54,** 1119-1131.

47. Koradi, R., Billeter, M. and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graphics*, **14,** 51-55.

48. Wijmenga, S.S. and van Buuren, B.N.M. (1998) The use of NMR methods for conformational studies of nucleic acids. *Prog NMR Spectr*, **32,** 287-387.

49. Gonzalez, R.L., Jr. and Tinoco, I., Jr. (2001) Identification and characterization of metal ion binding sites in RNA. *Methods Enzymol*, **338,** 421-443.

50. Cowan, J.A. (1993) Metallobiochemistry of RNA. $Co(NH_3)_6^{3+}$ as a probe for Mg^{2+}(aq)-binding sites. *J Inorg Biochem*, **49,** 171-175.

51. Feig, A.L. (2000) The use of manganese as a probe for elucidating the role of magnesium ions in ribozymes. *Met Ions Biol Syst*, **37,** 157-182.

52. Gonzalez, R.L., Jr. and Tinoco, I., Jr. (1999) Solution structure and thermodynamics of a divalent metal ion binding site in an RNA pseudoknot. *J Mol Biol*, **289,** 1267-1282.

53. Butcher, S.E., Allain, F.H. and Feigon, J. (2000) Determination of metal ion binding sites within the hairpin ribozyme domains by NMR. *Biochemistry*, **39,** 2174-2182.

54. 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.

55. Misra, V.K. and Draper, D.E. (1998) On the role of magnesium ions in RNA stability. *Biopolymers*, **48,** 113-135.

56. Bertini, I. and Luchinat, C. (1986) *NMR of paramagnetic molecules in biological systems*. Benjamin/Cummings Pub. Co., Menlo Park, Calif.

57. Butcher, S.E., Dieckmann, T. and Feigon, J. (1997) Solution structure of the conserved 16 S-like ribosomal RNA UGAA tetraloop. *J Mol Biol*, **268,** 348-358.

58. Ogle, J.M., Brodersen, D.E., Clemons, W.M., Jr., Tarry, M.J., Carter, A.P. and Ramakrishnan, V. (2001) Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science*, **292,** 897-902.

59. Silvian, L.F., Wang, J. and Steitz, T.A. (1999) Insights into editing from an ile-tRNA synthetase structure with tRNAile and mupirocin. *Science*, **285,** 1074-1077.

60. Tzokov, S.B., Murray, I.A. and Grasby, J.A. (2002) The role of magnesium ions and 2'-hydroxyl groups in the VS ribozyme-substrate interaction. *J Mol Biol*, **324,** 215-226.

61. von Ahsen, U., Green, R., Schroeder, R. and Noller, H.F. (1997) Identification of 2'-hydroxyl groups required for interaction of a tRNA anticodon stem-loop region with the ribosome. *RNA*, **3,** 49-56.

SUPPLEMENTARY MATERIAL

Figure 4.S1. Effect of $MgCl_2$ and $Co(NH_3)_6^{3+}$ on SL5 RNA. **a)** Overlay of the H6-C6/H8-C8 region of 2D ^{1}H-^{13}C CT-HSQC spectra of SL5 at selected points of the $MgCl_2$ titration in NMR buffer B. Spectra are shown for 0 mM (black), 1.0 mM (red), 2.0 mM (yellow), 3.0 mM (61), 5.0 mM (blue), 15 mM (magenta) and 46 mM (cyan). Peak labels are adjacent to the last point of the $MgCl_2$ titration. **b)** Overlay of H6-C6/C8-H8 region of 2D ^{1}H-^{13}C CT-HSQC spectra of $SL5^{Mg}$ recorded for two sets of buffer conditions: in NMR buffer A supplemented with 40 mM $MgCl_2$ (red) and in NMR buffer B supplemented with 20 mM $MgCl_2$ (blue). **c)** Overlay of the H6-C6/H8-C8 region of ^{1}H-^{13}C CT-HSQC spectra of SL5 at selected points of the $Co(NH_3)_6^{3+}$ titration in NMR Buffer A. Spectra are shown for 0 mM (black), 0.5 mM (red), 1.0 mM (green), 2.0 mM (blue), 5.0 mM (yellow), 6.0 mM (magenta). Peak labels are adjacent to the last point of the $Co(NH_3)_6^{3+}$ titration. The position of the peaks from the $MgCl_2$ titration in NMR buffer A at 40 mM $MgCl_2$ are shown in cyan (27). All spectra were collected at in 100% D_2O at 25 ˚C on a 600 MHz NMR spectrometer.

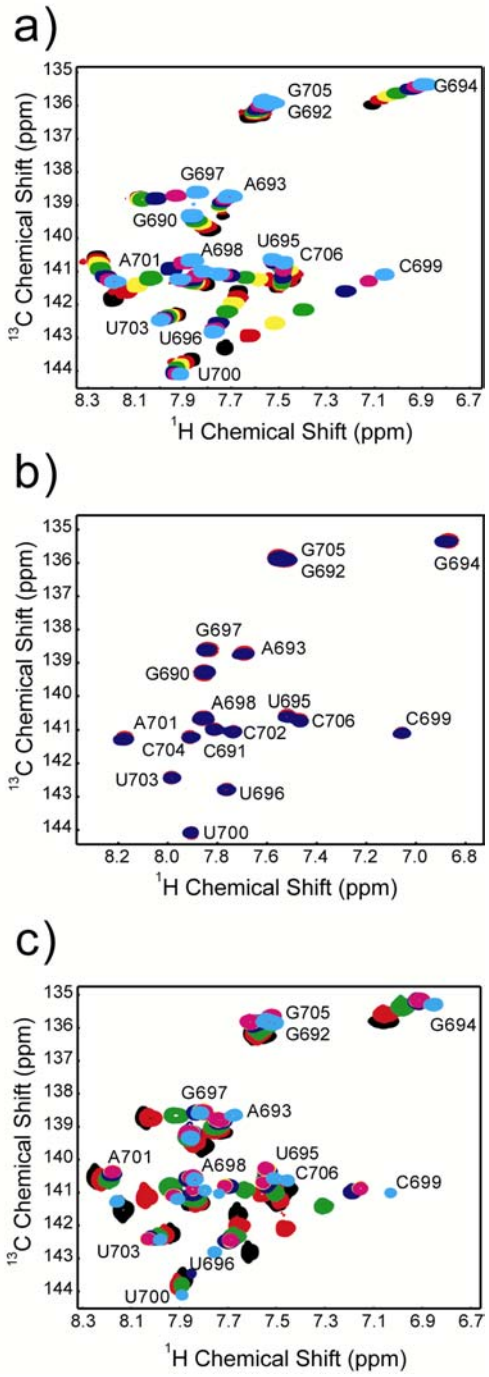

a)
b)
c)
^{13}C Chemical Shift (ppm)
^{1}H Chemical Shift (ppm)
G705
G692
G694
G697
A693
G690
A698
U695
C706
C699
A701
U703
U696
U700
C702
C704
C691

Table 4.S1. Resonance assignment of $SL5^{Mg}$.

NON-EXCHANGEABLE PROTON CHEMICAL SHIFTS (ppm)[a] (ppm +/- 0.09)										
	H1'	**H2'**	**H3'**	**H4'**	**H5'**	**H5''**	**H2**	**H5**	**H6**	**H8**
G690	5.48	4.51	4.42	4.08	3.84	3.76				7.84
C691	5.56	4.46	4.56	4.42	4.55	4.10		5.13	7.81	
G692	5.70	4.55	4.65	4.46	4.54	4.11				7.54
A693	5.89	4.63	4.57	4.50	4.56	4.10	7.47			7.69
G694	5.52	4.47	4.19	4.42	4.45	4.00				6.87
U695	5.37	4.23	4.34	4.34	4.41	4.08		5.10	7.51	
U696	5.56	4.70	4.60	4.27	4.38	4.10		5.64	7.77	
G697	5.53	4.78	4.61	4.32	4.33	4.07				7.86
A698	5.85	4.30	4.91	4.37	4.35	4.04	7.96			7.86
C699	5.53	4.34	4.66	4.39	4.46	4.01		5.04	7.06	
U700	6.06	4.66	4.76	4.53	4.24	4.23		5.85	7.91	
A701	5.95	4.90	4.74	4.56	4.57	4.27	7.91			8.18
C702	5.41	4.20	4.55	4.46	4.56	4.11		5.31	7.74	
U703	5.55	4.46	4.60	4.45	4.57	4.13		5.33	7.98	
C704	5.61	4.45	4.57	4.47	4.55	4.12		5.62	7.91	
G705	5.65	4.39	4.56	4.42	4.56	4.07				7.55
C706	5.59	3.97	4.04	4.15	4.51	4.02		5.23	7.47	
CARBON AND PHOSPHORUS CHEMICAL SHIFTS (ppm) (ppm +/- 0.8)										
	C1'	**C2'**	**C3'**	**C4'**	**C5'**	**C2**	**C5**	**C6**	**C8**	**5' P[b]**
G690	93.3	75.0	73.4	84.4	61.6				138.7	
C691	93.7	75.4	71.9	81.9	64.3		97.2	141.1		
G692	92.5	75.6	72.4	81.6	64.8				136.0	
A693	92.6	75.7	72.5	81.7	64.7	153.6			138.8	
G694	92.8	75.4	72.7	81.9	65.3				135.4	
U695	93.8	75.9	72.0	82.5	64.1		103.5	140.7		
U696	94.1	74.9	71.9	82.3	64.6		103.5	142.9		
G697	93.6	75.1	72.2	82.9	64.1				139.4	2.03
A698	92.6	76.2	72.5	82.7	64.4	154.9			140.8	1.00
C699	93.3	75.8	75.0	83.3	66.5		98.0	141.2		-1.09
U700	90.5	75.0	77.4	86.1	67.4		105.5	144.2		
A701	93.9	75.3	72.9	83.7	65.2	154.5			141.4	
C702	94.1	75.5	71.9	82.0	64.3		97.2	141.2		
U703	93.6	75.3	72.0	81.6	64.5		102.8	142.5		
C704	93.6	75.4	72.2	81.7	64.5		97.9	141.3		
G705	92.9	75.4	72.3	81.9	64.7				136.0	
C706	93.3	77.4	69.2	83.0	64.8		97.5	140.8		

BASE EXCHANGEABLE PROTON CHEMICAL SHIFTS (ppm)								
RESIDUES	H1[b]	H3[b]	H41[b]	H42[b]	H21[c]	H22[c]	H61[c]	H62[c]
G690								
C691			8.42	6.37				
G692	11.98							
A693								
G694	13.62							
U695								
U696								
G697					6.08	6.08		
A698							6.38	6.38
C699								
U700								
A701								
C702			8.23	6.92				
U703		14.07						
C704			8.22	6.47				
G705	12.98							
C706			8.19	6.67				
BASE NITROGEN CHEMICAL SHIFTS (ppm)								
RESIDUES	N1[d]	N2[d]	N3[e]	N4[d]	N6[d]	N7[d]	N9[d]	
G690						231.1	169.6	
C691				98.1				
G692	145.9						169.7	
A693	221.0		212.8			229.1	171.1	
G694	148.4					234.1	169.8	
U695			159.11					
U696								
G697		72.5			81.35	232.7	169.4	
A698	228.6		214.1			223.1	171.0	
C699								
U700								
A701	225.6		215.0			230.3	169.7	
C702				98.6				
U703			161.6					
C704				97.8				
G705	147.5					234.3	169.4	
C706				97.7				

a. Assignments of non-exchangeable ^{1}H, ^{13}C, and ^{31}P were obtained in NMR buffer A supplemented with 40 mM $MgCl_2$ and assignments of exchangeable ^{1}H and ^{15}N were obtained in NMR buffer B supplemented with 20 mM $MgCl_2$;
b. +/- 0.04 ppm;
c. +/- 0.02 ppm;
d. +/- 0.2 ppm;
e. +/- 0.2 ppm for adenines, 0.4 ppm for uridines.

Table 4.S2. Resolved SL5 protons that gave an NOE crosspeak to $Co(NH_3)_6^{3+}$ in ^{13}C-edited HMQC-NOESY spectra.

NOE observed at a mixing time of 90 ms	NOE observed only at a mixing time of 180 ms
690 H8	690 H4'
691 H41, H42, H5	692.H8
693 H8	694 H1'
694 H1, H3', H8	695 H6
695 H5	702 H1'
696 H5	703 H1', H5
702 H41, H42	706 H5
703 H3	
704 H41, H42	
706 H41, H42	

Table 4.S3. Results of $MnCl_2$ titration.

Nuclei and nuclei pairs that give rise to peaks broadened to baseline at 20 μM $MnCl_2$	Nuclei and nuclei pairs that give rise to peaks with little or no line-broadening at 80 μM $MnCl_2$
690 C8-H8, H8-N7, H8-N9	693 C1'-H1', C2-H2
691 N4-H41, N4-H42, C5-H5	694 C1'-H1'
692 H8-N7, H8-N9	695 C1'-H1'
693 C8-H8, H8-N7, H8-N9	698 C1'-H1', C2'-H2', C2-H2, C4'-H4'
694 C8-H8, H8-N7, H8-N9	699 C5-H5, C6-H6
695 C5-H5	700 C5-H5, C6-H6
696 C5-H5, C6-H6	701 C1'-H1' C2-H2
697 C8-H8, H8-N7, H8-N9, 5'-P	702 C1'-H1', C2'-H2'
698 H8-N7, H8-N9, 5'-P	703 C6-H6
699 5'-P	
701 H8-N7, H8-N9	
702 N4-H41, N4-H42	
704 N4-H41, N4-H42	
705 H8-N7, H8-N9	
706 N4-H41, N4-H42	

CHAPTER 5

CONCLUSIONS

Summary of the results of study

SL5 forms a loose U-turn motif in the absence of Mg^{2+}

The three-dimensional high-resolution structure of stem-loop V of the *Neurospora* VS ribozyme was solved by NMR spectroscopy (1). The structure of SL5 was shown to be a hairpin with a five-member terminal loop. The terminal loop of SL5 contained characteristics that resembled a U-turn motif. The U-turn motif in SL5 is not as compact as other characterized U-turn motifs (1). We presumed that the reason for the loose conformation of the U-turn motif in SL5 was due to the presence of a non Watson-Crick base pair being adjacent to the UNR sequence of the U-turn motif.

Mg^{2+} stabilizes the U-turn in SL5

The high-resolution solution structure of SL5 in the presence of magnesium ions, $SL5^{Mg}$, was solved by NMR spectroscopy (2). Inspection of the structure showed that the U-turn motif in $SL5^{Mg}$, however, is more compact than the one found in $SL5^{free}$ (SL5 in the absence of Mg^{2+} ions). The U-turn motif was stabilized as the result of a conformational change caused by the binding of four magnesium ions to the loop of SL5.

The U-turn in stem-loop V may facilitate binding to stem-loop I

The U-turn motif in SL5 may facilitate formation of an important loop-loop interaction between SL5 and stem-loop I (SL1) (1,2). The U-turn motif achieves this by making the residues of SL5 involved in loop-loop interaction available by exposing their Watson-Crick faces to the solvent.

Impact of research and future directions

A more general definition of the U-turn motif

There have been seven characteristics of the U-turn motif that have been published (3,4). Our research has shown that there are U-turn motifs in the structural database that do not fit all the published characteristics of the U-turn motif (1,2). What we have shown is that the presence of a non Watson-Crick base pair 5'of the UNR sequence and the presence of a hydrogen bond between the U H3 and the R 3'-P are not necessary for the formation of a U-turn motif. U-turn motifs without these two characteristics do not have a conformation that is as compact as those that do, however. These "loose U-turns" however, still carry out the structural function of the U-turn, which is to facilitate the participation of the bases 3' of the turn in tertiary interactions (4). In loose U-turn motifs the bases 3' of the turn have their Watson-Crick faces exposed to the solvent, as is the case in canonical U-turn motifs. As we have shown, the presence of the Watson-Crick base pair 5' of the UNR sequence destabilizes the U-turn motif. The hydrogen bond between the U H3 and the R 3'-P is a stabilizing interaction for the U-turn, so its absence will also destabilize the U-turn. What we have shown is that multivalent metal ions can provide the stabilizing force to counteract these two destabilizing effects (2) (unpublished data). So our research shows that the presence of a non Watson-Crick base pair 5' of the UNR sequence and

the presence of a hydrogen bond between the U H3 and the R 3'-P are important but not necessary for the definition of U-turn motifs.

A more comprehensive study of U-turn motifs in RNA

We have developed the tools to search the structural database for occurrences of U-turn motifs in order to compare them with the U-turn motif found in SL5. We have found several instances of uncharacterized U-turn motifs (unpublished data). These U-turn motifs are very diverse in their sequence, loop size, metal ions that are bound, and the types of tertiary interactions that they are involved in. To date there has been no reviews published on U-turn motifs involving a study of these factors, only one on predicting the existence of U-turns based on sequence analysis (4). We actually have these structures on hand and are in a position to do such a comprehensive review on U-turn motifs. We can gain insight into the numerous factors that may affect the formation of U-turn motifs, such as the size of the loop containing the motif and role of multivalent metal ion binding. We can also gain more information on the type of tertiary interactions that U-turn motifs are involved in and any possible role that these interactions have in stabilizing the U-turn motif. Also we can obtain information as to the extent to which U-turn motifs play vital roles in the functions of biologically important RNA molecules.

Role of multivalent metal ions in RNA solution structure determination

During the course of our investigation we found that a U-turn motif was postulated to be in the terminal loop of stem-loop I (5,6). However, a solution structure of the wild-type substrate was solved and terminal loop of SL1 was found to be disordered (7). By analogy to the results we obtained from our study, we postulate that reason the terminal loop of SL1 was unstructured is

due to the fact that there were no multivalent metal ions present in the NMR buffer. Another reason may be that SL1 has been postulated to have a Watson-Crick base pair 5' to the UNR sequence in its terminal loop, similar to the case seen in SL5. We have given examples in which the presence of a Watson-Crick base pair 5' of the UNR sequence hinders the formation of a canonical U-turn motif in terminal loops (1). In the case of SL5 we showed that the addition of magnesium ions to the NMR buffer allowed the terminal loop of SL5 to adopt a compact U-turn conformation (1). Other multivalent ions have been shown to stabilize U-turn motifs such as the case in which the addition of $Co(NH_3)_6^{3+}$ resulted in the stabilization of the U-turn motif in an anticodon stem-loop RNA (8).

Multivalent metal ions are typically not added to NMR buffers for the structure determination of an RNA molecule. This is due to the fact that there can be a reduction in the signal to noise ratio of the NMR signals. This reduction in signal to noise can be a result of non-specific line broadening caused by the metal ions and/or aggregation or precipitation of the RNA molecule. However, in many cases the presence of multivalent metal ions may be relevant to the structure and/or activity of the RNA molecule that is being studied. This becomes even more important if the RNA that is being studied contains structural elements that are known to be capable of binding metal ions. It is possible that the full story is not being told if the solution structures of such RNA molecules are solved in the absence of multivalent ions.

Our research suggests that in cases where the RNA molecule is known to have specific structural elements, such as a motif, that is known to bind multivalent metal ions in a specific manner, the binding of multivalent metal ions to the RNA should always be investigated. This could be carried out by performing chemical shift mapping and if feasible, solving the structure of the RNA in the absence and presence of the multivalent metal ions. The insight we will gain

by increasing the frequency of such investigations is that it will give a better understanding of the role of multivalent metal ions in stabilizing the structure of RNA molecules in general, and metal-ion binding motifs specifically.

The structure of the loop-loop interaction

We have determined the structure of one of the participants in the loop-loop interaction that is important for the activity of the VS ribozyme. We have already shown that a U-turn motif is responsible for facilitating the participation of SL5 in the loop-loop interaction. It would be interesting to see if this is also the case in stem-loop I as it has been postulated to posses a U-turn motif in solution (5,6). Small RNA molecules that reproduce the loop-loop interaction in solution have already been synthesized and purified (unpublished data). The structure of the loop-loop interaction could therefore be determined by NMR spectroscopy. The magnesium-ion binding sites of the loop-loop interaction could also be determined. This would give a better understanding of the role of magnesium ions in formation of the loop-loop interaction. A structure of the loop-loop interaction would give further information as to whether the extruded residues U700 from SL5 and C634 from SL1 are involved in any tertiary interactions. The information gained from examining the structure of the loop-loop interaction would no doubt further our understanding of its important role in the substrate recognition of the VS ribozyme.

References

1. Campbell, D.O. and Legault, P. (2004) NMR structure of the VS ribozyme stem-loop V RNA and magnesium-ion binding from chemical-shift mapping. *Journal of Molecular Biology***,** Submitted.

2. Campbell, D.O. and Legault, P. (2004) NMR structure of the VS ribozyme stem-loop V in the presence of magnesium and localization of the divalent metal ion binding sites. *To be submitted*.

3. Quigley, G.J. and Rich, A. (1976) Structural domains of transfer RNA molecules. *Science*, **194,** 796-806.

4. Gutell, R.R., Cannone, J.J., Konings, D. and Gautheret, D. (2000) Predicting U-turns in ribosomal RNA with comparative sequence analysis. *J Mol Biol*, **300,** 791-803.

5. Rastogi, T., Beattie, T.L., Olive, J.E. and Collins, R.A. (1996) A long-range pseudoknot is required for activity of the *Neurospora* VS ribozyme. *EMBO J*, **15,** 2820-2825.

6. Sood, V.D., Yekta, S. and Collins, R.A. (2002) The contribution of 2'-hydroxyls to the cleavage activity of the *Neurospora* VS ribozyme. *Nucleic Acids Res*, **30,** 1132-1138.

7. Flinders, J. and Dieckmann, T. (2001) A pH controlled conformational switch in the cleavage site of the VS ribozyme substrate RNA. *J Mol Biol*, **308,** 665-679.

8. Cabello-Villegas, J., Tworowska, I. and Nikonowicz, E.P. (2004) Metal ion stabilization of the U-turn of the A37 N6-dimethylallyl-modified anticodon stem-loop of Escherichia coli tRNA(Phe). *Biochemistry*, **43,** 55-66.