

Université de Montréal

**Études structurales par résonance magnétique nucléaire
du ribozyme VS de *Neurospora***

par

Éric Bonneau

Département de biochimie et médecine moléculaire

Faculté de Médecine

Thèse présentée à la Faculté de Médecine
en vue de l'obtention du grade de Philosophiae doctor (Ph.D)
en Biochimie

21 janvier 2015

© Éric Bonneau, 2015

Université de Montréal
Faculté des études supérieures et postdoctorales

Cette thèse intitulée :

Études structurales par résonance magnétique nucléaire du ribozyme VS de *Neurospora*

Présenté par :
Éric Bonneau

a été évaluée par un jury composé des personnes suivantes :

Léa Brakier-Gingras, président-rapporteur
Pascale Legault, directeur de recherche
Katherine Borden, membre du jury
Nancy Greenbaum, examinateur externe
Dindial Ramotar, représentant du doyen

Résumé

Le ribozyme VS de *Neurospora* catalyse des réactions de clivage et de ligation d'un lien phosphodiester spécifique essentielles à son cycle de réplication. Il est formé de six régions hélicales (I à VI), qui se divisent en deux domaines, soit le substrat (SLI) et le domaine catalytique (tiges II à VI). Ce dernier comprend deux jonctions à trois voies qui permettent de reconnaître le substrat en tige-boucle de façon spécifique. Ce mode de reconnaissance unique pourrait être exploité pour cibler des ARN repliés pour diverses applications. Bien que le ribozyme VS ait été caractérisé biochimiquement de façon exhaustive, aucune structure à haute résolution du ribozyme complet n'a encore été publiée, ce qui limite la compréhension des mécanismes inhérents à son fonctionnement. Précédemment, une approche de *divide-and-conquer* a été initiée afin d'étudier la structure des sous-domaines importants du ribozyme VS par spectroscopie de résonance magnétique nucléaire (RMN) mais doit être complétée.

Dans le cadre de cette thèse, les structures de la boucle A730 et des jonctions III-IV-V et II-III-VI ont été déterminées par spectroscopie RMN hétéronucléaire. De plus, une approche de spectroscopie RMN a été développée pour la localisation des ions divalents, tandis que diverses approches de marquage isotopique ont été implémentées pour l'étude d'ARN de plus grandes tailles. Les structures RMN de la boucle A730 et des deux jonctions à trois voies révèlent que ces sous-domaines sont bien définis, qu'ils sont formés de plusieurs éléments structuraux récurrents (*U-turn*, *S-turn*, triplets de bases et empilement coaxial) et qu'ils contiennent plusieurs sites de liaison de métaux. En outre, un modèle du site actif du ribozyme VS a été construit sur la base des similarités identifiées entre les sites actifs des ribozymes VS et *hairpin*. Dans l'ensemble, ces études contribuent de façon significative à la compréhension de l'architecture globale du ribozyme VS. De plus, elles permettront de construire un modèle à haute résolution du ribozyme VS tout en favorisant de futures études d'ingénierie.

Mots-clés : Ribozyme VS de *Neurospora*, structure de l'ARN, motif *S-turn*, motif *U-turn*, site actif, jonction à trois voies, interaction *kissing-loop* I/V, spectroscopie de résonance magnétique nucléaire (RMN), ions Mg²⁺

Abstract

The *Neurospora* VS ribozyme catalyzes the cleavage and the ligation of a specific phosphodiester bond, which is essential for its replication cycle. It is formed of six helical regions (I to VI) that are divided in two domains: the substrate (SLI) and the catalytic domain (stems II-VI). The latter contains two three-way junctions that allow recognition of the stem-loop substrate in a specific manner. This unique mode of substrate recognition could be exploited to target folded RNAs for diverse applications. Even though the VS ribozyme has been extensively characterized biochemically, no high-resolution structure of the complete ribozyme has been published yet and this limits our mechanistic understanding. A divide-and-conquer approach was previously initiated to study the structure of the important subdomains of the VS ribozyme by nuclear magnetic resonance (NMR), but this approach needs to be completed.

In this thesis, the structures of the A730 loop, the III-IV-V junction and the II-III-VI junction were determined by heteronuclear NMR spectroscopy. Moreover, a unique NMR approach was developed for localizing divalent metal ions, whereas several isotope-labeling strategies were implemented to facilitate the study of large RNA molecules. The NMR structures of the A730 loop and the two three-way junctions reveal that these subdomains are well defined, that they are formed by several recurrent structural elements (U-turn and S-turn motifs, base triples and coaxial stacking) and that they contain several metal-binding sites. Interestingly, structural similarities were identified between the VS and hairpin ribozymes, which allowed the modeling of the VS ribozyme active site. In summary, these studies significantly contribute to a better understanding of the global architecture of the VS ribozyme. In addition, they will allow the construction of a high-resolution model of the complete VS ribozyme and facilitate future engineering studies.

Keywords: *Neurospora* VS ribozyme, RNA structure, S-turn motif, U-turn motif, active site, three-way junction, I/V kissing-loop interaction, nuclear magnetic resonance (NMR) spectroscopy, Mg²⁺ ions

Table des matières

Résumé	i
Abstract	ii
Table des matières	iii
Liste des tableaux	vii
Liste des figures	viii
Liste des abréviations	xi
Remerciements	xiv
1. Introduction	19
1.1. Importance biologique de l'ARN	20
1.2 La structure de l'ARN	21
1.2.1. Les bases structurales de l'ARN	21
1.2.2. Les paires de bases et les hélices d'ARN	23
1.2.3. Les éléments de structure secondaire de l'ARN	25
1.2.3.1. Les boucles terminales	25
1.2.3.2. Renflements et boucles internes	27
1.2.3.3. Les jonctions multivoies	29
1.2.4. Les éléments de structure tertiaire de l'ARN	30
1.2.4.1. Interactions entre des domaines hélicaux	30
1.2.4.2. Interactions entre des domaines hélicaux et des régions non-pairées	31
1.2.4.3. Interactions entre des régions non-pairées	32
1.2.5. Contribution des métaux à la formation de structures tertiaires d'ARN	33
1.3. Les ribozymes naturels	35
1.3.1. L'importance de l'étude des ribozymes	35
1.3.2. Les différentes classes de ribozymes	36
1.3.3. Les petits ribozymes nucléolytiques	40
1.3.3.1. Généralités sur les petits ribozymes nucléolytiques	40
1.3.3.2. Caractéristiques des petits ribozymes nucléolytiques	42
1.4. Le ribozyme VS de <i>Neurospora</i>	45
1.4.1. Le ribozyme VS et son rôle dans le cycle de réplication	45
1.4.2. Caractérisation du ribozyme VS	46
1.4.3. Les domaines fonctionnels du ribozyme VS	47

1.4.4. L'interaction <i>kissing-loop</i> I/V et son rôle dans la reconnaissance et l'activation du substrat	49
1.4.4.1. La reconnaissance du substrat par l'interaction <i>kissing-loop</i> I/V	49
1.4.4.2. Activation du substrat par la formation de l'interaction <i>kissing-loop</i> I/V ...	50
1.4.4.3. Contribution de la jonction III-IV-V	52
1.4.5. Formation du site actif du ribozyme VS	53
1.4.5.1. Les boucles internes de SLI et de SLVI	53
1.4.5.2. Le site actif est formé par l'association de la boucle interne de SLI et de la boucle A730	54
1.4.5.3. Contribution de la jonction II-III-VI	55
1.4.6. Le mécanisme enzymatique du ribozyme VS	55
1.5. Structures tridimensionnelles du ribozyme VS de <i>Neurospora</i>	58
1.5.1. Structures RMN du substrat du ribozyme VS	59
1.5.1.1. Structures RMN de SLI inactif	59
1.5.1.2. Structure RMN de la forme active de SLI	60
1.5.2. Structures RMN de la tige-boucle V	62
1.5.3. Structure RMN de l'interaction <i>kissing-loop</i> I/V	63
1.5.4. Structure RMN de la boucle A730	64
1.5.5. Modèles à basse résolution des jonctions II-III-VI et III-IV-V	65
1.5.6. Modèles à basse résolution du ribozyme VS	67
1.6. La spectroscopie de résonance magnétique nucléaire	69
1.7. Objectifs de la thèse	74
2. Détermination de la structure RMN de la boucle A730 et modélisation du site actif du ribozyme VS de <i>Neurospora</i>	77
Mise en contexte	78
ABSTRACT	81
INTRODUCTION	82
MATERIALS AND METHODS	84
RESULTS	88
DISCUSSION	91
REFERENCES	97
TABLE	104
FIGURES	105
SUPPORTING INFORMATION	112
3. Localisation des ions Mg ²⁺ dans la structure RMN de la boucle A730	117

Mise en contexte	118
ABSTRACT	121
INTRODUCTION.....	122
EXPERIMENTAL PROCEDURES	124
RESULTS	128
DISCUSSION	134
REFERENCES.....	139
TABLES.....	148
FIGURES	152
SUPPORTING INFORMATION	159
4. Détermination de la structure RMN de la jonction III-IV-V et localisation des ions Mg^{2+} par relaxation paramagnétique induite par le Mn^{2+}	164
Mise en contexte	165
ABSTRACT	168
INTRODUCTION.....	169
EXPERIMENTAL PROCEDURES	171
RESULTS	174
DISCUSSION	178
REFERENCES.....	184
TABLE	193
FIGURES	194
SUPPORTING INFORMATION	201
5. Détermination de la structure RMN de la jonction II-III-VI et identification d'une interaction tertiaire critique	219
Mise en contexte	220
ABSTRACT	222
INTRODUCTION.....	223
EXPERIMENTAL PROCEDURES	225
RESULTS	227
DISCUSSION	231
REFERENCES.....	236
TABLE	242
FIGURES	243
SUPPORTING INFORMATION	249
6. Discussion	271

6.1. Retour sur les résultats principaux de l'étude	273
6.1.1. Le site actif du ribozyme VS est pré-organisé en présence d'ions Mg^{2+}	273
6.1.2. La jonction III-IV-V est un domaine architectural qui facilite la formation de l'interaction <i>kissing-loop</i> I/V	275
6.1.3. La jonction II-III-VI requiert la formation d'une interaction tertiaire critique pour se replier.....	276
6.1.4. Contribution à une meilleure compréhension des interactions ARN-métal	278
6.1.5. Implémentation d'approches RMN pour l'étude d'ARN de grande taille.....	279
6.2. Contribution aux bases de données et prédictions de structure d'ARN.....	280
6.3. Nouvelle vision de la structure du ribozyme VS et construction d'un modèle RMN complet à haute résolution	282
6.3.1. L'architecture globale du ribozyme VS	282
6.3.2. Éléments structuraux récurrents du ribozyme VS.....	284
6.3.3. Vers un modèle RMN complet à haute résolution du ribozyme VS.....	286
6.3.4. Obtention de données expérimentales sur le site actif	288
6.3.5. Localisation des ions Mg^{2+} dans le ribozyme VS	289
6.4. Ingénierie du ribozyme VS	290
7. Références	292

Liste des tableaux

Table 1.1. Liste des ribozymes naturels connus, de leur mécanisme enzymatique et de la disponibilité d'une structure tridimensionnelle.....	36
Table 2.1. Structural statistics of the SLVI RNA.....	104
Supplementary Table 2.1. Resonance assignment of SLVI.....	112
Table 3.1. ^{31}P NMR CSP (Δ_p) following Cd^{2+} titrations of phosphorothioate SLVI RNAs.	148
Table 3.2. T_m values for the unmodified and phosphorothioate SLVI RNAs.	149
Table 3.3. Structural statistics of the SLVI RNA with $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ complexes.	150
Table 3.4. Description of $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ -binding sites in SLVI.	151
Table S3.1. Distance restraints derived from Mn^{2+} -induced PRE.....	162
Table S3.2. Results from the WebFR3D motif search.....	163
Table 4.1. Structural statistics for J345 and J345 ^{Mg}	193
Table S4.1. Resonance assignment of exchangeable ^1H and their attached ^{15}N for the J345 RNA.	209
Table S4.2. Resonance assignment of non-exchangeable ^1H and their attached ^{13}C for the J345 RNA.....	212
Table S4.3. Distance restraints derived from Mn^{2+} -induced PRE used for the J345 ^{Mg} structure calculation.	215
Table S4.4. Structural characteristics of $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ -binding sites in J345 ^{Mg}	216
Table 5.1. Structural statistics for J236 and J236 ^{Mg}	242
Table S5.1. Chemical shifts (ppm) of base exchangeable ^1H and their attached ^{15}N for the J236 RNA.....	261
Table S5.2. Chemical shifts (ppm) of non-exchangeable ^1H and their attached ^{13}C for the J236 RNA.....	263
Table S5.3. Distance restraints derived from Mn^{2+} -induced PRE and used for the J236 ^{Mg} structure calculation.	266
Table S5.4. Structural characteristics of $\text{Mg}(\text{H}_2\text{O})_6^{2+}$ -binding sites in J236 ^{Mg}	268

Liste des figures

Figure 1.1. Structure chimique des ribonucléosides standards composant l'ARN.....	22
Figure 1.2. Représentation de divers éléments de structure secondaire.....	23
Figure 1.3. Représentation schématique des faces d'interactions de nucléosides et de deux paires de bases canoniques.....	24
Figure 1.4. Représentations schématiques et structurales de boucles terminales.	26
Figure 1.5. Représentation structurale d'une boucle interne.	29
Figure 1.6. Représentations schématiques et structurales des familles de jonctions à trois voies.	30
Figure 1.7. Représentations schématiques et structurales de triplets de bases.	32
Figure 1.8. Représentations schématiques et structurales d'interactions tertiaires entre des régions non-pairées.	33
Figure 1.9. Représentations des différents types d'interactions entre l'ARN et le Mg^{2+} ..	34
Figure 1.10. Structures secondaires et tridimensionnelles de différents ribozymes naturels.	38
Figure 1.11. Mécanismes réactionnels d'auto-clivage et d'auto-épissage.....	39
Figure 1.12. Structures du site actif de différents ribozymes nucléolytiques.	43
Figure 1.13. Mécanisme de réplication par <i>rolling circle</i> de l'ARN VS.....	46
Figure 1.14. Structures primaire et secondaire du ribozyme VS de <i>Neurospora</i>	48
Figure 1.15. Structures primaire et secondaire du substrat dans les conformations <i>unshifted</i> et <i>shifted</i>	51
Figure 1.16. Mécanisme enzymatique proposé du clivage du ribozyme VS de <i>Neurospora</i>	57
Figure 1.17. Structures RMN du substrat dans la conformation <i>unshifted</i> et inactive.	60
Figure 1.18. Structure RMN du substrat dans la conformation <i>shifted</i> et active.	61
Figure 1.19. Structures RMN de la tige-boucle V.	63
Figure 1.20. Structure RMN de l'interaction <i>kissing-loop</i> I/V.	64
Figure 1.21. Structure RMN de la boucle A730.	65
Figure 1.22. Modèles de FRET des jonctions II-III-VI et III-IV-V.....	66
Figure 1.23. Modèles à basse résolution du ribozyme VS de <i>Neurospora</i>	68
Figure 2.1. The <i>Neurospora</i> VS ribozyme and its A730 internal loop.....	105

Figure 2.2. Stabilization of the A730 loop by Mg ²⁺ ions.	106
Figure 2.3. NMR solution structure of the SLVI RNA fragment.	107
Figure 2.4. Formation of a <i>cis</i> WC/WC G9-A20 base pair in the A730 loop.....	108
Figure 2.5. S-turn motif in the A730 loop of the SLVI RNA.	109
Figure 2.6. Determination of adenine pK _a 's in SLVI.	110
Figure 2.7. Homology modeling of the VS ribozyme active site.....	111
Figure S2.1. Hill plot of the pH titration data for (a) A20 and (b) A8.....	116
Figure 3.1. Primary and secondary structures of the <i>Neurospora</i> VS ribozyme and the SLVI RNA.	152
Figure 3.2. Probing divalent-metal-ion-binding sites in SLVI through Mn ²⁺ -induced PRE.	153
Figure 3.3. Effect of Cd ²⁺ ions on the 1D ³¹ P NMR spectrum of selected SLVI RNAs containing single phosphorothioate modifications.....	154
Figure 3.4. Lowest-energy structure of the SLVI RNA determined in the presence of Mg(H ₂ O) _n ²⁺ complexes.....	155
Figure 3.5. Mg(H ₂ O) ₆ ²⁺ -binding sites associated with sites 2, 4 and 5.	156
Figure 3.6. Mg(H ₂ O) ₅ ²⁺ -binding sites associated with sites 1 and 3.	157
Figure 3.7. Superposition of metal-binding sites in SLVI with PDB structures containing similar metal-binding sites identified using WebFR3D.....	158
Figure S3.1. Effect of Cd ²⁺ ions on SLVI RNA.....	159
Figure S3.2. Effect of the addition of Cd ²⁺ ions on the ³¹ P NMR spectrum of selected SLVI RNAs containing single phosphorothioate modifications.....	160
Figure S3.3. Effect of the addition of Cd ²⁺ on the imino region of the ¹ H NMR spectra of SLVI RNAs containing single phosphorothioate modifications.....	161
Figure 4.1. J345 RNA adopts a single stable structure in the presence of Mg ²⁺ ions.	194
Figure 4.2. Unambiguous assignment of the uridine imino protons from the core of J345 using ¹⁵ N segmental labeling.....	195
Figure 4.3. NMR solution structure of the J345 RNA.	196
Figure 4.4. The cCUUGg tetraloop of the J345 RNA.....	197
Figure 4.5. NMR structure of the core domain of the J345 RNA.....	198
Figure 4.6. Base triples within the core domain of the J345 RNA.	199
Figure 4.7. Mg(H ₂ O) _n ²⁺ -binding sites within the J345 ^{Mg} RNA.	200

Figure S4.1. The J345 RNA adopts a single stable structure in the absence of Mg^{2+} ions that is similar to that formed in the presence of Mg^{2+} ions.	204
Figure S4.2. The S12/S3 complex adopts a structure similar to that of the J345 RNA..	205
Figure S4.3. NMR evidence for a highly organized core and for helical stacking between stems III and IV.....	206
Figure S4.4. A ribose zipper within the core domain of the J345 RNA.	207
Figure S4.5. Additional $Mg(H_2O)_n^{2+}$ -binding sites within the $J345^{Mg}$ RNA.....	208
Figure 5.1. The J236 RNA adopts a stable structure in the presence of Mg^{2+} ions.	243
Figure 5.2. NMR solution structure of the $J236^{Mg}$ RNA.....	244
Figure 5.3. NMR structure of the core domain of $J236^{Mg}$	245
Figure 5.4. Base pairing within the core domain of $J236^{Mg}$	246
Figure 5.5. NMR structure of the II-VI bulge-bulge interaction of the $J236^{Mg}$ RNA.....	247
Figure 5.6. Schematic of the tertiary structure of a <i>cis</i> -cleaving VS ribozyme derived from NMR solution structures of isolated domains.....	248
Figure S5.1. The J236 RNA is unstable in the absence of Mg^{2+} ions and when the A_6 bulge of stem II is modified.	251
Figure S5.2. NMR evidence of the <i>cis</i> WC-WC $A_{10} \bullet G_{53}$ base pair in stem II and of continuous helical stacking between stems II and III.	253
Figure S5.3. NMR solution structure of the J236 RNA.....	254
Figure S5.4. NMR structure of the core domain of J236.....	255
Figure S5.5. Base pair and base triples within the core domain of J236.	256
Figure S5.6. NMR structure of the II-VI bulge-bulge interaction of the J236 RNA.....	257
Figure S5.7. Preferential $Mg(H_2O)_6^{2+}$ -binding sites within the $J236^{Mg}$ RNA.	258
Figure S5.8. NMR evidence of the <i>cis</i> Sugar edge/WC $A_6 \bullet A_{32}$ base pair defining the II-VI bulge-bulge interaction.	260
Figure 6.1. Structure secondaire révisée du ribozyme VS de <i>Neurospora</i>	284
Figure 6.2. Structures RMN pouvant être utilisées pour modéliser la structure complète à haute résolution du ribozyme VS dans une conformation pré-catalytique.	287

Liste des abréviations

δ : déplacement chimique

1D : une dimension

2D : deux dimensions

3D : trois dimensions

5'-PO₄⁻ : 5'-phosphate

5'-PS : 5'-phosphorothiolate

A : adénine

ADN : acide désoxyribonucléique

AMP : adénosine mono-phosphate

ARN : acide ribonucléique

ARNm : ARN messenger

ARNmi : micro-ARN

ARNnc : ARN non-codant

ARNpi : ARN interagissant avec Piwi

ARNr : ARN ribosomal

ARNsf : ARN sous-génomique de flavivirus

ARNsi : petit ARN interférant

ARNsn : petit ARN nucléaire

ARNsno : petit ARN nucléolaire

ARNt : ARN de transfert

C : cytosine

Ca²⁺ : ion calcium

Cd²⁺ : ion cadmium

CLEC2 : *C-type lectin type II*

Co(NH₃)₆³⁺ : cobalt hexamine

COSY : *correlation spectroscopy*

CoTC : *Co-Transcriptional Cleavage*

CPEB3 : *cytoplasmic polyadenylation element binding protein 3*

Da : Dalton

ENU : *N-ethyl-nitrourea* (N-nitroso N-éthylurée)

FRET : *Fluorescence resonance energy transfer* (transfert d'énergie entre molécules fluorescentes)

G : guanine

GlcN6P : glucosamine-6-phosphate

GMP : guanosine mono-phosphate

HBV : virus de l'hépatite B

HDV : virus de l'hépatite D

HMQC : *heteronuclear multiple quantum coherence*

HSQC : *heteronuclear single quantum coherence*

Hz : Hertz

K⁺ : ion potassium

K_d : constante de dissociation

Mg²⁺ : ion magnésium

M : molaire

Mn²⁺ : ion manganèse

N : nucléotide

Na⁺ : ion sodium

NMP : nucléotide mono-phosphate

NOE : *Nuclear Overhauser effect* (effet nucléaire Overhauser)

NOESY : *NOE spectroscopy*

nt : nucléotide

PDB : *Protein Data Bank* (banque de données de protéines)

ppm : partie par million

R : purine

RDC : *residual dipolar coupling* (couplage résiduel dipolaire)

RMN : résonance magnétique nucléaire

rmsd : *root-mean-square deviation* (racine de la moyenne des déviations au carré)

RNAse P : ribonucléase P

SAXS : *small-angle X-ray scattering* (diffusion de rayons X aux petits angles)

SELEX : *systematic evolution of ligands by exponential enrichment* (Évolution systématique de ligand par enrichissement exponentiel)

SL : *stem-loop* (tige-boucle)

S_N2 : substitution nucléophile bimoléculaire

TERC : *telomerase RNA component* (ARN composant la télomérase)

TOCSY : *total correlation spectroscopy*

U : uracil

V : *Varkud*

VIH : virus de l'immunodéficience humaine

VS : *Varkud Satellite*

WC : Watson-Crick

Y : pyrimidine

Remerciements

J'aimerais commencer cette section particulièrement importante par remercier ma directrice de thèse, la Dr. Pascale Legault. J'ai eu l'honneur et la chance de pouvoir travailler sous sa tutelle et de profiter de ses enseignements durant de nombreuses années. Pascale a fait preuve d'énormément de patience et de soutien à mon égard. Je suis passé d'un jeune homme faisant souvent preuve d'un excès d'optimisme et de manque de recul à un scientifique plus posé, plus critique. S'il y a bien une qualité que Pascale m'a transmise, c'est la rigueur scientifique. Elle m'a appris à bien faire les choses, autant techniquement qu'intellectuellement.

Je tiens aussi à te remercier, Pascale, pour ton dévouement. On ne se cachera pas que la dernière année et demie a été particulièrement intense et qu'elle nous a fait faire des cauchemars et de l'angoisse. Malgré tout, tu as su m'appuyer dans mon rythme effréné et au final, ce que nous avons accompli ensemble est incroyable. Tu as été pour moi plus qu'une directrice de thèse, tu as été une partenaire de travail, un des piliers de ma réussite. Bref, je tiens à te remercier du fond du cœur pour avoir cru en moi et pour tout le support que tu m'as accordé durant ces six années de doctorat. Merci!

La beauté de la science, c'est qu'en plus de repousser les limites du savoir, nous avons la chance de rencontrer tellement de gens magnifiques. Parmi ceux-ci, je suis reconnaissant d'avoir eu la chance de côtoyer tous les membres des laboratoires Legault (Pascale, Patricia, Alex, Nico, Alix, Jo Bou, Pierre, Geneviève DT, Julie, Dominique, Élyse, Ao, Vanessa, Mélanie, Aurélie et Geneviève Desjardins) et Omichinski (Jim, Mathieu, Julien, Haytham, Phil, Thomas, Geneviève Arseneault, Paola, Jimmy, Luca, Caroline, Normand, Lauriane, et Ryan). Je vous remercie pour toutes les soirées, les soupers, les bières, les partys, les congrès, les parties de squash ou de badminton et toutes nos pauses café. Je vous remercie de m'avoir écouté quand c'était plus difficile et d'avoir été à mes côtés pour célébrer. Sans vous, tout ce périple aurait été beaucoup plus difficile. Merci de m'avoir accueilli, pour la première fois, par une froide journée de janvier 2008. Vous avez tout de suite su me faire sentir à l'aise, comme dans une nouvelle famille. Ça a été un plaisir de travailler à vos côtés !

En plus d'avoir eu la chance d'être entouré de mes amis du laboratoire, j'ai eu la chance de faire mon doctorat dans un département très actif. Je tiens à remercier tous les membres du conseil d'administration de l'association des étudiants et étudiantes aux cycles supérieurs en biochimie de l'Université de Montréal (AECSBUM) dont j'ai fais parti de 2009 à 2013. Merci entre autre à vous tous, Étienne, Éric, Samuel, Sébastien, Alex, Mathieu Vernier, Lian, Emmanuelle, Camille, Xavier, Stéphanie, Mathieu Coinçon, Jean-Sébastien et Patricia. Avoir eu la chance de participer et de contribuer à la vie étudiante et à l'atmosphère du département de biochimie à vos côtés fut un honneur. Je tiens aussi à remercier tous les amis du département (Geneviève Huot, Mathieu Coincon, Mathieu Vernier, Franck Galardo, Xavier, Stéphanie Gravel, Nicolas Montpas et Paul) ainsi que tout le personnel de soutien, particulièrement Sylvie Beauchemin, Philippe Lampron et Valérie De Rop, ainsi que les membres de mon comité de thèse, les Dr Desgroseillers et Sygusch. Avoir eu la chance de me développer dans un tel environnement fût une réelle chance. Merci à vous tous.

Faire un doctorat, ce n'est pas toujours facile et ça demande parfois du soutien extérieur. J'ai la chance d'avoir les meilleurs amis que quelqu'un puisse espérer, des gens extraordinaires, avec le cœur sur la main, qui sont toujours prêts à écouter et qui sont toujours là, les uns pour les autres. Sans ces gens là, sans leur support, je ne serais pas là où j'en suis aujourd'hui. Je tiens à vous remercier haut et fort pour votre présence et ce, même si j'ai dû m'absenter souvent par excès de travail. Merci à vous tous Martin Leduc-Poirier, Guillaume Harpin, Simon B. Lavallée, Danny Racine, Simon Lachance, Michaël Demers, Philippe Lamarre, Louis-Étienne Jolin, Anthony Belisle, Vincent Gendreau, Maxime Blanchard, Olivier Houle, Valérie Trudel, Jonathan Gagnon, Louis-Michel Perrier, Christophe Bundock De Muri, et Maxime Benoit. Sans votre support, sans votre amour inconditionnel, cette longue aventure aurait été très difficile.

J'ai aussi eu la chance d'avoir le support de ma famille à partir du premier jour de tout ce long périple. Ils m'ont toujours appuyés, ils m'ont écouté, ils ont lu certains de mes textes et ils ont corrigés mon français. Ce sont des gens dévoués qui ont crus en moi. Merci à ma mère, Francine, à mon père, Yves et à mon frère Marc-André. Merci à vous pour votre soutien particulièrement vif, ça a été une des clés de mon succès. Merci aussi à mes grands-parents, Jean-Paul, Céline, Ghislain et Yvette, ainsi qu'à mon oncle Jean-Yves et ma tante

Chantale ainsi qu'à ma belle-famille Lise, Sylvain et Pascal. Parfois, sans que vous ne le sachiez, vos mots d'encouragements et vos questionnements sur mon cheminement ont su me redonner courage et détermination. Un énorme merci!

Depuis environ cinq ans, je partage aussi ma vie avec une femme fantastique, Jessica Langlois. Son support n'a jamais faibli durant toutes ces années. Elle m'a supporté dans tous les moments; autant dans les creux de vague, où elle a su me motiver, me redonner courage et me relancer, que dans les meilleurs où elle a su célébrer avec moi. Durant toutes ces années, elle m'a écouté pratiquer des présentations, elle a corrigé mes textes, elle m'a entendu ressasser les mêmes histoires, jour après jour, semaine après semaine, année après année. Elle doit maintenant connaître le ribozyme VS aussi bien que moi et cette thèse est aussi un peu la sienne, juste pour les sacrifices qu'elle a dû faire pour mon succès. Elle est restée tout sourire et inébranlable autant des les mauvaises journées où planaient mal de tête et humeur massacrant, que dans les belles journées où je croyais pouvoir déplacer des montagnes. Et pendant que cette thèse naissait dans mon esprit, un petit bébé poussait dans son ventre. Cette thèse est donc l'accomplissement d'une personne mais provient aussi et surtout de la réussite d'un couple. Jess, je tiens à te remercier du plus profond de mon cœur pour ta patience d'or et pour ton amour. Sans toi, je ne sais pas si j'y serais parvenu.

Ainsi, à tous et à toutes qui m'avez accompagnés de près ou de loin, je tiens à vous dire, encore une fois, merci.

The function is determining the form; the form is expressing the function

*Louis H. Sullivan
Architecte*

À Céline, Yvette et Maxime qui nous avez quitté,

À ceux qui m'accompagnent encore,

Et à toi, Élizabeth, que j'attends impatiemment...

1. Introduction

1.1. Importance biologique de l'ARN

L'ARN est une molécule essentielle qui se retrouve dans tous les domaines du vivant et qui remplit des fonctions capitales pour la survie des cellules. Voilà 50 ans, le rôle connu de l'ARN dans la biologie moléculaire se limitait au transfert de l'information génétique en tant qu'ARN messenger (ARNm). Cependant, bien que les ARNm jouent un rôle central dans les processus biologiques, la découverte des nombreux ARN non-codants (ARNnc) a révélé l'importance fonctionnelle de l'ARN.

Depuis plusieurs années, le nombre d'ARNnc découverts ne cesse d'augmenter et leur découverte a mis en lumière la diversité et l'importance de leurs rôles dans plusieurs processus biochimiques. Puisque ces ARN remplissent des fonctions précises, que ce soit par eux-mêmes ou en association avec des protéines, ils sont considérés comme des ARN fonctionnels. Les ARNnc jouent par exemple des rôles importants dans la machinerie traductionnelle (ARNt et ARNr; [1, 2]), la réplication de l'ADN (ARN Y; [3]), le maintien de l'intégrité chromosomique (TERC; [4]) et la régulation de l'expression génique (ARNmi, ARNsi, ARNpi, CRISPR et riboswitch; [5-9]). Les ARNnc sont aussi impliqués dans la modification (ARNsno; [10]), l'épissage (ARNsn; [11]) et le clivage (ARNsi et ribozymes; [12]) des ARNm. De plus, l'étude des ARN viraux ne cesse de révéler comment ces virus, poussés par une grande pression évolutive, développent de nouvelles stratégies pour se maintenir et se répliquer dans les cellules hôtes (ARNci, ARNsf; [13, 14]).

Une des découvertes les plus significatives sur l'ARN a eu lieu durant les années 80 lorsque les scientifiques Sydney Altman et Thomas Cech ont découvert les premiers ARN ayant des propriétés catalytiques [15, 16]. La découverte de ces enzymes ribonucléiques, aussi appelées ribozymes, a révolutionné la perception des scientifiques sur l'origine de la vie [17, 18]. En effet, comme l'ARN peut à la fois contenir du matériel génétique et catalyser des réactions chimiques, il est possible que les premières proto-cellules aient été essentiellement composées d'ARN. Cette hypothèse de l'évolution de la vie à partir d'un monde à ARN est également soutenue par le vaste rôle des ARN fonctionnels dans toutes les sphères biochimiques, suggérant également qu'il reste certainement beaucoup à découvrir sur les rôles de l'ARN et sur les mécanismes sous-jacents à ces fonctions.

Bien que l'ARN soit omniprésent et qu'il possède des fonctions cellulaires variées, les détails moléculaires de sa polyvalence fonctionnelle sont moins bien compris. Malgré

tout, il est clair que la disparité entre la fonction des différents ARN réside dans le lien étroit qui existe entre la structure et la fonction des ARN. Par exemple, un ribozyme doit adopter une architecture précise qui contribue à l'adoption de sa structure active, facilitant ainsi la formation du site actif où se déroule la réaction enzymatique. En conséquent, comme la structure adoptée par un ARN dicte généralement sa fonction, il est essentiel d'étudier la structure des ARNnc, et plus particulièrement celle des ribozymes, afin de comprendre les bases structurales de leur fonctionnement et de leur diversité.

1.2 La structure de l'ARN

1.2.1. Les bases structurales de l'ARN

L'élément de base de l'ARN est le ribonucléotide. Celui-ci est constitué de trois groupements distincts soit un groupement phosphate, un ribose et une base azotée [19]. Les bases azotées standards (adénine, guanine, cytosine, uracil) sont des hétérocycles qui se divisent en deux classes dépendamment de leur composition chimique [19]. L'adénine (A) et la guanine (G) sont des purines, leur base azotée étant formée par l'association d'un cycle imidazole et d'un cycle pyrimidique. La classe des pyrimidines contient quant à elle la cytosine (C) et l'uracil (U) dont les bases azotées sont formées de cycles pyrimidiques à six atomes. Les bases azotées sont liées aux riboses par un lien glycosidique pour former les ribonucléosides, qui sont respectivement nommés adénosines, guanosines, cytidines et uridines (Figure 1.1; [19]). Les ribonucléotides, ou simplement nucléotides, sont quant à eux définis comme des ribonucléosides phosphorylés en position 5' du ribose [19]. Une molécule d'ARN est formée par une suite de nucléotides reliés entre eux par un lien covalent 3'-5'-phosphodiester [19]. Par convention, le début d'une chaîne d'ARN est défini comme étant l'extrémité 5' et la fin comme l'extrémité 3'.

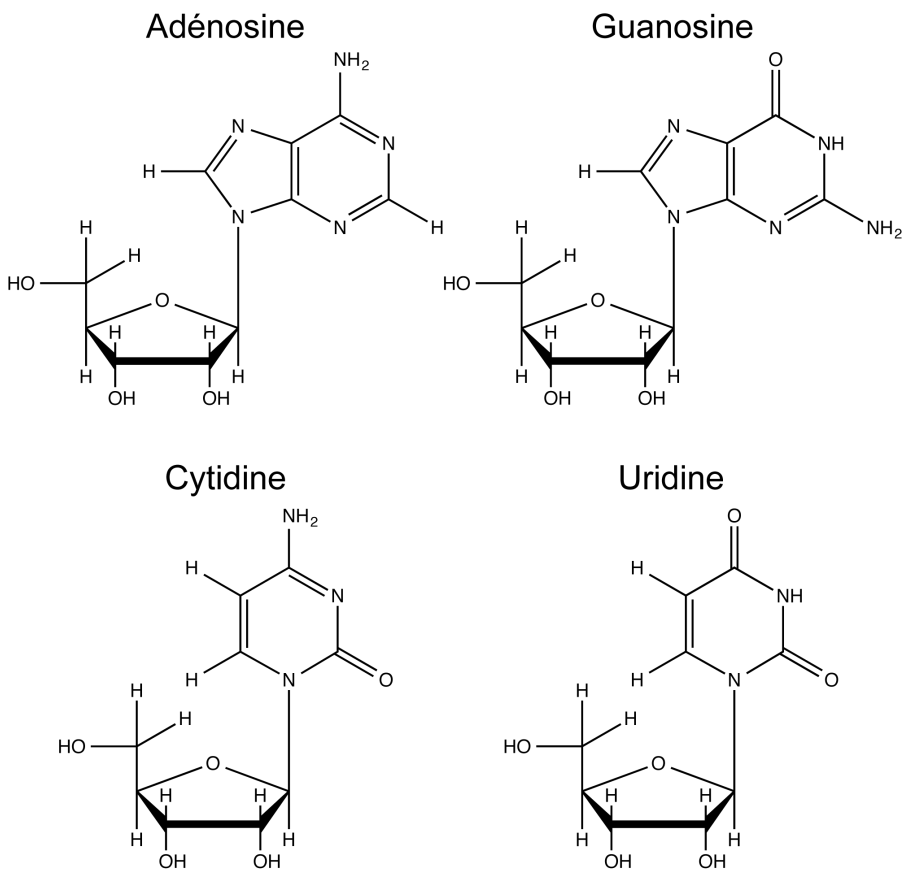


Figure 1.1. Structure chimique des ribonucléosides standards composant l'ARN. L'adénosine et la guanosine ont pour bases des purines tandis que la cytidine et l'uridine ont pour bases des pyrimidines.

L'ARN adopte quatre niveaux organisationnels fondamentaux différents qui définissent sa structure. La structure primaire correspond à la séquence d'un ARN, c'est-à-dire à l'ordre et à la composition en nucléotides de celui-ci. La structure secondaire correspond à l'agencement des paires de bases canoniques et des régions simple brin présentes dans un ARN (Sections 1.2.2 et 1.2.3). Les éléments de structure secondaire comprennent les régions simple brin, les tiges, les boucles terminales, les renflements, les boucles internes et les jonctions multivoies (Figure 1.2; [20]). La structure tertiaire est l'organisation tridimensionnelle (3D) d'un ARN qui correspond généralement à sa structure active dans la cellule (Section 1.2.4). La structure tertiaire est caractérisée par une organisation spatiale précise des nucléotides les uns par rapport aux autres et se distingue de la structure secondaire par les interactions à longue distance qui se forment entre les éléments de structure secondaire. Finalement, certains ARN sont caractérisés par la

formation d'une structure quaternaire qui correspond à l'association de plusieurs moléculaires les unes avec les autres.

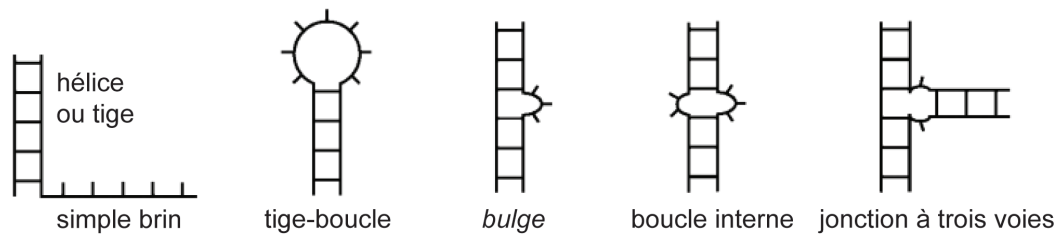


Figure 1.2. Représentation de divers éléments de structure secondaire. Tiré de [20].

1.2.2. Les paires de bases et les hélices d'ARN

Une des interactions les plus communes dans l'ARN est la paire de bases, où deux nucléotides interagissent ensemble via des ponts hydrogène. De par leur composition, les nucléotides possèdent trois faces d'interactions. Les faces Watson-Crick (WC) et *Sugar edge* sont communes à tous les résidus tandis que les purines possèdent une face Hoogsteen et les pyrimidines une face C-H (Figure 1.3A; [21]). Ces faces présentent des groupements chimiques cycliques (N) et exocycliques (NH₂ et C=O) qui, avec le groupement du 2'-OH du ribose, favorisent la formation de ponts hydrogène entre les nucléotides (Figure 1.3B). Selon la nomenclature de Leontis et Westhof, une paire de bases est définie par la nature des nucléotides engagés dans l'interaction, par les faces d'interactions impliquées et par l'orientation relative des liaisons glycosidiques des nucléotides [22]. La géométrie est en *cis* lorsque les liens glycosidiques des deux nucléotides pointent dans la même direction et en *trans* lorsque les liens glycosidiques pointent dans des directions opposées [22].

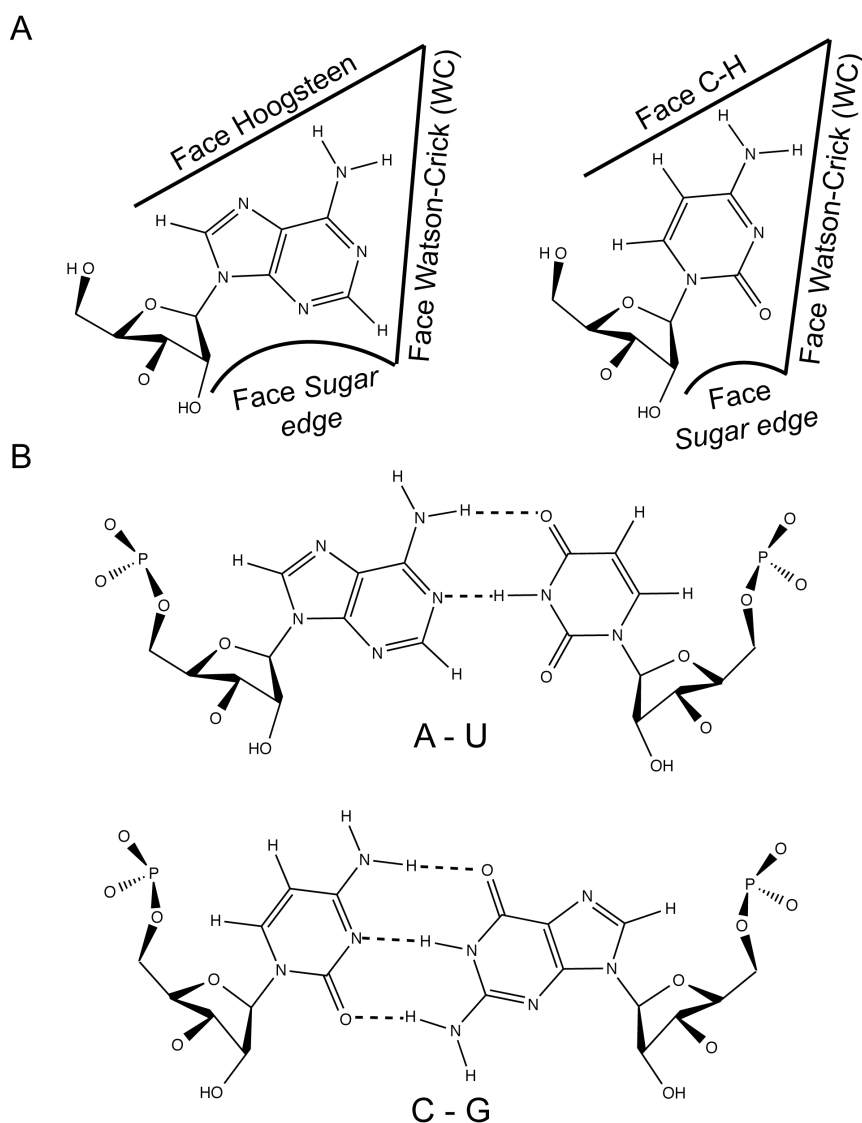


Figure 1.3. Représentation schématique des faces d'interactions de nucléosides et de deux paires de bases canoniques. (A) Les faces d'interactions des purines et des pyrimidines représentées sur une adénosine et une cytidine, respectivement et (B) les paires de bases canoniques *cis* WC/WC A-U et C-G. Adapté de [21].

Certaines paires de bases se forment de façon préférentielle à cause de la nature hautement complémentaire de leur face Watson-Crick (WC). Il s'agit des paires de bases *cis* WC/WC G-C et A-U qui sont définies comme des paires de bases canoniques de l'ARN (Figure 1.3B; [23, 24]). La paire de bases *cis* WC/WC G-U (ou *wobble* G-U) est aussi considérée comme canonique car elle couramment retrouvée dans l'ARN. Les hélices d'ARN sont principalement constituées de ces trois paires de bases. Cependant, plusieurs

autres paires de bases, nommées paires de bases non-canoniques, peuvent également se former dans d'autres contextes en supportant la formation de ponts hydrogène entre les différentes faces des nucléotides. Par exemple, la paire de bases *trans Sugar edge/Hoogsteen* G-A est retrouvée dans certaines tétraboucles [25, 26], tandis que les paires de bases *trans Sugar edge/Hoogsteen* A-A et *cis WC/WC* G-A, A-A et U-U sont fréquemment retrouvées dans les jonctions multivoies et les boucles internes d'ARN [27, 28]. En fait, dans la majorité des régions simple brin des structures secondaires d'ARN, la plupart des nucléotides forment de telles paires de bases non-canoniques [21, 22].

La formation de paires de bases canoniques consécutives génère une hélice d'ARN double brin. L'empilement de ces paires de bases contribue de façon significative à la stabilisation des domaines hélicaux. Les hélices d'ARN typiques sont de types A; celles-ci sont droites (*right-handed*) et ont un tour complet d'hélice accompli en 11 paires de bases [19]. Les hélices de type A sont caractérisées par un sillon majeur profond et étroit et un sillon mineur large et peu profond. La présence du groupement 2'-OH favorise l'adoption d'une configuration C3'-endo des riboses dans ces hélices.

1.2.3. Les éléments de structure secondaire de l'ARN

Bien que les hélices constituent la majeure partie des molécules d'ARN, les hélices sont souvent interrompues ou coiffées par des segments simple brin. Ces segments forment alors des boucles terminales, des renflements, des boucles internes et des jonctions multivoies qui sont des éléments importants de la structure secondaire de l'ARN. Ils favorisent le repliement et la stabilisation de l'ARN et sont considérés comme des motifs lorsqu'ils ont des séquences récurrentes qui forment des structures bien définies [29]. Les prochaines sections présenteront les motifs de structure secondaire retrouvés dans les boucles terminales, boucles internes et jonction multivoies.

1.2.3.1. Les boucles terminales

Les boucles terminales sont des segments simple brin qui coiffent les hélices d'ARN. Certaines boucles terminales assurent des structures bien définies aux extrémités des hélices et les plus communes sont les boucles UNR, GNRA, UNCG et CUYG. Selon

cette nomenclature, N correspond à n'importe quel nucléotide; R est une purine (A ou G) et Y correspond à une pyrimidine (C ou U). Toutes ces boucles terminales ont été largement caractérisées par des études phylogénétiques, thermodynamiques et structurales et sont fréquemment utilisées afin de stabiliser des hélices d'ARN lors d'études structurales [30, 31].

Le motif *U-turn* est formé d'une boucle d'au moins trois résidus ayant la séquence consensus UNR (Figure 1.4A). Cette boucle a initialement été caractérisée dans la boucle anticodon de l'ARNt de la phénylalanine [32]. Le motif *U-turn* assure un virage brusque de la chaîne ribose-phosphate permettant de maintenir un empilement de base entre les résidus N et R et l'hélice. Deux ponts hydrogène impliquant l'uridine stabilisent la boucle UNR (2'-OH du U et N7 du R; H3 du U et le 3'-phosphate du résidu R; Figure 1.4A). Grâce au virage brusque présent dans le *U-turn*, les faces WC des résidus N et R sont dirigées vers le solvant dans le sillon mineur de l'hélice, facilitant la formation d'interactions tertiaires [33, 34].

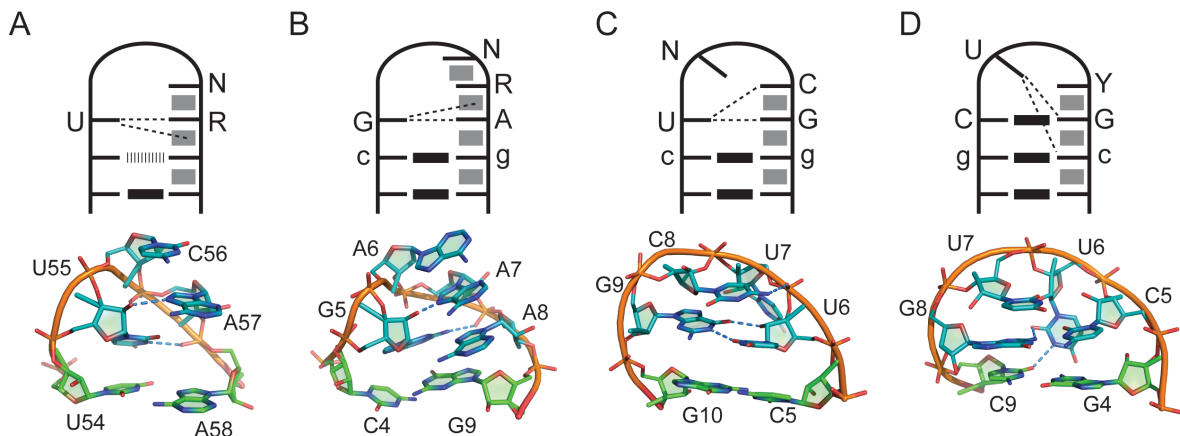


Figure 1.4. Représentations schématiques et structurales de boucles terminales. (A) Boucle UNR (PDB : 1FFY; [35]), (B) cGNRAg (PDB : 1ZIF; [26]), (C) boucle cUNCGg (PDB : 2KOC; [36]) et (D) boucle gCUUGc (PDB : 1RNG; [37]). Dans les représentations schématiques, les rectangles noirs représentent les appariements WC, les rectangles pointillés les appariements non WC, les rectangles gris l'empilement entre les bases et les traits pointillés les ponts hydrogène. Dans les représentations structurales, les nucléotides impliqués dans les boucles sont représentés en bleus, les paires de bases fermantes sont en vert tandis que les lignes pointillées représentent les ponts hydrogène.

La boucle cGNRAg a été identifiée par des analyses comparatives de séquences de l'ARN ribosomal 16S d'*Escherichia coli* (*E. coli*; [30]). Il s'agit d'une tétraboucle qui partage plusieurs similarités structurales avec le motif *U-turn*, incluant un virage brusque entre le G et le N et l'empilement des bases à partir du résidu N et des résidus suivants [25, 26, 38]. La boucle GNRA est aussi stabilisée par la formation de ponts hydrogène (2'-OH du G et le N7 du R; H1 du G et le 3'-phosphate du R; Figure 1.4B; [38]). Une paire de bases canonique C-G conservée ferme la boucle GNRA. De plus, les résidus N, R et A pointent en direction du solvant dans le sillon mineur, favorisant la formation d'interactions tertiaires.

La boucle cUNCGg est une autre tétraboucle qui a été identifiée à partir d'analyses phylogéniques [30] et d'expériences de dénaturation thermique et de transcription inverse sur l'ARNm du bactériophage T4 [39]. Contrairement aux boucles UNR et GNRA, la chaîne ribose-phosphate de la boucle UNCG adopte un virage beaucoup moins abrupt (Figure 1.4C; [40]). Dans cette boucle, seul le résidu N se retrouve dans le sillon mineur tandis que les bases des résidus C et G maintiennent l'empilement avec l'hélice. Une des particularités de la boucle est que le G adopte une configuration *syn* plutôt que *anti* [40, 41]. Des ponts hydrogène entre les bases du U et du G ainsi qu'entre le NH₂ du C et le 5'-phosphate du N stabilisent la structure de la boucle. De plus, tout comme dans la boucle GNRA, une paire de bases canonique C-G conservée stabilise et ferme la boucle cUNCGg [39].

Finalement, la boucle gCUYGc est la moins fréquente des tétraboucles [30]. Similairement à la boucle UNCG, la chaîne ribose-phosphate définit un virage qui positionne le résidu U dans le sillon mineur où il est stabilisé par des ponts hydrogène entre le H3 du U et le O2 du C de la paire de bases fermante et le O2 du U et le NH₂ du G (Figure 1.4D; [37]). Les résidus C et G de la boucle gCUYGc forment une paire de bases *cis* WC/WC, tout comme les résidus G et C de la paire de bases fermante.

1.2.3.2. Renflements et boucles internes

Les renflements et les boucles internes sont définis comme des régions de la structure secondaire d'un ARN présentant des résidus qui ne sont pas impliqués dans des paires de bases canoniques. Les résidus non-pairés peuvent se retrouver sur un seul brin ou sur les deux brins d'ARN; ils sont alors respectivement définis comme un renflement ou

une boucle interne (Figure 1.2). Bien que les renflements et les boucles internes aient des séquences variées, ils adoptent généralement une structure précise qui favorise la formation d'interactions non-canoniques [42-44]. D'ordre général, les renflements et les boucles internes engendrent des irrégularités dans les hélices d'ARN qui peuvent favoriser la formation d'interactions tertiaires ARN-ARN et ARN-protéines (Figure 1.5; [42, 45, 46]). Les renflements et boucles internes sont parfois stabilisés par certains motifs structuraux tel que les motifs *S-turn* et les empilements de purines inter-brin (*cross-strand stacking interaction*).

Tout d'abord, le motif *S-turn* a initialement été identifié dans la boucle E de l'ARNr 5S d'eucaryotes [47]. Bien que sa séquence soit généralement variable, le *S-turn* adopte toujours une structure tridimensionnelle similaire qui est caractérisée par une large distorsion de la chaîne ribose-phosphate adoptant une forme de S (Figure 1.5). Cette distorsion entraîne généralement l'exclusion d'au moins un nucléotide à l'extérieur de l'hélice (Figure 1.5). Le motif *S-turn* est aussi caractérisé par la formation de paires de bases non-canoniques, par l'inversion de la configuration d'un ribose de C3'-endo à C2'-endo, et engendre le renversement (*flipping*) d'au moins un ribose [48]. Le *S-turn* est retrouvé dans le ribozyme *hairpin* [49, 50], la boucle Sarcin-Ricin [51] et de nombreux autres ARN [48, 52-55].

Par ailleurs, les boucles internes peuvent également inclure un motif d'empilement de purines inter-brin [29]. Normalement, les nucléotides d'une hélice tendent à s'empiler sur les nucléotides du même brin d'ARN. Cependant, dans certains contextes de boucles internes ou de jonctions, les purines provenant de brins opposés peuvent s'empiler, stabilisant ainsi les structures d'ARN en réduisant le contact entre le solvant et les surfaces hydrophobes des nucléotides [29].

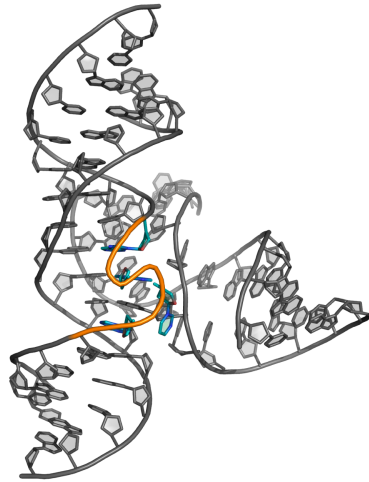


Figure 1.5. Représentation structurale d'une boucle interne. Motif *S-turn* dans la boucle interne du ribozyme *hairpin* (PDB : 2D2L; [50]). Le *S-turn* est représenté en couleur.

1.2.3.3. Les jonctions multivoies

Une jonction d'ARN est formée par la rencontre de plusieurs hélices distinctes à un même point, que l'on nomme point de branchement [56], et ces hélices sont souvent connectées entre elles par des segments simple brin [57]. Les jonctions d'ARN jouent un rôle crucial dans le repliement des ARNt, des ARNr [58], des riborégulateurs [59-62] et des ribozymes [58, 63-70]. Les jonctions agissent aussi comme des plateformes d'interactions avec les protéines dans différents contextes comme dans le ribosome [71, 72] et dans la particule de reconnaissance du signal (*signal recognition particle*; [73]).

La classe de jonctions la plus abondante est formée des jonctions à trois voies où trois hélices se connectent au même point de branchement [74]. Les jonctions à trois voies sont organisées en famille selon leur structure tertiaire, selon la longueur des segments simple brin de la jonction et selon l'orientation de leurs hélices (hélices P1, P2 et P3; Figure 1.6). Les familles A et B sont les moins abondantes et sont exclusivement retrouvées chez les ARNr [75]. Les familles A, B et C de jonctions à trois voies sont toutes caractérisées par un empilement coaxial des hélices P1 et P2. Plus spécifiquement, les jonctions à trois voies de la famille A sont caractérisées par un arrangement perpendiculaire de l'hélice P3 par rapport aux deux autres tandis que la famille B adopte plutôt une structure où l'hélice P3 tend vers l'hélice P2 mais sans devenir parallèle (Figure 1.6; [75]). La famille C est la plus vaste et la plus diversifiée des trois classes. Les jonctions de cette famille sont caractérisées par un arrangement où les hélices P1 et P3 sont côtes-à-côtes et sont généralement

stabilisées par des interactions tertiaires (Figure 1.6; [75, 76]). Le segment simple brin entre ces deux hélices forme souvent des motifs de boucles terminales qui coiffent l'hélice P3 [76]. Il s'agit de la seule famille qui contienne des jonctions appartenant à des ARN non-ribosomiaux [75].

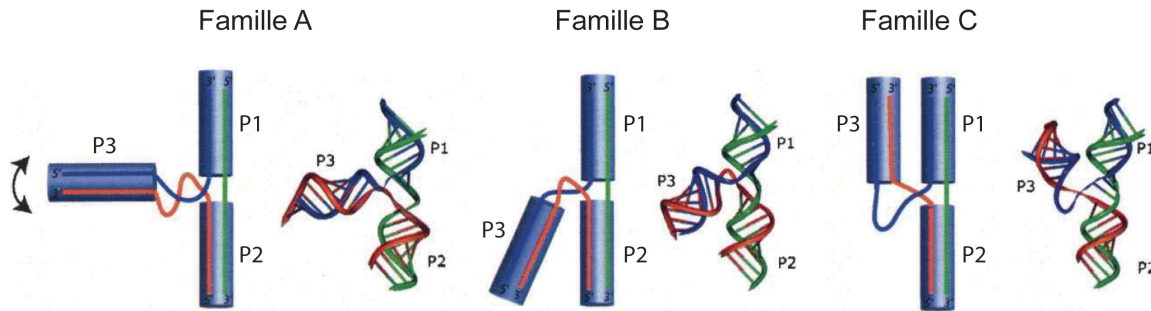


Figure 1.6. Représentations schématiques et structurales des familles de jonctions à trois voies. Tiré de [75].

1.2.4. Les éléments de structure tertiaire de l'ARN

La structure tertiaire de l'ARN correspond à sa forme fonctionnelle, c'est-à-dire la forme lui permettant d'interagir avec d'autres molécules ou encore de catalyser des réactions enzymatiques. Le repliement tridimensionnel de l'ARN passe par la formation d'éléments de structure tertiaire. Ces éléments sont généralement dépendants de l'interaction entre des éléments de structure secondaire et ils se divisent en trois classes : les interactions entre des domaines hélicaux, les interactions entre des domaines hélicaux et des régions non-pairées ainsi que les interactions entre des régions non-pairées [20].

1.2.4.1. Interactions entre des domaines hélicaux

Une des interactions principales entre les nucléotides est l'empilement consécutif de bases. La stabilisation procurée par cet empilement provient de l'interaction entre les systèmes d'électrons π des bases azotées [19]. Dans la structure tertiaire des molécules d'ARN, les hélices préformées tendent à s'associer entre elles afin de maintenir cet empilement de base. Il s'agit du phénomène d'empilement coaxial des hélices d'ARN [20]. L'empilement coaxial dépend généralement du repliement des boucles internes et des

jonctions multivoies et est favorisé par la présence de cations [20]. Cet empilement entre les domaines hélicaux est particulièrement important chez les ARN fonctionnels qui doivent adopter une structure tertiaire précise et compacte. De nombreux ARN sont donc stabilisés par l'empilement coaxial d'hélices d'ARN [77-82].

1.2.4.2. Interactions entre des domaines hélicaux et des régions non-pairées

Un autre élément de structure tertiaire est l'interaction entre certaines boucles terminales et des régions hélicales [72, 81]. Par exemple, les trois adénosines de la boucle GAAA (Section 1.2.3.2) peuvent interagir de façon spécifique avec un récepteur de boucle GAAA, formé par le sillon mineur d'un motif hélicale particulier [81, 83-85]. De plus, les résidus N et R des motifs *U-turn* (Section 1.2.3.1) peuvent aussi interagir avec le sillon mineur de domaines hélicaux [64, 71, 72, 76, 86-91]. Les interactions impliquant les sillons mineurs sont récurrentes et semblent expliquer la différence fonctionnelle entre l'ARN et l'ADN puisque les hélices régulières d'ARN possèdent un sillon mineur large et peu profond tandis que celles d'ADN possèdent un sillon mineur profond et étroit. De plus, les groupements 2'-OH de l'ARN se situent dans les sillons mineurs, ce qui favorise la formation de ponts hydrogène [20].

Un autre type d'interaction entre des domaines hélicaux et des régions non-pairées est le triplet de bases. Les triplets de bases sont formés par l'interaction d'un nucléotide dans le sillon majeur ou mineur d'une paire de bases. D'ordre général, les interactions dans les sillons majeurs sont plutôt rares mais existent tout de même comme c'est le cas dans le long ARN non-codant MALAT1 (Figure 1.7A; [77, 92-94]). Les triplets de bases dans les sillons mineurs ont jusqu'à maintenant été beaucoup plus fréquemment observés [77, 81, 95]. Plus spécifiquement, les adénosines non-pairées forment un motif d'interaction particulier dans les sillons mineurs : il s'agit de l'interaction A-mineur (Figure 1.7B; [95]). Ce motif d'interaction a été grandement caractérisé chez les ARNr où quatre géométries d'interactions A-mineur sont répertoriées (Figure 1.7B; [95]). L'absence de groupement chimique sur la face *Sugar edge* des adénosines favorise l'insertion de ce résidu dans les sillons mineurs par la formation de ponts hydrogène et de contacts de van der Waals [2, 95].

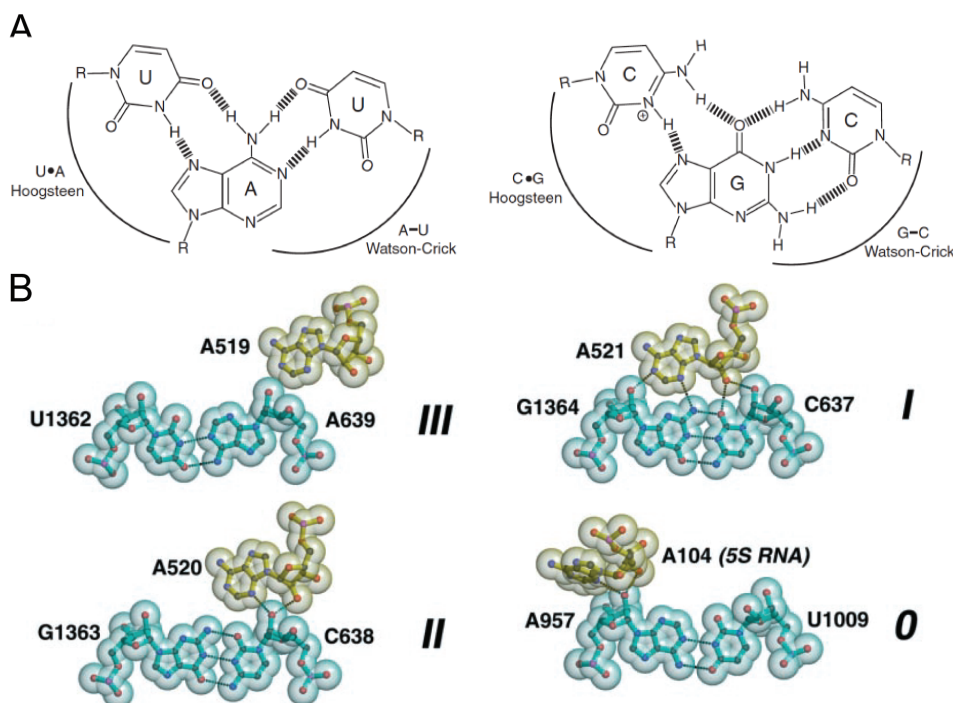


Figure 1.7. Représentations schématiques et structurales de triplets de bases. (A) Triplets de bases formés par une interaction dans le sillon majeur et (B) triplets de bases formés par une interaction A-mineur. La figure en (A) est tirée de [93] et celle en (B) est tirée de [95].

La formation de triplets de bases, particulièrement les motifs A-mineur, favorisent aussi la formation d'un autre type d'interaction tertiaire, les *ribose zippers* [96]. Un *ribose zipper* est formé par des ponts hydrogène entre les riboses de deux chaînes antiparallèles, impliquant principalement les groupements 2'-OH [96]. Le réseau d'interactions multiples retrouvé dans les *ribose zipper* a le potentiel de condenser les hélices d'ARN. Ainsi, les *ribose zipper* sont abondamment retrouvés dans les ARNr [96], dans les introns du groupe I [81] et dans les ribozymes *hammerhead* [97] et *Hepatitis Delta Virus* [82].

1.2.4.3. Interactions entre des régions non-pairées

Les nucléotides non-pairés des boucles terminales peuvent participer à des interactions avec différents éléments structuraux. Par exemple, un pseudonoeud est formé par l'interaction entre une boucle terminale et un segment simple brin complémentaire tandis qu'un *kissing-loop* résulte de la formation de paires de bases entre deux boucles terminales (Figure 1.8A-B). Les pseudonoeuds ont été identifiés dans plusieurs ARN viraux [82, 98-101], tandis que des *kissing-loop* jouent des rôles cruciaux dans les ARNt [32], les

riborégulateurs [59, 60, 102-105], les ribozymes [65, 91, 106-108] et dans l'initiation de la dimérisation du génome viral du VIH-1 [109]. Ces interactions permettent le rapprochement de domaines éloignés dans la structure secondaire et leur stabilisation se fait par la formation de paires de bases, de triplets de bases et d'empilements coaxiaux.

Finalement, des interactions tertiaires peuvent aussi se former entre des boucles internes. Bien qu'il ne s'agisse pas d'un motif tertiaire répandu, ce genre d'interaction joue un rôle crucial dans plusieurs ARN comme dans le ribozyme *hairpin* dont le site actif est formé par l'interaction de deux boucles internes qui contiennent chacune un nucléotide catalytique [110, 111]. Dans ce cas précis, l'affinité de l'interaction entre les deux boucles internes est suffisamment forte pour qu'elle se produise en absence d'éléments structuraux supplémentaires [112].

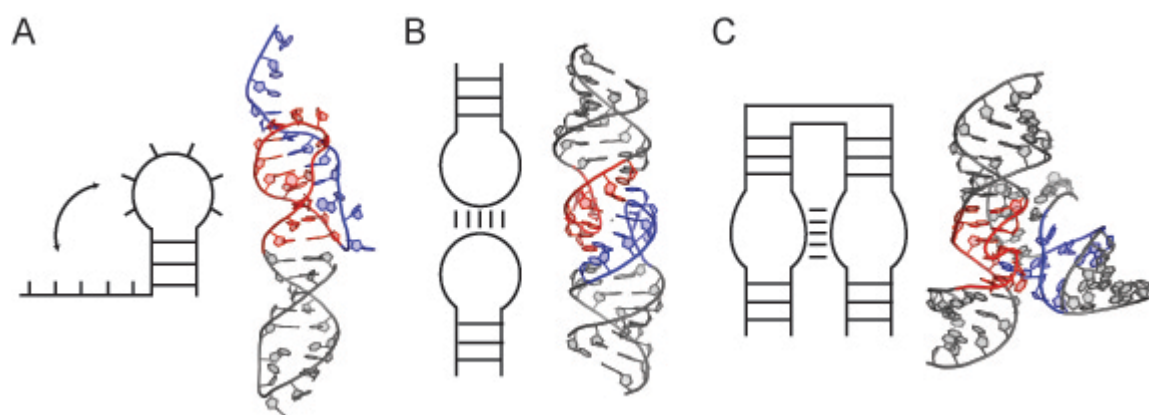


Figure 1.8. Représentations schématisques et structurales d'interactions tertiaires entre des régions non-pairées. (A) Pseudonoeud (PDB : 1A60; [99]), (B) interaction *kissing-loop* (PDB : 2B8S; [113]) et (C) interaction entre des boucles internes (PDB : 2D2L; [50]). Les régions d'interactions sont colorées en rouge et bleu.

1.2.5. Contribution des métaux à la formation de structures tertiaires d'ARN

Chaque groupement phosphate composant un nucléotide comporte une charge négative à pH neutre, conférant aux molécules d'ARN un caractère anionique. À cause de cette charge, l'ARN en solution est entouré d'une couche de molécule d'eau, appelée couche d'hydratation. Le repliement en structure complexe, nécessaire pour la majorité des ARN fonctionnels, requiert généralement la présence de cations qui favorisent le rapprochement des chaînes ribose-phosphates en équilibrant les charges négatives. Les

cations monovalents (Na^+ , K^+) et divalents (Mg^{2+} , Ca^{2+}) sont biologiquement prédominants tandis que certains cations trivalents, comme le cobalt hexamine ($\text{Co}(\text{NH}_3)_6^{3+}$) possèdent des avantages expérimentaux [114, 115]. Malgré tout, le Mg^{2+} est le cation le mieux adapté pour le repliement de l'ARN puisqu'il a un petit rayon ionique, une densité de charge élevée et une géométrie en octaèdre qui permet de coordonner six ligands.

En solution, le Mg^{2+} est entouré d'une couche de six molécules d'eau. Sous cette forme hexahydratée, le Mg^{2+} peut interagir de trois façons avec l'ARN [116]. Tout d'abord, le Mg^{2+} et l'ARN peuvent interagir de façon diffuse à travers leur couche d'hydratation respective (Figure 1.9A). Une interaction diffuse signifie que les ions sont délocalisés autour de l'ARN et qu'ils ne favorisent le repliement de l'ARN que par des interactions Coulombiennes non-spécifiques.

Les deux autres modes d'interactions sont définis comme étant spécifiques à des sites de liaison précis. La concentration des forces d'attraction de l'ARN peut favoriser la localisation des ions à un site précis où l'ion Mg^{2+} et l'ARN partage la même couche d'hydratation. Il s'agit donc d'une interaction indirecte, nommée interaction *outer-sphere* (Figure 1.9B). Le dernier type d'interaction implique une interaction encore plus forte entre l'ARN et le Mg^{2+} . Dans ce cas, l'ion Mg^{2+} perd une de ses molécules d'eau et interagit directement avec l'ARN. Cette interaction directe est nommée interaction *inner-sphere* (Figure 1.9C). Dans ces deux derniers cas, les ions Mg^{2+} favorisent localement la condensation de l'ARN en permettant le rapprochement de groupements chimiques chargés.

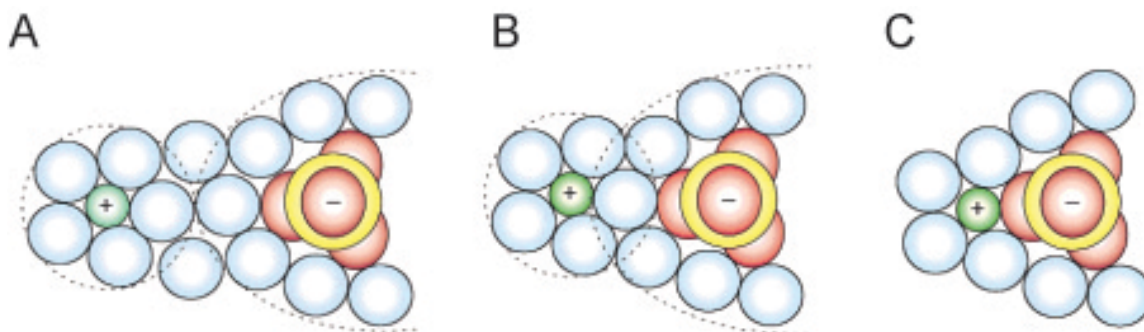


Figure 1.9. Représentations des différents types d'interactions entre l'ARN et le Mg^{2+} . (A) La liaison diffuse s'effectue via les molécules d'eau des couches d'hydratation de l'ARN et du magnésium. (B) La liaison indirecte implique la couche d'hydratation du magnésium qui interagit avec un groupement de l'ARN. (C) La liaison directe s'effectue entre le magnésium et l'ARN sans impliquer de molécule d'eau. Tiré de [116].

1.3. Les ribozymes naturels

1.3.1. L'importance de l'étude des ribozymes

Tel que mentionné précédemment, les ribozymes sont des enzymes ribonucléiques pouvant être associés avec des protéines mais dont le catalyseur direct est toujours l'ARN. Il a été proposé que les ribozymes aient joués un rôle déterminant dans l'apparition de la vie et l'étude des ribozymes naturels permet de mieux comprendre les mécanismes ayant pu contribuer au monde à ARN [117]. De plus, les ribozymes constituent des éléments actifs dans la biologie actuelle qui catalysent certaines réactions chimiques essentielles au métabolisme cellulaire [117]. Plusieurs ribozymes naturels ont été identifiés jusqu'à présent et ceux-ci catalysent différents types de réactions enzymatiques. La diversité de ces réactions concorde avec la diversité de structures adoptées par les ribozymes [118]. Ils sont donc caractérisés par une étroite relation structure-fonction qui fait de ceux-ci des systèmes d'études modèles. Les ribozymes sont aussi des catalyseurs relativement simples comparativement aux enzymes protéiques. Ainsi, l'étude des ribozymes facilite la compréhension moléculaire des mécanismes catalytiques, ce qui pourrait en retour permettre de comprendre le fonctionnement de certaines enzymes protéiques [117].

Par ailleurs, différentes approches sont disponibles pour faire de l'ingénierie et modifier les ribozymes existants [117]. L'utilisation d'approche de sélection *in vitro* (SELEX) ou d'ingénierie rationnelle pourrait faciliter la conception et l'utilisation des ribozymes en tant qu'outils biochimiques ou encore comme traitement thérapeutique. Par exemple, des ribozymes *hammerhead* modifiés (Section 1.3.3) ont été conçus afin de cibler l'ARNm du gène VEGFR-1, démontrant un effet anti-tumoral et anti-métastatique [119], tandis qu'un autre a été conçu pour cibler l'ARNm du récepteur à chimiokine CCR5 afin de combattre les infections par le virus VIH-1 [120]. Ainsi, l'étude des ribozymes est hautement importante pour comprendre plusieurs processus biochimiques et leurs bases moléculaires et pourrait éventuellement mener à de nouveaux développements thérapeutiques.

1.3.2. Les différentes classes de ribozymes

Présentement, plusieurs ribozymes naturels ont été identifiés et caractérisés, et ceux-ci jouent pour la plupart des rôles cellulaires essentiels. Les ribozymes sont catégorisés selon le type de réaction chimique catalysée et les ribozymes connus sont le ribosome [121], la ribonucléase P (RNase P; [122, 123]), les introns du groupe I [124, 125] et du groupe II [126], le ribozyme GIR1 [127], le spliceosome [128], les ribozymes *CPEB3* [129], *CoTC* [130] et *CLEC2* [131] ainsi que la classe des petits ribozymes nucléolytiques (Table 1 et Figure 1.10; [70, 111, 118]). Chacun de ces ribozymes sera brièvement présenté ci-dessous.

Table 1.1. Liste des ribozymes naturels connus, de leur mécanisme enzymatique et de la disponibilité d'une structure tridimensionnelle. Adapté de [117].

Ribozymes	Mécanisme enzymatique	Structure	Références
<i>Synthèse peptidique</i>			
Ribosome	Transfert peptidique	Oui	[121]
<i>Maturation des ARNt</i>			
RNase P	Hydrolyse	Oui	[122, 123]
<i>Épissage</i>			
Introns du groupe I	Double transestérification	Oui	[124, 125]
Introns du groupe II	Double transestérification	Oui	[126, 132]
GIR1	Double transestérification	Oui	[127]
Spliceosome	Double transestérification	Non	[128]
<i>Ribozymes nucléolytiques</i>			
<i>Hammerhead</i>	Transestérification simple	Oui	[90]
<i>Hairpin</i>	Transestérification simple	Oui	[133]
HDV	Transestérification simple	Oui	[82]
<i>gms</i>	Transestérification simple	Oui	[67]
<i>Twister</i>	Transestérification simple	Oui	[70]
<i>Varkud Satellite (VS)</i>	Transestérification simple	Non	[134, 135]
<i>Présents dans ARNm</i>			
<i>CoTC</i>	Transestérification simple	Non	[130]
<i>CLEC2</i>	Transestérification simple	Non	[131]
<i>CPEB3</i>	Transestérification simple	Non	[129]

Le ribosome et la RNase P sont des molécules universelles qui sont essentielles à la survie de toutes les cellules. Ces deux ribozymes forment des complexes ribonucléoprotéiques et jouent des rôles clés qui permettent le fonctionnement de la machinerie traductionnelle cellulaire. Tout d'abord, le ribosome permet la traduction des ARNm en protéines. Le site actif du ribosome est formé d'une poche symétrique d'environ 180 nucléotides qui est responsable de la transpeptidation des acides aminés sur la chaîne protéique naissante et qui facilite la translocation de l'ARNt du site A au site P [121, 136-141]. Quant à elle, la RNase P est une endoribonucléase qui permet la maturation des ARNt par hydrolyse de leur l'extrémité 5' (Figure 1.10I; [122, 142]). Comme pour le ribosome, le site actif de la RNase P est constitué d'ARN [143]. Le ribosome et la RNase P supportent donc des fonctions cellulaires essentielles, sont ubiquitaires et ont un arbre phylogénique très ancien, suggérant que ces molécules pourraient provenir d'un monde prébiotique où ils auraient joué un rôle clé dans l'apparition de la vie [121, 144].

Une autre classe de ribozyme est celle des ribozymes catalysant une réaction d'épissage. Cette classe contient les introns du groupe I et II, le ribozyme GIR1 ainsi que le spliceosome (Figure 1.10E et J; [145]). Les introns du groupe I et II et le ribozyme GIR1 catalysent leur auto-épissage à partir de gène d'ARN (ARNm, ARNt et ARNr; [145]), tandis que le spliceosome est responsable de l'épissage des ARNm des cellules eucaryotes [146]. Ces ribozymes catalysent tous une double réaction de transestérification qui permet l'excision des introns sous forme de lariat (lasso) tout en permettant la ligation des régions exoniques flanquantes (Figure 1.11A; [118, 127, 147, 148] [145]). Leurs mécanismes enzymatiques requièrent la présence d'ions métalliques qui jouent un rôle direct dans la réaction puisqu'ils stabilisent le développement des charges et des intermédiaires réactionnels : ces ribozymes sont donc des métalloenzymes [126, 128, 148-150]. Les introns se retrouvent dans à peu près tous les domaines du vivant et il a été proposé que les introns du groupe II pourraient être les ancêtres du spliceosome [151]. De plus, les introns peuvent contenir des séquences codantes et leur capacité d'insertion et d'excision dans les molécules d'ARN et d'ADN leur confère un potentiel thérapeutique intéressant [152-154].

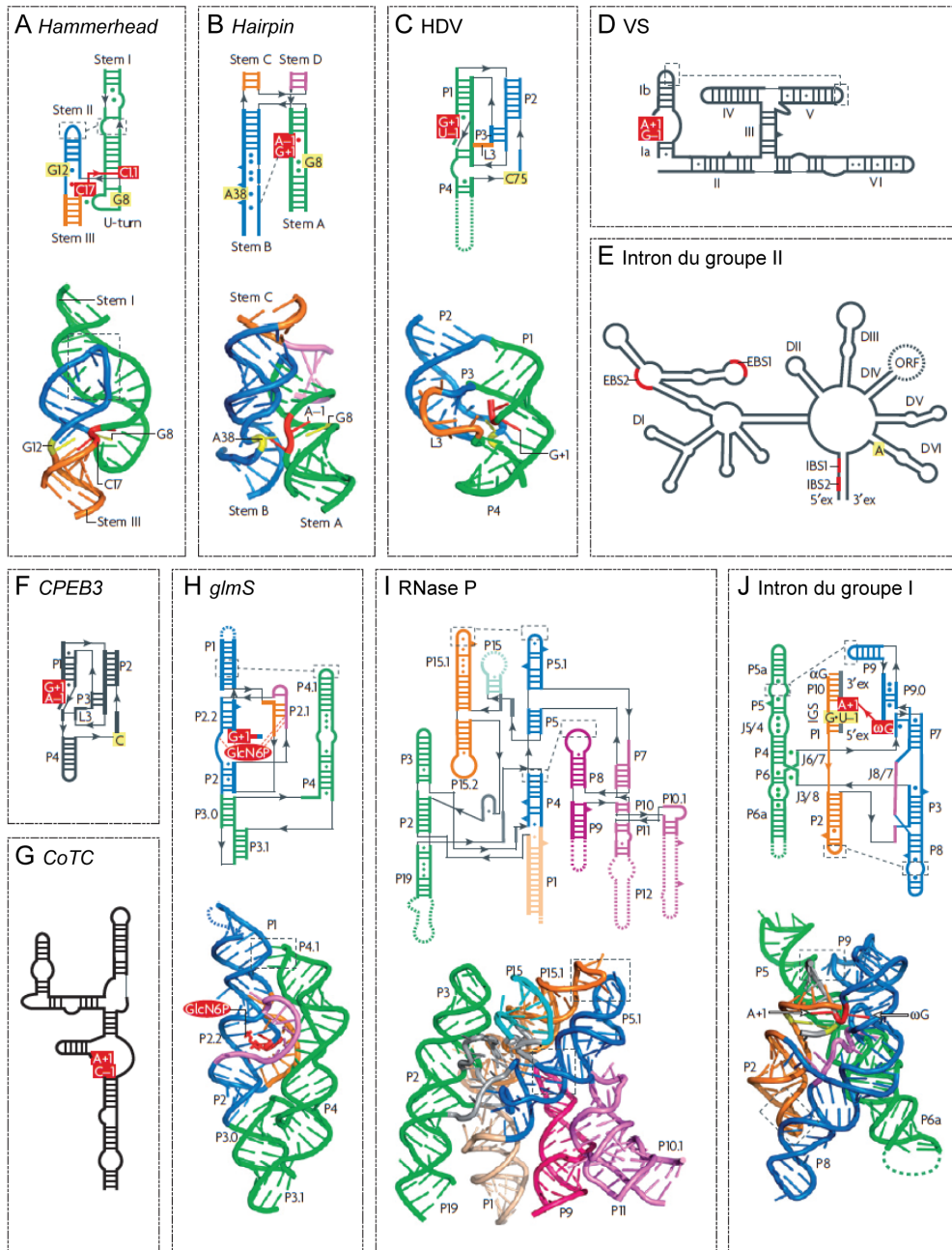


Figure 1.10. Structures secondaires et tridimensionnelles de différents ribozymes naturels.

Dans les structures secondaires, les traits pleins et les cercles représentent les paires de bases, les triangles les nucléotides non-pairés et les lignes pointillées les interactions tertiaires. Les structures tridimensionnelles connues en 2007 sont présentées avec les mêmes couleurs que les structures secondaires. (A) *hammerhead*, (B) *hairpin*, (C) virus de l'hépatite D (HDV), (D) VS, (E) intron du groupe II, (F) *CPEB3*, (G) *CoTC*, (H) *glmS*, (I) RNase P et (J) intron du groupe I. Récemment, les ribozymes *CLEC2* et *twister* ont été découverts et les structures de l'intron du groupe II et du *twister* ont été obtenues [70, 131, 132, 155-157]. Adapté de [158].

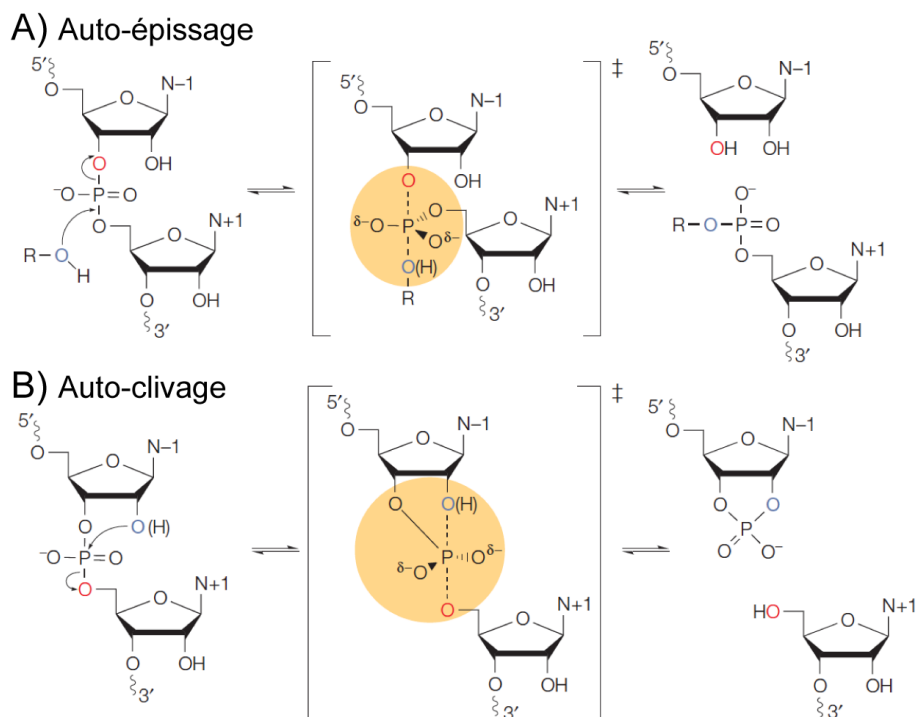


Figure 1.11. Mécanismes réactionnels d'auto-clivage et d'auto-épissage. (A) Mécanisme réactionnel des introns du groupe I et II et (B) mécanisme réactionnel des ribozymes nucléolytiques. Les oxygènes nucléophiles sont en bleus, les groupes partants en rouge et l'intermédiaire réactionnel trigonal bipyramidal est encerclé en jaune. Adapté de [159].

Il est aussi intéressant de noter que quelques ribozymes ont récemment été découverts dans certains ARNm où ils régularaient l'expression génique et l'épissage de certains gènes. Par exemple, le ribozyme *CoTC* (*Co-Transcriptional Cleavage*) est retrouvé dans le gène de la β -globine et participe à la terminaison de la transcription de ce gène (Figure 1.10G; [130]). Un ribozyme au rôle encore méconnu se retrouve également dans le deuxième intron du gène *CPEB3* qui code pour le *cytoplasmic polyadenylation element binding protein 3* (Figure 1.10F; [129]), tandis qu'un autre ribozyme est présent dans le gène *CLEC2* (*C-type lectin type II*) où il régule l'expression de ce gène [131]. Encore relativement peu d'information est disponible sur ces nouveaux ribozymes mais leur découverte suggère que plusieurs autres ribozymes toujours inconnus pourraient se retrouver dans les ARNm et participeraient au contrôle de l'expression des gènes [158].

1.3.3. Les petits ribozymes nucléolytiques

1.3.3.1. Généralités sur les petits ribozymes nucléolytiques

La classe des petits ribozymes nucléolytiques est la plus grande classe de ribozymes connus. Elle contient les ribozymes *hammerhead*, *hairpin*, *glmS*, *twister*, du virus de l'hépatite D (*Hepatitis Delta Virus*; HDV) et *Varkud Satellite* (VS). Tous ces ribozymes sont formés de moins de 200 nucléotides et catalysent la même réaction enzymatique, soit une réaction de transestérification simple menant à l'auto-clivage ou à l'auto-ligation d'un lien phosphodiester spécifique [111, 158]. La classe des petits ribozymes nucléolytiques est particulièrement intéressante pour l'étude de l'ARN. En effet, la simplicité de ces molécules et l'absence d'interaction avec des protéines facilitent grandement l'analyse des relations structure-fonction de l'ARN.

Les petits ribozymes nucléolytiques sont retrouvés dans différents domaines du vivant où ils jouent des rôles distincts. Par exemple, les ribozymes *hammerhead* et *twister* sont retrouvés autant chez les bactéries que dans les cellules eucaryotes [70, 160]. De plus, les ribozymes *hairpin*, HDV et VS et certains ribozymes *hammerhead* sont retrouvés dans des ARN satellites où leur activité catalytique s'inscrit dans le processus de réplication de leur génome, tandis que les ribozymes *glmS* et *twister* sont quant à eux retrouvés dans certains ARNm bactériens [111, 159]. Le *glmS* permet de réguler l'expression génique du gène *glmS* tandis que le ribozyme *twister*, récemment identifié, n'a pas encore de fonction connue [70, 111, 159].

Comme tous les petits ribozymes nucléolytiques catalysent la même réaction de transestérification, il a été proposé qu'ils partagent également le même mécanisme acide-base général [111, 159, 161]. Ce mécanisme, initialement caractérisé chez la ribonucléase A [162], se base sur des transferts de protons et mène au clivage ou à la ligation d'un lien phosphodiester spécifique. Ce mécanisme enzymatique acide-base général fait appel à un mécanisme réactionnel de type substitution nucléophile bimoléculaire (S_N2) où la réaction se déroule en une seule étape concertée. En considérant une réaction de clivage, la réaction débute par l'attaque nucléophile d'une base générale sur le groupement 2'-OH du résidu en position N_{-1} du site de clivage (Figure 1.11B). L'oxygène en 2', maintenant déprotoné, fait à son tour une attaque nucléophile sur le phosphate scissile. La réaction passe alors par un intermédiaire réactionnel où le groupement phosphate est sous forme trigonal bipyramidale.

L'oxyanion ensuite formé sur l'oxygène en 5' prend finalement le proton de l'acide général pour terminer la réaction enzymatique. La réaction chimique produit des extrémités 2'-3' cyclique phosphate pour le résidu N_{-1} et 5'-OH pour le résidu N_{+1} , tandis que l'acide général et la base générale ont inter-converti leur état de protonation. Lorsque les ribozymes nucléolytiques catalysent la ligation, le mécanisme enzymatique se fait dans l'ordre inverse [161]. Différentes études biochimiques et structurales ont aussi démontré que chez la plupart des petits ribozymes nucléolytiques, les rôles d'acide général et de base générale sont joués par des nucléotides.

Pour que le mécanisme enzymatique puisse prendre place, les petits ribozymes nucléolytiques, comme tout enzyme, requièrent la formation d'un site actif. Ce site actif est un microenvironnement où les conditions chimiques et physiques peuvent différer de celles du solvant. Ces différences physicochimiques peuvent inclure des changements locaux de pH ou de force ionique [117]. En effet, lorsque les rôles d'acides et de bases sont joués par des nucléotides, ceux-ci adoptent souvent des pK_a modifiés [163]. La formation du site actif peut donc altérer l'état de protonation de certains nucléotides, favorisant le mécanisme enzymatique. Les sites actifs peuvent aussi former des sites de liaison de métaux [164]. Ces métaux peuvent stabiliser les intermédiaires réactionnels ou favoriser certains changements conformationnels requis pour l'activité catalytique [117, 165]. Les études structurales de certains petits ribozymes nucléolytiques ont permis d'identifier un de ces réarrangements conformationnels. En effet, les nucléotides qui se trouvent de part et d'autre du site de clivage (résidus N_{-1} et N_{+1}) doivent adopter une conformation étendue, nommée conformation *splayed*. Cette conformation étendue au site de clivage favorise une configuration alignée (*in-line*) de l'oxygène en 2', du phosphore et de l'oxygène en 5' [117, 166, 167] et a entre autre été observée chez les ribozymes *glms*, HDV, *hammerhead* et *hairpin* (Figure 1.12). Quant aux ribozymes nucléolytiques pour lesquels la structure tridimensionnelle n'est pas encore connue (comme pour le ribozyme VS), ils doivent probablement adopter une conformation étendue dans leur site actif. Il est donc essentiel d'étudier tous les ribozymes de cette classe afin de pouvoir identifier les généralités fonctionnelles partagées qui favorisent la catalyse.

1.3.3.2. Caractéristiques des petits ribozymes nucléolytiques

Chacun des petits ribozymes nucléolytiques possède des propriétés qui le distinguent. Ces propriétés sont la base de la diversité catalytique de l'ARN. L'étude détaillée des ribozymes nucléolytiques est donc importante, autant pour comprendre les stratégies catalytiques propres à chaque ribozyme que dans une optique d'ingénierie rationnelle de ceux-ci.

Tout d'abord, le ribozyme *glmS* se retrouve dans la région 5' non-traduite de l'ARNm du gène bactérien *glmS* qui code pour le gène L-glutamine:D-fructose-6-phosphate amidotransférase qui produit la glucosamine-6-phosphate (GlcN6P). Cette enzyme fait partie de la voie métabolique de l'UDP-GlcNAc, un composant de la paroi bactérienne [168-170]. L'ARN *glmS* a la particularité d'être à la fois un riborégulateur et un ribozyme [169, 170]. En effet, lorsque la concentration de GlcN6P augmente dans la bactérie, l'aptamère du *glmS* lie le GlcN6P [169-172]. Cette liaison positionne le GlcN6P dans le site actif du ribozyme *glmS* où il participe à la réaction enzymatique comme acide général. L'auto-clivage de l'ARNm du *glmS* permet donc de réguler l'activité métabolique menant à la formation du GlcN6P [169]. Le ribozyme *glmS* est formé de trois domaines hélicaux qui sont stabilisés par deux pseudonoeuds (Figure 1.10H; [67, 172-174]). Ce ribozyme requiert la présence d'ions divalents pour son repliement, ce qui permet l'adoption de sa structure active [169, 172]. Différentes données biochimiques et structurales supportent le fait que la base générale, impliquée dans la déprotonation du groupement 2'-OH, serait le résidu G₃₃ tandis que l'acide général, qui facilite la protonation du groupement 5'-O⁻ partant, serait le groupement C2-amine du GlcN6P (Figure 1.12A; [67, 174]). Étant donné son rôle dans le métabolisme bactérien, le ribozyme *glmS* a le potentiel de servir de cible au développement d'antibiotiques [175, 176].

Le ribozyme HDV est un ribozyme contenu dans l'ARN génomique et antigénomique du virus de l'hépatite D (HDV). Le HDV est lui-même considéré comme étant un virus satellite du virus de l'hépatite B (HBV) qui se propage uniquement en présence du HBV, augmentant sa virulence et les symptômes de l'infection [177]. Le HDV est un virus formé d'un génome d'ARN circulaire qui code entre autre pour un ribozyme catalysant les réactions d'auto-clivage et d'auto-ligation nécessaires à son maintien [178]. Sa réplication est assurée par un mécanisme de *rolling circle* (cercle roulant) qui requiert l'ADN polymérase II de la cellule hôte [177-179]. Le ribozyme HDV minimal est formé de

90 nucléotides (5 nucléotides en amont et environ 85 nucléotides en aval du site de clivage) qui se répartissent en cinq domaines hélicaux (Figure 1.10C; [82, 180-182]). Deux pseudonoeuds stabilisent la structure et permettent de former le site actif [82, 183, 184]. Il est proposé qu'une molécule d'eau coordonné à ion Mg^{2+} joue le rôle de base générale et que le résidu C_{75} joue le rôle d'acide général dans le mécanisme enzymatique (Figure 1.12B; [111, 185]). Il est aussi intéressant de constater que le ribozyme retrouvé dans le gène *CPEB3* présenté précédemment adopte une structure *HDV-like*, ce qui soulève l'hypothèse que d'autres ribozymes semblables pourraient être retrouvés dans les génomes de mammifères (Section 1.3.2; [129]). Finalement, le ribozyme HDV est particulièrement stable, ce qui lui confère des propriétés utiles pour le développement d'outils biochimiques [186].

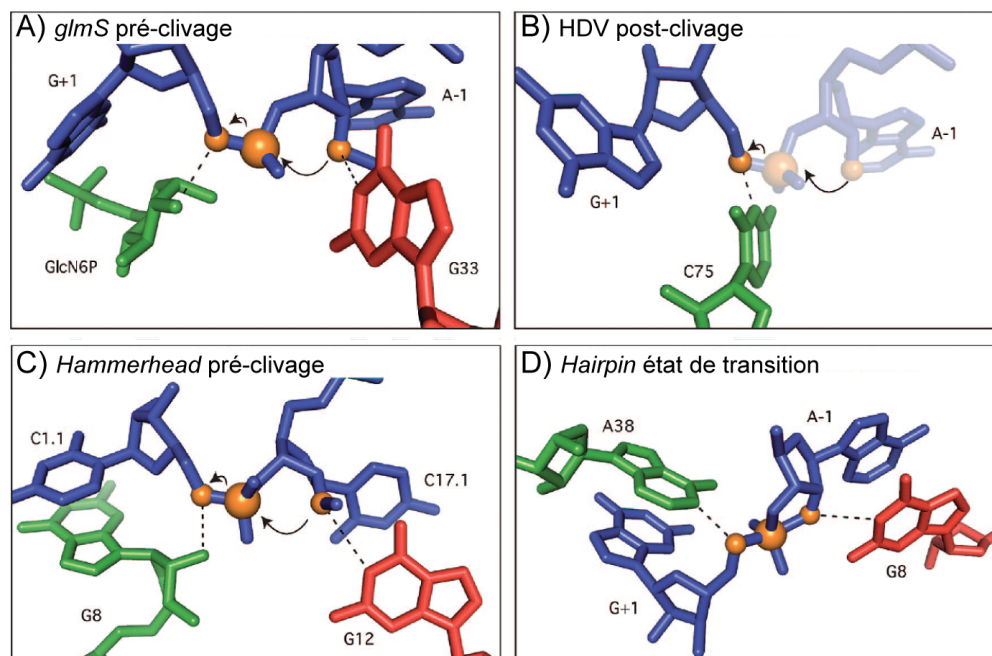


Figure 1.12. Structures du site actif de différents ribozymes nucléolytiques. (A) Site actif du ribozyme *glmS* pré-clivage, (B) site actif du ribozyme HDV post-clivage, (C) site actif du ribozyme *hammerhead* pré-clivage et (D) site actif du ribozyme *hairpin* en état de transition. Les nucléotides de part et d'autre du site de clivage sont en bleu, le phosphate scissile est en orange, la base générale et l'acide général sont en rouge et vert, respectivement. Tiré de [111].

Le ribozyme *hammerhead* a été initialement découvert dans les génomes de pathogènes de plantes, comme des viroïdes et des virusoïdes, où il participe au mécanisme de répllication par *rolling circle* [187-193]. Des analyses bioinformatiques ont plus tard

révélé que le ribozyme *hammerhead* est retrouvé dans pratiquement tous les domaines du vivant [160, 194, 195]. Il est formé de trois hélices s'orientant autour d'une jonction à trois voies impliquant environ 15 nucléotides conservés au site actif (Figure 1.10A; [191]). Durant de nombreuses années, les structures du *hammerhead* ne coïncidaient pas avec les données biochimiques [196]. L'utilisation d'un ribozyme plus long a mis en évidence que des interactions tertiaires se forment entre une boucle terminale et une boucle interne de deux hélices en présence de cations [90]. Ces contacts tertiaires engendrent un réarrangement du domaine catalytique qui stabilise la forme active du ribozyme et celle-ci a permis de réconcilier les données biochimiques et structurales. Ainsi, le rôle de base générale serait joué par le N1 du G₁₂ tandis que le rôle d'acide général serait rempli par le 2'-OH du G₈ (Figure 1.12C; [90, 165, 197]). Récemment, un ribozyme *hammerhead-like* a été identifié dans le gène *CLEC2* ce qui suggère que de nombreux autres ribozymes *hammerhead* pourraient être découverts dans certains ARNm [131, 198].

Le ribozyme *twister* est le plus récent membre de la famille des petits ribozymes nucléolytiques [70]. Il a été découvert par une analyse bioinformatique visant à identifier des ARN structurés par la comparaison des génomes de différentes espèces [199]. La séquence minimale du ribozyme *twister* est d'une quarantaine de nucléotides qui se répartissent en trois hélices, deux boucles internes et une boucle terminale [70]. Des hélices supplémentaires peuvent aussi se greffer aux boucles internes [70]. En présence de cations, la structure tertiaire est stabilisée par deux pseudonoeuds qui permettent l'adoption d'une structure active où une guanine jouerait le rôle de base générale [155-157]. Plusieurs autres études seront nécessaires pour caractériser le ribozyme *twister*, notamment pour identifier l'acide général participant à la catalyse [155, 156].

Le ribozyme *hairpin* est retrouvé dans certains ARN satellites de pathogènes de plantes, dont le *Tobacco ringspot virus*, où il participe à la réplication par *rolling circle* [187]. Sa séquence naturelle minimale contient environ 75 nucléotides qui s'orientent autour d'une jonction à quatre voies (Figure 1.10B; [133]). Deux boucles internes conservées (les boucles A et B) se retrouvent sur des tiges adjacentes. Ces deux boucles interagissent ensemble en présence de cations pour former le site actif du ribozyme *hairpin*. Chacune de ces boucles internes contient un nucléotide impliqué dans le mécanisme enzymatique [133, 200, 201]. En effet, la boucle A contient le résidu G₈ qui remplit la fonction de base générale tandis que la boucle B contient le résidu A₃₈ qui joue le rôle

d'acide général (Figure 1.12D; [202-204]). Il a été proposé que la stratégie catalytique du ribozyme *hairpin* soit similaire à celle du ribozyme VS de *Neurospora* malgré que, comme nous le verrons ci-dessous, ce dernier soit unique en son genre [205, 206].

1.4. Le ribozyme VS de *Neurospora*

1.4.1. Le ribozyme VS et son rôle dans le cycle de réplication

Le ribozyme VS de *Neurospora* a été identifié dans le laboratoire de Richard Collins à l'Université de Toronto en 1990. Il est un élément fonctionnel essentiel contenu dans l'ARN *Varkud Satellite* (VS), un ARN satellite présent dans les mitochondries de certains isolats du champignon *Neurospora* [134]. L'ARN VS monomérique, qui est constitué de 881 nucléotides, alterne entre une forme linéaire et circulaire selon l'étape de son cycle de réplication. L'ARN VS est retrouvé uniquement lorsque le plasmide *Varkud* (V) ou un plasmide homologue est également présent dans la mitochondrie [207, 208]. Le plasmide V est un rétro-plasmide qui encode une rétro-transcriptase nécessaire pour le cycle de réplication de l'ARN VS [207]. L'ARN VS se maintient dans la mitochondrie par un mécanisme de réplication par *rolling circle* où le ribozyme catalyse les réactions d'auto-clivage et d'auto-ligation essentielles au mécanisme. Le cycle de réplication de l'ARN VS débute par la transcription de l'ADN VS par l'ARN polymérase mitochondriale. La transcription génère un ARN multimérique qui inclut plusieurs ARN VS les uns à la suite des autres (Figure 1.13A; [134, 209]). L'activité de clivage du ribozyme VS permet la scission entre les résidus G₆₂₀ et A₆₂₁ ce qui engendre la monomérisation des ARN VS et la production d'extrémités 5'-OH et 2'-3' cyclique phosphate (Figure 1.13B; [134, 209]). Les monomères se referment ensuite sur eux-mêmes en se circularisant par la réaction inverse d'auto-ligation (Figure 1.13C). La rétro-transcriptase du plasmide V permet ensuite de synthétiser le brin d'ADN de polarité négative (-), ou ADN complémentaire (ADNc; Figure 1.13D). Le brin d'ARN est ensuite déplacé pour permettre la synthèse du brin d'ADN de polarité positive (+), ce qui mène à la production d'un ADN VS double brin, complétant ainsi le cycle de réplication (Figure 1.13E). Outre sa contribution à son mécanisme de réplication, l'ARN VS n'a aucune fonction connue dans la mitochondrie du champignon *Neurospora* [210].

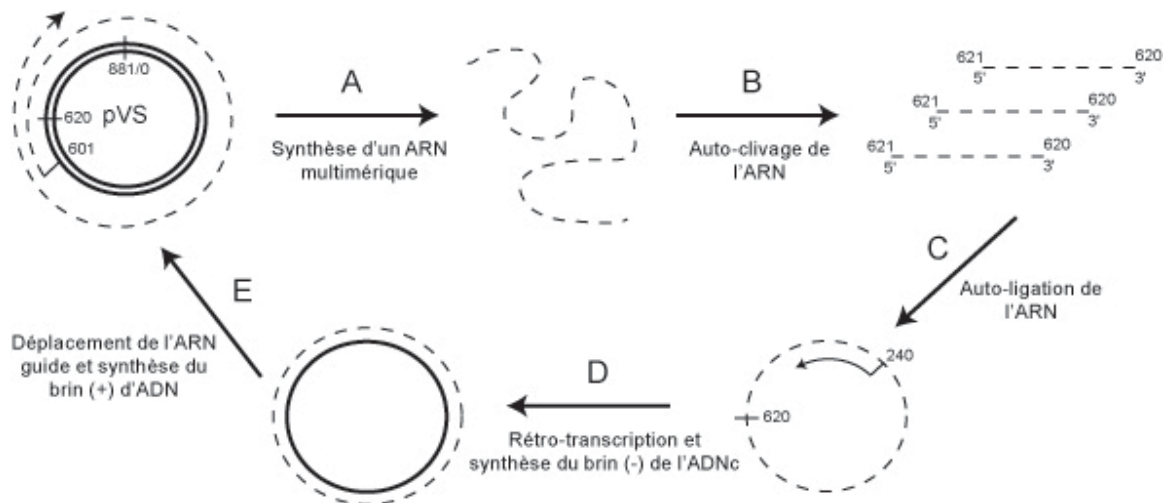


Figure 1.13. Mécanisme de répllication par *rolling circle* de l'ARN VS. (A) Synthèse de l'ARN multimérique, (B) auto-clivage, (C) auto-ligation par le ribozyme VS, (D) rétro-transcription et synthèse du brin (-) d'ADN et (E) déplacement du brin d'ARN et synthèse du brin (+) d'ADN. Les lignes pleines représentent l'ADN et les lignes pointillées l'ARN. Adapté de [207].

1.4.2. Caractérisation du ribozyme VS

À partir d'analyses complémentaires de phylogénie, de délétion de séquences et d'essais enzymatiques, le domaine minimal contigu responsable de l'activité de clivage a été identifié [211]. Ce domaine, nommé le ribozyme VS de *Neurospora*, est formé d'environ 150 nucléotides, dont un seul résidu (le G₆₂₀) est requis en amont du site de clivage (Figure 1.14; [211]). Le ribozyme VS se comporte de la même façon que dans l'ARN VS puisqu'il possède le même site de clivage (entre le G₆₂₀ et le A₆₂₁). De plus, le ribozyme VS requiert la présence de cations pour être actif et ce, autant dans son état isolé que dans l'ARN VS [209, 212]. Plus précisément, le ribozyme VS utilise préférentiellement les ions Mg²⁺ mais est également actif en présence de Mn²⁺, de Ca²⁺ et de concentrations élevées d'ions monovalents tandis que le cobalt hexamine (Co(NH₃)₆³⁺) inhibe son activité [212-215]. De plus, il a aussi été démontré que le ribozyme VS peut cliver un substrat isolé ou bien le substrat d'un autre ribozyme en *trans*, en accord avec son rôle physiologique dans le cycle de répllication de l'ARN VS [216, 217]. Ainsi, comme les caractéristiques précédentes sont partagées par le ribozyme VS autant sous sa forme isolée que dans l'ARN

VS, les études biochimiques et structurales subséquentes ont été réalisées sur la forme minimale contigüe ou des formes similaires du ribozyme VS (Figure 1.14; [209, 212, 214, 218, 219]).

1.4.3. Les domaines fonctionnels du ribozyme VS

La structure secondaire du ribozyme VS a été proposée à partir d'expériences de mutagénèse et de modifications chimiques (Figure 1.14; [218]). Le ribozyme VS est formé de six régions hélicales, nommées I à VI [218]. Ces six régions hélicales forment deux domaines principaux, soit le substrat et le domaine catalytique. Le substrat est formé par la tige-boucle I (ou SLI, pour *stem-loop* I) tandis que le domaine catalytique est formé par les régions hélicales II à VI. Le domaine du substrat est composé de deux tiges, soit les tiges Ia et Ib, d'une boucle terminale et d'une boucle interne où se situe le site de clivage entre les résidus G₆₂₀ et A₆₂₁ (Figure 1.14). Le domaine catalytique est quant à lui composé de cinq régions hélicales (II-VI) qui s'axent autour de deux jonctions à trois voies, nommées jonctions II-III-VI et III-IV-V (Figure 1.14). Certains domaines hélicaux comme les tiges II et III et la tige-boucle VI (SLVI) incluent des renflements et des boucles internes qui entrecoupent ces hélices, formant ainsi plusieurs sous-domaines hélicaux (IIa et IIb, IIIa et IIIb, VIa à VIc; [220]).

Sous sa forme naturelle, le ribozyme VS catalyse les réactions de clivage et de ligation sur son substrat qui est positionné en 5' du domaine catalytique. Il s'agit de l'activité en *cis* du ribozyme VS. Cependant, l'activité catalytique est aussi possible lorsque le substrat est positionné à l'extrémité 3' (permutation circulaire; [221]) ou encore lorsque le substrat est séparé du domaine catalytique (activité en *trans*; [214]).

Bien que la plupart des domaines hélicaux du ribozyme VS soient importants pour l'activité catalytique, des études de délétion dirigée ont démontré que certaines régions ne sont pas nécessaires. En particulier, la tige-boucle IV (SLIV) et le sous-domaine VIc peuvent subir une réduction de leurs paires de bases sans affecter la catalyse [221]. Aussi, lorsque le ribozyme est sous sa forme *trans* ou en permutation circulaire, la tige IIa peut aussi être réduite sans diminution d'activité [221]. Ainsi, à l'exception de ces régions, tous les autres domaines du ribozyme VS sont considérés comme des sous-domaines fonctionnels [221, 222].

Ces sous-domaines fonctionnels participent à la formation de deux interactions tertiaires qui sont cruciales pour l'activité catalytique du ribozyme VS. Premièrement, les boucles terminales des tiges-boucles I et V (SLI et SLV) interagissent ensemble pour former l'interaction *kissing-loop* I/V (Figure 1.14; [108]). Cette interaction, qui permet la reconnaissance et l'activation du substrat, est aussi facilitée par le repliement de la jonction III-IV-V [223]. Ensuite, le site actif du ribozyme VS est formé par l'association des boucles internes de SLI et de la boucle A730 retrouvée dans SLVI (Figure 1.14). Chacune de ces boucles internes contient un nucléotide catalytique, soit le G₆₃₈ dans SLI et le A₇₅₆ dans la boucle A730, qui participe directement au mécanisme enzymatique. En plus d'être facilitée par le repliement de la jonction II-III-VI, la formation du site actif est essentielle à l'activité catalytique du ribozyme VS [223]. Ces deux interactions tertiaires importantes, leurs rôles et les sous-domaines fonctionnels impliqués seront décrites en plus de détails aux sections 1.4.4 et 1.4.5.

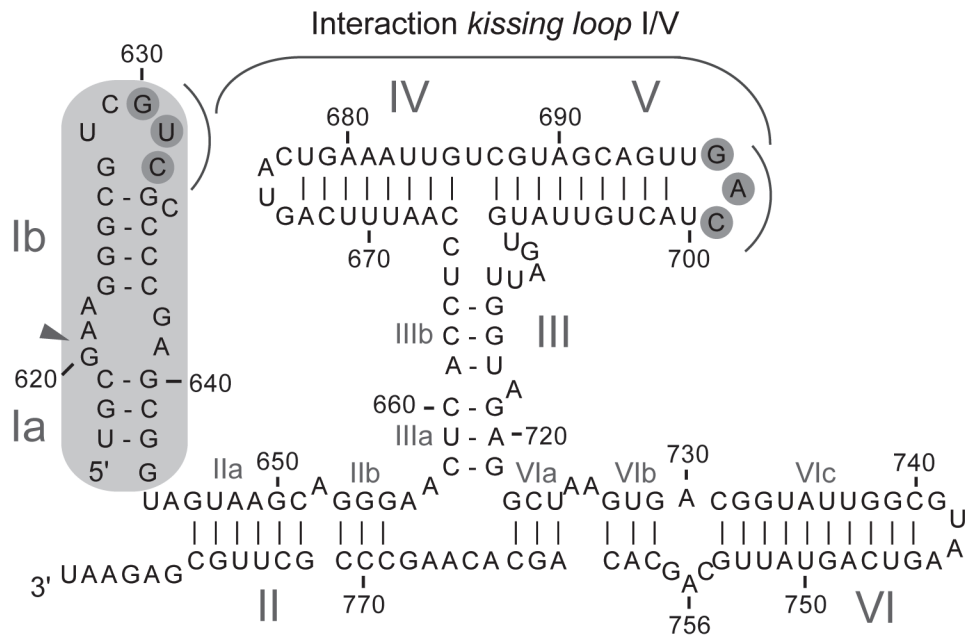


Figure 1.14. Structures primaire et secondaire du ribozyme VS de *Neurospora*. Les domaines hélicaux sont numérotés de I à VI. Le substrat, qui contient les domaines Ia et Ib, est ombragé en gris pâle, les domaines II à VI forment le domaine catalytique et les résidus impliqués dans l'interaction *kissing-loop* I/V sont ombragés en gris foncé. Le site de clivage est présenté par une flèche gris foncé. Les sous-domaines hélicaux sont désignés par une lettre. Adapté de [108].

1.4.4. L'interaction *kissing-loop* I/V et son rôle dans la reconnaissance et l'activation du substrat

Une des caractéristiques du ribozyme VS est que le domaine catalytique reconnaît son substrat par la formation d'une interaction *kissing-loop* entre les boucles terminales des domaines I et V [108]. Cette interaction, qui est dépendante du Mg^{2+} , permet de lier et de positionner le substrat dans le domaine catalytique, contribuant ainsi à l'adoption de la structure tridimensionnelle active du ribozyme VS [224, 225]. De plus, l'interaction *kissing-loop* I/V induit un changement conformationnel dans le substrat, lui conférant ainsi sa conformation active [225, 226]. L'interaction *kissing-loop* I/V est donc très importante puisqu'elle facilite l'activité catalytique du ribozyme VS, autant au niveau du clivage que de la ligation. Les sections suivantes présentent les détails moléculaires de la reconnaissance et de l'activation du substrat par l'interaction *kissing-loop* I/V.

1.4.4.1. La reconnaissance du substrat par l'interaction *kissing-loop* I/V

L'interaction *kissing-loop* I/V implique la formation d'une interaction à longue distance entre les boucles terminales des hélices I et V. Cette interaction se définit par la formation de trois paires de bases canoniques entre les résidus G_{630} , U_{631} et C_{632} de SLI et G_{697} , A_{698} et C_{699} de SLV [108]. Des nombreuses études de mutagenèse ont démontré l'importance de ces résidus dans la formation de l'interaction *kissing-loop* I/V [108, 225]. En effet, la mutation de ces résidus peut diminuer l'activité catalytique jusqu'à un facteur de 1000 fois [108]. De plus, des études de modifications chimiques ont montré que l'interaction *kissing-loop* I/V favorise le repliement compact du ribozyme tandis que les mutations qui affectent ces six résidus altèrent aussi le repliement du ribozyme [108].

Afin de former les paires de bases canoniques de l'interaction *kissing-loop* I/V, les boucles terminales des domaines I et V forment des motifs *U-turn* [108, 227]. Pour SLI, il a été proposé à partir d'études biochimiques que le motif *U-turn* soit formé par les résidus U_{628} à G_{630} [108]. Des études structurales ont par ailleurs démontré que ce motif *U-turn* n'est pas présent préalablement à la formation de l'interaction *kissing-loop* I/V [228] mais que celui-ci se structure simultanément à la formation des paires de bases entre SLI et SLV (Sections 1.5.1 et 1.5.3; [227, 229]). Chez SLV, ce sont plutôt les résidus U_{696} à A_{698} qui adoptent un motif *U-turn* (Section 1.5.2; [108, 230, 231]). Contrairement au motif *U-turn*

de SLI, celui de SLV est formé sous sa forme libre et contient des sites de liaison de métaux qui favorisent la compaction de ce motif [230, 231].

Bien que la nature des résidus des boucles de SLI et de SLV soit importante pour la formation de l'interaction *kissing-loop* I/V, il est intéressant de constater que des mutations compensatoires entre les boucles terminales de SLI et de SLV qui maintiennent les paires de bases et qui répondent aux nécessités structurales des *U-turn* permettent de supporter l'activité catalytique [225]. L'affinité de l'interaction *kissing-loop* I/V est particulièrement élevée, permettant à ces deux tige-boucles d'interagir ensemble indépendamment du reste du ribozyme VS [225, 229, 232].

De plus, l'interaction *kissing-loop* I/V dépend de la présence d'ions Mg^{2+} . En effet, en absence d'ions divalents, l'interaction *kissing-loop* n'est généralement pas détectée [108, 224]. Cette dépendance aux ions divalents est supportée par des études d'interférence aux phosphorothioates et de restauration avec le Mn^{2+} qui ont suggéré que les boucles terminales de SLI et SLV forment des sites de liaisons de métaux [108, 233]. En plus de nécessiter la présence d'ions Mg^{2+} , l'interaction *kissing-loop* I/V permettrait d'établir un équilibre catalytique entre le clivage et la ligation dépendant de la concentration d'ions Mg^{2+} [234].

Finalement, il a été démontré que le changement du nombre de paires de bases de la tige V par l'addition ou la soustraction de plus d'une paire de bases inhibe fortement l'activité catalytique du ribozyme VS, possiblement en empêchant la formation du site actif [221, 235]. Cependant, il est intéressant de constater que la compensation de la longueur des hélices entre les tiges Ib et V permet de maintenir l'activité catalytique du ribozyme VS [235]. Ainsi, le rôle principal de l'interaction *kissing-loop* I/V est de permettre la reconnaissance et la liaison du substrat mais aussi de positionner le substrat dans le domaine catalytique afin de former le site actif.

1.4.4.2. Activation du substrat par la formation de l'interaction *kissing-loop* I/V

Les premières expériences enzymologiques sur le ribozyme VS ont suggéré que l'étape limitante de la réaction enzymatique serait un réarrangement conformationnel du substrat [212, 214]. Il a ensuite été démontré que ce réarrangement conformationnel s'effectue en parallèle avec la formation de l'interaction *kissing-loop* I/V [225-227]. Préalablement à la formation de l'interaction *kissing-loop*, le substrat se présente sous sa

forme inactive ou *unshifted* (Figure 1.15A). Dans cette conformation, la boucle interne de SLI est composée d'une boucle symétrique de six nucléotides (G₆₂₀-A₆₂₂ et C₆₃₇-A₆₃₉) tandis que la tige Ib est formée de cinq paires de bases et la boucle terminale contient quatre résidus (Figure 1.15A). Sous cette forme, le substrat possède une faible affinité pour le domaine catalytique et l'activité catalytique est pratiquement inexistante [236]. Cependant, parallèlement à la formation de l'interaction *kissing-loop* I/V, qui est possible en présence d'ions Mg²⁺, le substrat subit un réarrangement conformationnel qui correspond à sa forme active ou *shifted* (Figure 1.15B; [225-227]). Ce réarrangement conformationnel implique une réorganisation des paires de bases de la tige Ib et l'exclusion du résidu C₆₃₄ (Figure 1.15B; [225-227]). De plus, supportant l'importance du réarrangement conformationnel, il a été montré que certaines mutations de la tige Ib diminuent ou éliminent l'activité catalytique en empêchant la réorganisation des paires de bases de la tige Ib [218, 226].

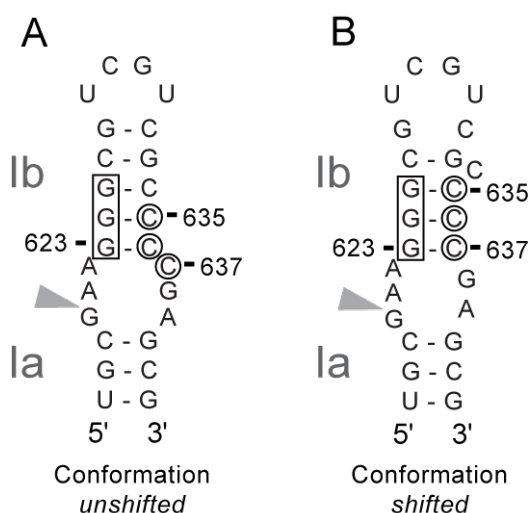


Figure 1.15. Structures primaire et secondaire du substrat dans les conformations *unshifted* et *shifted*. (A) Substrat en conformation *unshifted* et (B) en conformation *shifted*. Les trois guanines dont les partenaires changent lors du réarrangement conformationnel sont encadrées tandis que les trois cytidines qui se réorganisent sont encerclées. Le site de clivage est présenté par une flèche grise. Adapté de [226].

Parallèlement à la formation de l'interaction *kissing-loop* I/V, le réarrangement conformationnel permet à la boucle interne de SLI d'adopter une configuration asymétrique puisque le résidu C₆₃₇ est déplacé dans la tige Ib (Figure 1.15B). Cette conformation de la boucle interne de SLI correspond à sa forme catalytique active [225, 226]. En effet, les

mutations créant un substrat constitutivement *shifted* sont actifs tandis que ceux qui ne peuvent adopter cette conformation sont inactifs [236]. De plus, plusieurs études biochimiques ont proposé que le G₆₃₈ joue le rôle de base générale dans le mécanisme enzymatique du ribozyme VS (Section 1.4.6; [237-239]). Comme le C₆₃₇ est déplacé dans la tige Ib, ce réarrangement pourrait permettre de positionner le G₆₃₈ pour jouer son rôle dans la catalyse. Finalement, la conformation active contient un site de liaison de métaux unique qui pourrait faciliter la catalyse en stabilisant les intermédiaires réactionnels du mécanisme enzymatique [226, 240, 241].

1.4.4.3. Contribution de la jonction III-IV-V

La jonction III-IV-V est formée de résidus simple brin qui relie les tiges III, IV et V. Bien que cette jonction ne soit pas absolument nécessaire pour l'activité catalytique du ribozyme VS, elle favorise la catalyse en facilitant la liaison du substrat en positionnant SLV adéquatement afin de former l'interaction *kissing-loop* I/V [223]. La jonction III-IV-V est donc un domaine architectural important qui détermine l'orientation et le positionnement des domaines hélicaux III, IV et V [242, 243]. Dans le même ordre d'idée, le changement de la longueur de la tige III, qui connecte les jonctions III-IV-V et II-III-VI diminue l'activité catalytique, possiblement en modifiant l'orientation de SLV et par conséquent, le positionnement de SLI pour former le site actif [243]. De plus, des études de mutagenèse dirigée ont démontré que la mutation de la plupart des nucléotides de la jonction ainsi que de certaines paires de bases qui ferment les hélices diminue l'activité catalytique du ribozyme VS [242, 243]. Ces données de mutagenèse ont permis de proposer que les résidus U₇₁₀ à A₇₁₂ de la jonction III-IV-V formeraient un motif *U-turn* (Figure 1.14; [242]). Il est aussi intéressant de noter que le remplacement de ce motif *U-turn* par une tige, créant une jonction à quatre voies, permet de maintenir l'activité catalytique [242]. En résumé, ces données suggèrent que la jonction III-IV-V adopte une structure précise qui permet le positionnement de SLV, favorisant le bon positionnement du substrat dans le domaine catalytique.

1.4.5. Formation du site actif du ribozyme VS

1.4.5.1. Les boucles internes de SLI et de SLVI

Tel que mentionné précédemment, le site actif du ribozyme VS est formé par l'association de la boucle interne de SLI et de la boucle A730. Tout d'abord, la boucle interne de SLI est un élément essentiel pour l'activité catalytique du ribozyme VS. Sous sa forme active, il s'agit d'une boucle interne asymétrique riche en purines qui inclue les résidus G₆₂₀, A₆₂₁ et A₆₂₂ sur un brin et les résidus G₆₃₈ et A₆₃₉ sur le second brin (Figure 1.15B). La boucle interne de SLI forme une partie importante du site actif puisqu'elle inclut le site de clivage qui se situe entre les nucléotides G₆₂₀ et A₆₂₁. De plus, l'importance de la nature de plusieurs des résidus de la boucle interne de SLI a été démontrée par mutagenèse. En effet, la mutation des résidus A₆₂₁, A₆₂₂, et G₆₃₈ réduit significativement l'activité catalytique du ribozyme VS [226, 237]. Plus précisément, il est proposé que le résidu G₆₃₈ joue le rôle de base générale dans le mécanisme acide-base général [237, 239]. En effet, le remplacement du G₆₃₈ par des nucléotides analogues suggère que le groupement fonctionnel responsable de l'activité catalytique serait la position N1 de ce résidu [237, 244].

La boucle A730 est aussi un élément essentiel de l'activité du ribozyme VS [220, 221, 223, 245]. La boucle A730 est une boucle interne asymétrique constituée du résidu A₇₃₀ sur un brin et des résidus C₇₅₅, A₇₅₆ et G₇₅₇ sur le second brin (Figure 1.14). Les résidus de cette boucle sont importants pour l'activité du ribozyme VS puisque la mutation de chacun des nucléotides de la boucle A730 mène à une réduction significative de l'activité catalytique [220, 223]. De plus, l'ablation de la boucle A730 réduit la catalyse sous le seuil de détection, confirmant le rôle essentiel de cette boucle dans la formation du site actif [221]. Parmi toutes les études de mutations, de substitutions et de délétion, le nucléotide le plus sensible est le A₇₅₆ [220, 221, 223, 245]. En effet, la mutation ou la délétion de ce nucléotide peut diminuer l'activité catalytique jusqu'à 1000 fois [220, 223]. Plusieurs études biochimiques suggèrent d'ailleurs que le A₇₅₆ joue le rôle de l'acide général dans le mécanisme acide-base général du ribozyme VS [220, 239, 245]. De plus, le remplacement du A₇₅₆ par des nucléotides analogues suggère que le groupement fonctionnel responsable de l'activité catalytique serait la position N1 de ce résidu [245-247]. Finalement, des études d'interférence aux phosphorothioates ont également montré que la boucle A730 pourrait former un site de liaison de métaux [233]. Les expériences d'interférence et de restauration

(*rescue*) de l'activité par l'utilisation du Mn^{2+} suggèrent que les groupements phosphates des G_{757} et C_{758} pourraient lier directement un ion Mg^{2+} [233].

Un autre élément intéressant à propos des boucles internes de SLI et de SLVI est qu'elles partagent certaines similarités avec les boucles A et B du ribozyme *hairpin*. En effet, la boucle interne de SLI, qui contient le G_{638} s'apparente à la boucle A du ribozyme *hairpin* qui contient le G_8 qui joue également le rôle de base générale dans le mécanisme enzymatique du ribozyme *hairpin*. En ce qui concerne la boucle A730, qui contient le résidu A_{756} , celle-ci s'apparente à la boucle B du ribozyme *hairpin* qui contient quant à elle le résidu A_{38} qui joue aussi le rôle d'acide général dans le mécanisme de la réaction de clivage [205, 206, 237]. Ainsi, ces deux ribozymes semblent partager certaines caractéristiques architecturales et mécanistiques [205, 206, 237].

1.4.5.2. Le site actif est formé par l'association de la boucle interne de SLI et de la boucle A730

Puisque les boucles internes de SLI et de SLVI contiennent chacun un nucléotide catalytique, il est nécessaire que ces boucles s'associent pour former le site actif du ribozyme VS. Ainsi, il est nécessaire que le ribozyme VS adopte une structure qui permette à ces boucles internes d'interagir ensemble. Cette structure implique le repliement des deux jonctions à trois voies ainsi que la formation de l'interaction *kissing-loop* I/V qui permet de positionner le substrat dans le domaine catalytique. Malgré tout, très peu de données expérimentales ont permis de déterminer directement l'association de ces deux boucles internes. Une des évidences provient de l'utilisation d'un nucléotide photo-activable (4-thio-uridine) qui a été substitué au A_{621} (3' du site de clivage) dans la boucle interne de SLI et qui a permis, en faisant une expérience de *cross-linking*, de former un lien covalent entre ce résidu et le A_{756} , qui se trouve dans la boucle A730 [248]. Ainsi, dans des conditions favorables au clivage, la boucle interne de SLI et la boucle A730 se rapprochent suffisamment pour former un *cross-link* [248]. Ce rapprochement permettrait ainsi de positionner les nucléotides catalytiques à proximité du phosphate clivable [239].

Contrairement à l'interaction *kissing-loop* I/V, aucune donnée ne supporte que la boucle interne de SLI et la boucle A730 puissent interagir ensemble de façon stable avec des tige-boucles isolées, donc indépendamment du reste du ribozyme. De plus, il est suggéré que la formation du site actif est thermodynamiquement instable [229]. Ainsi,

l'absence d'interaction suffisamment forte entre ces deux boucles internes explique l'importance de l'interaction *kissing-loop* I/V dans la liaison et le positionnement du substrat dans le domaine catalytique.

1.4.5.3. Contribution de la jonction II-III-VI

La jonction II-III-VI, similairement à la jonction III-IV-V, est formée par les résidus simple brin qui connectent les tiges II, III et VI (Figure 1.14). Ce domaine architectural est essentiel pour l'activité du ribozyme VS puisqu'il facilite la formation du site actif [223]. Une des particularités de la jonction II-III-VI est que chacune des hélices qui s'y rattachent possède un renflement qui se situe à trois paires de bases de la jonction (Figure 1.14). Bien que ces renflements tolèrent certaines mutations, leur délétion réduit significativement l'activité catalytique, suggérant un rôle important dans le repliement du ribozyme [220, 243]. De plus, l'ajout de paires de bases dans la tige IIb ou dans la tige III diminue fortement l'activité catalytique [220, 243].

Par ailleurs, la nature des résidus qui forment les régions simple brin de la jonction est très importante puisque leur mutation ou l'effet de modifications chimiques réduit considérablement l'activité du ribozyme, signifiant qu'ils sont importants pour le repliement de cette jonction [222, 249]. Plus précisément, des études de mutagenèse dirigée ainsi que le remplacement de la jonction II-III-VI par la jonction 7-5-6 de l'ARNr 23S de *Haloarcula marismortui* suggèrent que les résidus A₆₅₆ et G₇₆₈ pourraient former un motif *G-wedge* [249]. Finalement, les études de modifications et de protections chimiques démontrent aussi que la jonction II-III-VI requiert la présence d'ions Mg²⁺ pour se replier [218, 222, 224, 233, 249, 250]. Ainsi, ces données suggèrent que la jonction II-III-VI adopte une structure précise qui facilite la formation du site actif du ribozyme VS de *Neurospora*.

1.4.6. Le mécanisme enzymatique du ribozyme VS

Le repliement des jonctions III-IV-V et II-III-VI et la formation des deux interactions tertiaires critiques permettent l'adoption de la structure tridimensionnelle active du ribozyme VS [250]. Pour qu'il y ait clivage, le substrat doit subir un réarrangement

conformationnel et la boucle interne de SLI et la boucle A730 de SLVI doivent s'associer pour former le site actif. Ce site actif contient les nucléotides G₆₂₀ et A₆₂₁, qui sont séparés par le phosphate clivable, ainsi que les résidus catalytiques G₆₃₈ et A₇₅₆. Comme tout site actif, celui du ribozyme VS forme un microenvironnement où les conditions physicochimiques sont altérées. Dans ce site actif, il faut donc que les résidus G₆₃₈ et A₇₅₆ possèdent les états d'ionisation nécessaires au transfert de protons du mécanisme enzymatique acide-base général. L'étude de l'activité catalytique du ribozyme VS sous différent pH révèle une dépendance du pH qui suit une courbe en forme de cloche (*bell-shaped*), correspondant à un modèle mécanistique à double ionisation [237, 238, 246, 251]. Ces états d'ionisation coïncident avec une base déprotonée et un acide protoné pour une réaction de clivage et aux états inverses pour la ligation [237, 238, 246, 251]. Bien que les profils de dépendance au pH supporte l'implication du G₆₃₈ et du A₇₅₆ dans le mécanisme enzymatique, leur rôle exact ne peut être attribué sur la base de ces résultats. L'utilisation de groupements phosphorothiolates sur le groupement phosphate du A₆₂₁ a permis de confirmer que le A₇₅₆ joue le rôle d'acide général [239]. En effet, en présence d'un groupement phosphate naturel (5'-PO₄), la mutation du A₇₅₆ réduit l'activité de plus de 1000 fois puisque la présence de l'acide général est nécessaire pour protoner le groupe partant (5'-oxyanion) et ainsi achever le mécanisme enzymatique. Cependant, comme la substitution d'un oxygène par un soufre dans un groupement phosphorothiolate (5'-PS) en fait un meilleur groupe partant, l'activité catalytique est maintenue en présence d'un A₇₅₆ muté. Ces résultats confirment donc que le G₆₃₈ joue le rôle de base générale et le A₇₅₆ le rôle d'acide général dans le mécanisme enzymatique du ribozyme VS [239].

Pour jouer leur rôle respectif, les G₆₃₈ et A₇₅₆ doivent adopter un état d'ionisation adéquat. La valeur des pK_a des groupements N1 d'adénosines mono-phosphate (AMP) et de guanosines mono-phosphate (GMP) en solution est de 3.5 et de 9.2, respectivement [159]. Ainsi, à pH neutre, le N1 d'une adénosine est généralement déprotoné et celui d'une guanosine est généralement protoné. Cependant, dans le site actif du ribozyme VS, le A₇₅₆ doit être protoné et le G₆₃₈ doit être déprotoné pour permettre la réaction enzymatique. Selon des études de dépendance au pH et de substitution par des nucléotides analogues, le pK_a du N1 du A₇₅₆ serait d'environ 5.2-5.8 tandis que celui du G₆₃₈ de 8.3-8.4, [238, 239, 251]. Ces données suggèrent que la proportion de ribozymes ayant un A₇₅₆ protoné et un

G_{638} déprotoné à pH neutre augmente par rapport aux NMP, ce qui contribue à augmenter la vitesse de réaction [252].

Bien que ces résultats démontrent que les valeurs de pK_a des G_{638} et A_{756} se rapprochent de la neutralité, la manière dont l'altération de ces pK_a est accomplie dans le site actif du ribozyme VS est moins bien comprise. L'altération des pK_a s pourrait être rempli par la coordination d'un ion métallique dans le site actif. Différentes études ont démontré que la boucle interne de SLI sous la forme *shifted* et que la boucle A730 forment des domaines de liaison de métaux [226, 233, 241, 253]. De plus, des études de dépendance au pH en présence de différents métaux ont démontré que la nature de l'ion coordonné au site actif modifie le pK_a des nucléotides catalytiques en leur permettant de s'approcher de la neutralité [254, 255].

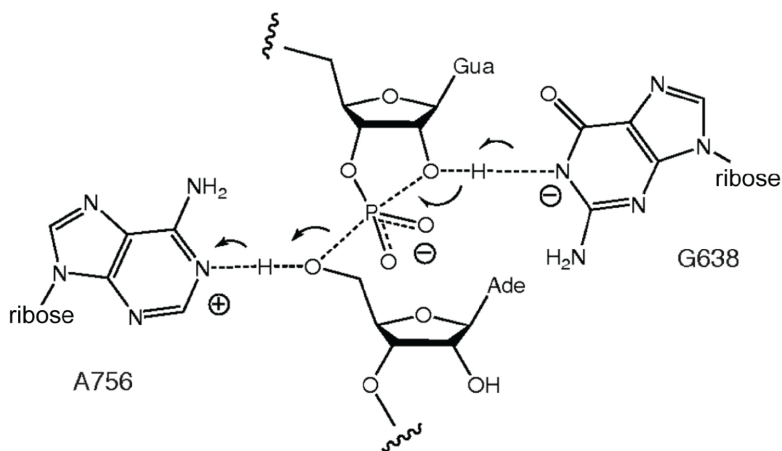


Figure 1.16. Mécanisme enzymatique proposé du clivage du ribozyme VS de *Neurospora*. Le G_{638} joue le rôle de base générale et le A_{756} le rôle d'acide général. Les flèches représentent les transferts de protons et les attaques nucléophiles tandis que les lignes pointillées représentent les liens chimiques formés ou défaits. Adapté de [239].

Finalement, le mécanisme enzymatique proposé du ribozyme VS commence par un transfert de proton du 2'-OH du G_{620} au G_{638} $N1^-$ (Figure 1.16A; [237, 239]). Le groupement 2'-O⁻ activé fait ensuite une attaque nucléophile sur le groupement phosphate adjacent pour former l'intermédiaire réactionnel sous forme trigonale bipyramidale. Ensuite, le groupement N1 protoné du A_{756} , qui joue le rôle d'acide général dans la réaction, fait un transfert de proton au 5'-oxyanion du A_{621} . La réaction produit des extrémités 5'-OH pour le A_{621} et 2'-3' cyclique phosphate pour le G_{620} . Les deux nucléotides catalytiques ont

quant à eux inversés leur état de protonation respectif. Dans la réaction de ligation, le A₇₅₆ fait une attaque nucléophile sur le groupement 5'-OH du A₆₂₁ pour initier le mécanisme enzymatique tandis que l'oxyanion en 2' du G₆₂₀ fait un transfert de proton au G₆₃₈.

1.5. Structures tridimensionnelles du ribozyme VS de *Neurospora*

L'obtention de la structure des macromolécules biologiques a toujours été un aspect important de caractérisation biochimique de ces molécules. En effet, la visualisation de ces structures permet de mieux comprendre leur mécanisme tout en permettant d'expliquer les données biochimiques. Présentement, plusieurs techniques permettent d'obtenir des données structurales sur des macromolécules. Entre autre, les techniques de *fluorescence resonance energy transfer* (FRET) et de *small-angle X-ray scattering* (SAXS) sont des techniques biophysiques qui sont utilisées pour obtenir des données à basse résolution sur un système. Elles sont souvent utilisées sur l'ARN et permettent de définir la structure globale d'un système mais sans pouvoir présenter le détail atomique des structures.

Les deux techniques principales pour déterminer la structure à haute résolution de macromolécules sont la cristallographie par diffraction de rayons X et la résonance magnétique nucléaire (RMN). Les structures d'ARN à haute résolution disponibles ont été obtenues pratiquement à parts égales par ces deux méthodes. Un des avantages de la RMN est qu'il s'agit d'une technique qui permet l'étude des systèmes dynamiques. Malgré que cette approche ait été historiquement limitée à l'étude d'ARN de moins de 30 nucléotides, des avancées techniques importantes permettent maintenant l'études de systèmes beaucoup plus gros [256]. Aussi, la RMN possède des avantages importants en permettant la caractérisation des systèmes à l'étude (dynamique, étude de pK_a, localisation de métaux; Section 1.6).

Plusieurs études structurales ont été accomplies sur le ribozyme VS de *Neurospora*. Bien qu'aucune structure cristalline ne soit rapportée jusqu'à maintenant, des études de FRET, de SAXS et de RMN ont quant à elles permis d'obtenir des données structurales sur le ribozyme VS. Afin de faciliter l'étude par RMN du ribozyme VS une approche de *divide-and-conquer* a été utilisée dans notre laboratoire. Selon cette approche, les différents sous-domaines importants du ribozyme sont étudiés de façon isolée. Ci-dessous, une brève description des structures RMN et des modèles de FRET et de SAXS est présentée.

1.5.1. Structures RMN du substrat du ribozyme VS

Le substrat du ribozyme VS est le domaine qui a été le plus étudié d'un point de vue structural. En effet, trois études RMN se sont penchées sur la structure adoptée par le substrat sous sa forme inactive et sous sa forme active.

1.5.1.1. Structures RMN de SLI inactif

Deux études de RMN se sont intéressées à la structure du substrat sous sa forme inactive (*unshifted*). Dans un cas, la boucle terminale a été remplacée par une boucle GNRA (Figure 1.17A; [257]) tandis que dans l'autre, le substrat naturel a été utilisé (Figure 1.17B; [228]). Dans les deux cas, les mêmes motifs structuraux ont été identifiés dans la boucle interne. Ces structures ont révélé que les résidus G₆₂₀ et A₆₃₉ ainsi que A₆₂₁ et G₆₃₈ forment des paires de bases *trans Sugar edge*/Hoogsteen G-A similaires à celles retrouvées dans les boucles GNRA (Figure 1.17). Les résidus A₆₂₂ et C₆₃₇ forment quant à eux une paire de bases *cis WC/WC A(+)*•C (Figure 1.17). Cette paire de bases est dépendante du pH puisqu'elle nécessite la protonation de l'adénine. En effet, le pK_a du A₆₂₂ est altéré à une valeur de 6.1-6.4 selon ces études [228, 257]. Comme la protonation du A₆₂₂ est requise pour que la paire de bases *cis WC/WC A(+)*•C se forme, il a été démontré que la déprotonation de ce résidu engendre un changement conformationnel important résultant en une forme ouverte de la paire de bases [228]. Cet équilibre conformationnel serait possiblement important pour permettre le réarrangement conformationnel du substrat qui se fait parallèlement à la formation de l'interaction *kissing-loop* I/V [228].

La structure de la boucle terminale naturelle de SLI a également été obtenue en absence de SLV (Figure 1.17B; [228]). Cependant, il semble que la boucle terminale de SLI soit désordonnée en absence d'interaction *kissing-loop*. De plus, il semble que la paire de bases G₆₂₇-C₆₃₂ de la tige Ib soit instable et que le motif *U-turn* proposé par des études biochimiques ne soit pas formé. Finalement, étant donné que les résidus de la boucle ne sont pas pairés, ils sont donc disponibles pour former des paires de bases avec la boucle terminale de SLV.

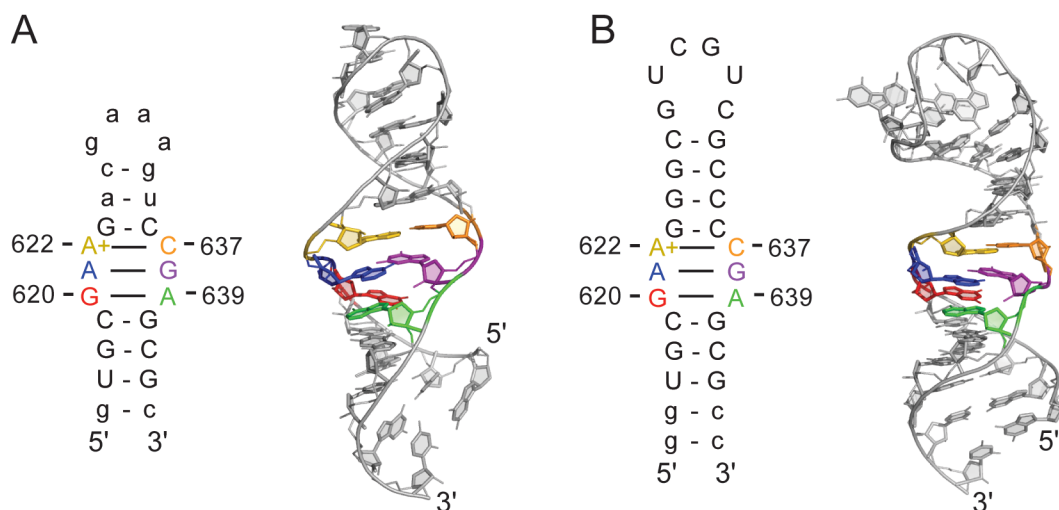


Figure 1.17. Structures RMN du substrat dans la conformation *unshifted* et inactive. Structure primaire, secondaire et tertiaire du substrat provenant de l'étude (A) de Michiels et al. (PDB : 1E4P; [257]) et (B) de Flinders et al. (PDB : 1HWQ; [228]). Les nucléotides en lettres minuscules ne font pas partie de la séquence naturelle. Le même système de couleur est utilisé pour les deux structures : G₆₂₀ en rouge, A₆₂₁ en bleu, A₆₂₃ en jaune, C₆₃₇ en orange, G₆₃₈ en mauve et A₆₃₉ en vert. Adapté de [228, 257].

1.5.1.2. Structure RMN de la forme active de SLI

Une seule étude structurale s'est intéressée au substrat sous sa forme active (*shifted*) [253]. Dans cette étude, la boucle terminale a été remplacée par une boucle GNRA et l'étude s'est principalement concentrée sur la structure RMN de la boucle interne de SLI (Figure 1.18). Dans cette structure, les résidus G₆₂₀ et A₆₃₉ forment encore une paire de bases *trans Sugar edge*/Hoogsteen G-A identique à la structure du substrat dans sa forme inactive. Cependant, des changements majeurs impliquent les résidus A₆₂₁, A₆₂₂ et G₆₃₈ puisque le C₆₃₇ forme désormais une paire de bases canonique avec le G₆₂₃ dans la tige Ib (Figure 1.18). Le G₆₃₈ forme deux paires de bases simultanément avec les A₆₂₁ et A₆₂₂, soit une paire de bases *trans Sugar edge*/Hoogsteen G-A partagée avec ces deux résidus (Figure 1.18). Le pK_a estimé du A₆₂₂ est d'environ 4.0 dans cette structure, suggérant que la déprotonation du A₆₂₂ de la forme inactive, en parallèle avec la formation de l'interaction *kissing-loop* I/V, favorise le réarrangement de la boucle interne de SLI. Il est particulièrement intéressant de noter que dans la forme active de SLI, la face WC du G₆₃₈ est positionnée à proximité du groupement 2'-OH du G₆₂₀, en accord avec son rôle de base générale dans le mécanisme enzymatique proposé du ribozyme VS.

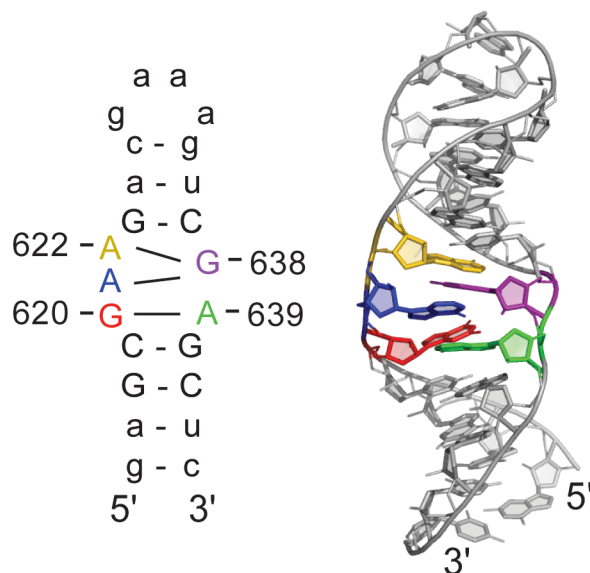


Figure 1.18. Structure RMN du substrat dans la conformation *shifted* et active. Structure primaire, secondaire et tridimensionnelle du substrat provenant de l'étude de Hoffmann et al. (PDB : 1OW9; [253]). Le système de couleur est le même qu'à la Figure 1.17 à l'exception du C₆₃₇ qui est noir puisqu'il forme maintenant une paire de bases canonique. Adapté de [253].

Les études RMN ont également permis d'identifier un site de liaison de métaux potentiel dans la boucle interne de SLI sous sa forme active [253]. Ce site s'apparente à un site retrouvé dans le ribozyme *hammerhead* et est supporté par des données de mutagenèse et d'interférence aux phosphorothioates [241]. Ce site de liaison de métaux qui n'est pas observé dans la conformation inactive pourrait être important pour la formation du site actif ou encore pour la stabilisation de l'intermédiaire réactionnel durant la réaction enzymatique. De plus, la recherche de l'occurrence de cette structure de boucles internes dans d'autres structures d'ARN a permis d'identifier deux boucles internes similaires dans des ARNr. Ces deux boucles internes forment des interactions tertiaires avec d'autres motifs d'ARN, supportant la formation du site actif du ribozyme VS via des interactions tertiaires similaires. Ainsi, la comparaison des structures inactives et actives de la boucle interne de SLI permet d'avancer des hypothèses sur les caractéristiques structurales potentiellement importantes pour l'activité, soit l'orientation du G₆₃₈ par rapport au phosphate clivable, la liaison d'un métal et la possibilité de former une interaction tertiaire.

1.5.2. Structures RMN de la tige-boucle V

Des analyses structurales par RMN ont aussi permis d'obtenir la structure à haute résolution de la boucle terminale de SLV en absence et en présence d'ions Mg^{2+} [230, 231]. Dans les deux cas, la paire de bases canonique qui ferme la boucle est composée des résidus U_{695} et A_{701} (Figure 1.19). En absence d'ions Mg^{2+} , les résidus U_{696} , G_{697} et A_{698} forment un motif *U-turn* mais celui-ci possède quelques différences significatives avec un motif *U-turn* canonique (Figure 1.19B; [230]). Ces différences résultent en un motif *U-turn* qui est relâché au lieu d'être compact (Figure 1.19B). L'ajout d'ions Mg^{2+} engendre des changements conformationnels qui permettent principalement de stabiliser la boucle terminale de SLV (Figure 1.19C; [231]). Entre autre, le *U-turn* adopte, dans cette condition, une structure beaucoup plus compacte qui est en accord avec la majorité des caractéristiques des motifs *U-turn*. De plus, quatre sites de liaison de métaux ont été identifiés par RMN en accord avec les données d'interférence aux phosphorothioates [233]. En présence d'ions Mg^{2+} , le U_{700} est davantage exclu de la structure de la boucle. Malgré tout, dans les deux structures, les faces WC des résidus G_{697} , A_{698} et C_{699} sont exposées au solvant et sont disponibles pour former des paires de bases avec la boucle terminale de SLI. Ainsi, la structure RMN de la boucle terminale de SLV en présence d'ions Mg^{2+} indique que cette boucle est repliée dans une conformation compatible avec la formation de l'interaction *kissing-loop I/V*.

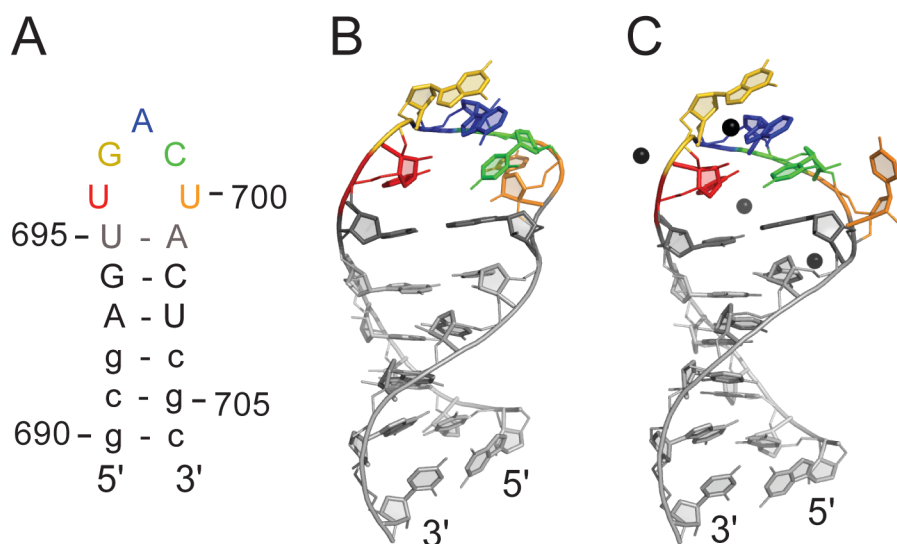


Figure 1.19. Structures RMN de la tige-boucle V. (A) Structures primaire et secondaire de la tige-boucle V, (B) structure tridimensionnelle de la tige-boucle V en absence de métaux divalents (PDB : 1TBK; [230]) et (C) structure tridimensionnelle en présence d'ions Mg^{2+} (PDB : 1YN2; [231]). Les résidus U_{695} et A_{701} sont gris foncé, le U_{696} est rouge, le G_{697} est jaune, le A_{698} est bleu, le C_{699} est vert et le U_{700} est orange. En (C), les ions Mn^{2+} sont en noir. Adapté de [230, 231].

1.5.3. Structure RMN de l'interaction *kissing-loop* I/V

Suite à l'obtention de la structure à haute résolution des tige-boucles SLI et SLV, une autre étude RMN s'est intéressée aux bases structurales permettant la reconnaissance du substrat via la formation de l'interaction *kissing-loop* I/V [227]. En plus de déterminer la structure de l'interaction boucle-boucle, cette étude a aussi démontré par RMN le réarrangement conformationnel de SLI parallèlement à l'interaction avec SLV [227]. Au niveau de l'interaction, les deux boucles terminales forment des *U-turn* compacts permettant la formation de trois paires de bases canoniques entre les boucles terminales (Figure 1.20). En effet, les protons iminos de ces paires de bases sont visibles et des réseaux de NOEs apparaissent dans les spectres RMN (Section 1.6), confirmant la formation des paires de bases intermoléculaires $G_{630}-C_{699}$, $U_{631}-A_{698}$ et $C_{632}-G_{697}$ au contact des deux boucles terminales (Figure 1.20). Cette structure révèle aussi que l'interaction est stabilisée par l'empilement des paires de bases de l'interaction *kissing-loop* avec les paires de bases des hélices SLI et SLV (Figure 1.20). De plus, des triplets de bases se forment à la

jonction entre les trois hélices, soit les hélices des deux tiges et celle de l'interaction boucle-boucle. Ainsi, le C₆₂₉ de SLI interagit dans le sillon mineur de la paire de base U₆₉₅-A₇₀₁ de SLV tandis que le G₆₃₄ interagit avec la paire de bases C₆₂₆-G₆₃₃ de SLI. L'interaction *kissing-loop* I/V forme également un angle interhélicale variant entre 127-175°. Cet angle peut être important pour le positionnement du substrat dans le domaine catalytique et pour la formation du site actif.

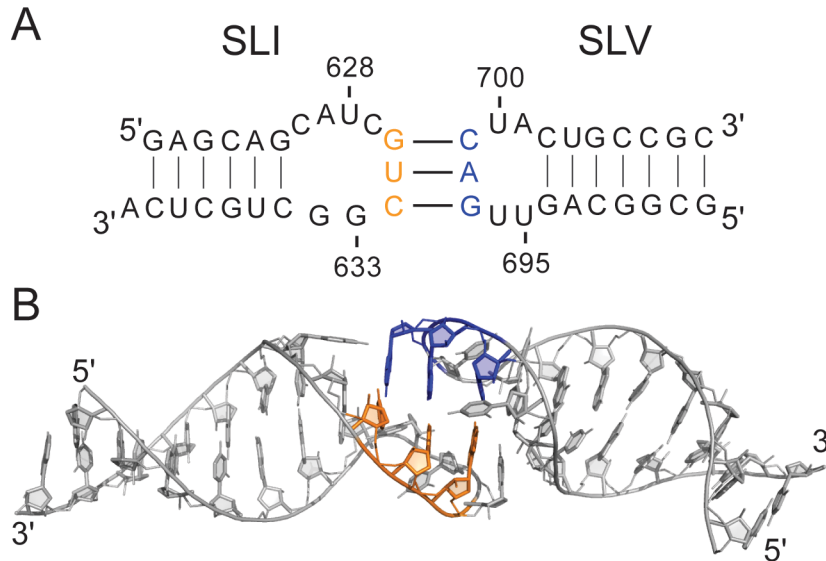


Figure 1.20. Structure RMN de l'interaction *kissing-loop* I/V. (A) Structure primaire et secondaire des tige-boucles I et V et (B) structure tridimensionnelle de l'interaction *kissing-loop* I/V (PDB : 2MI0; [227]). Les résidus G₆₃₀, U₆₃₁ et C₆₃₂ de l'hélice I sont colorés en orange tandis que les résidus G₆₉₇, A₆₉₈ et C₆₉₉ de l'hélice V sont colorés en bleu. Adapté de [227].

1.5.4. Structure RMN de la boucle A730

Une structure à haute résolution de la boucle A730 de SLVI a aussi été obtenue par RMN mais celle-ci ne concorde pas avec les données biochimiques [258]. Une première partie de cette étude a démontré que la présence du renflement formé des résidus A₇₂₅-A₇₂₆ engendre un réarrangement conformationnel dans SLVI menant à une réorganisation de la boucle A730 [258]. La boucle A730 devient alors symétrique et est formée de trois paires de bases dont une paire de base *cis* WC/WC G-U (Figure 1.21; [258]). Les résidus A₇₃₀ et C₇₅₅ forment une paire de bases *cis* WC/WC A(+) \cdot C, tandis que les résidus G₇₂₉ et A₇₅₆

forment une paire de bases *trans Sugar edge*/Hoogsteen G-A (Figure 1.21). Cette structure est stable en absence et en présence d'ions Mg^{2+} . Bien que cette structure soit supportée par les données RMN, elle est en désaccord avec l'ensemble des données biochimiques disponibles sur le ribozyme VS. En effet, le réarrangement de la structure secondaire de SLVI n'est supporté par aucune autre étude. De plus, le A_{756} qui doit participer à la réaction enzymatique forme une paire de bases et n'est donc pas disponible pour l'activité catalytique. Ainsi, cette structure RMN ne permet pas d'expliquer les données biochimiques. Il est donc important d'obtenir une structure à haute résolution de la boucle A730 qui concorderait avec les données biochimiques.

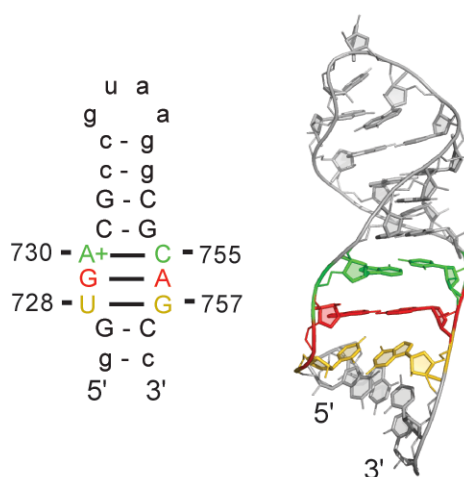


Figure 1.21. Structure RMN de la boucle A730. Structure primaire, secondaire et tertiaire de la boucle A730 de l'hélice VI (PDB : 1TJZ; [258]). Les résidus U_{728} et G_{757} sont en jaune, les résidus G_{729} et A_{756} sont en rouge et les résidus A_{730} et C_{755} sont en vert. Adapté de [258].

1.5.5. Modèles à basse résolution des jonctions II-III-VI et III-IV-V

Des modèles à basse résolution des deux jonctions à trois voies ont été obtenus par des études de FRET (Figure 1.22; [243, 249]). Le modèle de FRET de la jonction II-III-VI est dérivé d'un ARN qui inclut uniquement les résidus non paillés de la jonction qui sont connectés à des hélices modèles (Figure 1.22A; [249]). Ainsi, tous les renflements des hélices II, III et VI ont été retirés. Les analyses par gels natifs et par FRET démontrent qu'en présence d'ions Mg^{2+} , les hélices III et VI forment un empilement coaxial tandis que l'hélice II est positionnée près de l'hélice VI, ces deux tiges définissant un angle aigu

d'environ 60°. Le repliement de la jonction II-III-VI est dépendant de la présence d'ions Mg^{2+} [249]. Cette dépendance a également été observée dans une autre étude de FRET impliquant le ribozyme VS complet [250].

Le modèle global de la jonction III-IV-V a également été obtenu par FRET [243]. Selon ce modèle, les hélices III et IV formeraient un empilement coaxial et les hélices III et V seraient positionnées côtes-à-côtes (Figure 1.22B; [243]). L'ajout d'ions Mg^{2+} engendrerait le repliement de la jonction III-IV-V. Cependant, ces données ne sont pas supportées par une autre étude de FRET à molécule unique qui présente la jonction III-IV-V comme une jonction stable en absence et en présence d'ions Mg^{2+} [250]. De plus, des études de gel natif ont démontré que la mutation du U_{686} , qui est le seul résidu non-pairés entre les hélices IV et V, affecte significativement le repliement de la jonction III-IV-V [243]. En effet, la mutation de ce résidu favorise plutôt l'empilement coaxial des hélices IV et V au lieu des hélices III et IV [243].

Bien que les modèles de FRET des deux jonctions soient informatifs sur le repliement global de ces domaines architecturaux, ils ne donnent qu'une idée générale de la structure. Le détail atomique de ces jonctions est manquant, ce qui ne permet pas d'expliquer précisément les données biochimiques. Des structures à haute résolution sont donc nécessaires pour comprendre le rôle de chacun des nucléotides dans le repliement de ces jonctions.

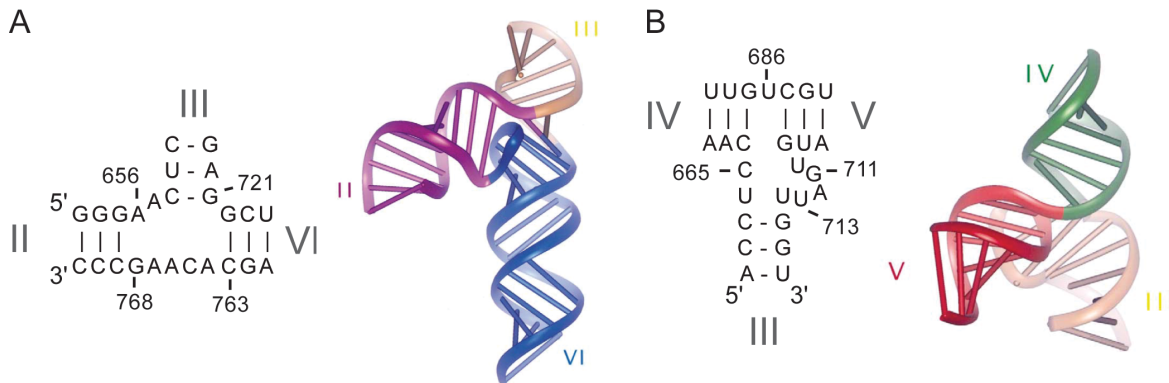


Figure 1.22. Modèles de FRET des jonctions II-III-VI et III-IV-V. Structure primaire et secondaire et modèle tridimensionnel de (A) la jonction II-III-VI et (B) la jonction III-IV-V. Adapté de [243, 249, 259].

1.5.6. Modèles à basse résolution du ribozyme VS

Trois différents modèles à basse résolution du ribozyme VS complet ont été obtenus à partir de données biochimiques et d'études par FRET et par SAXS (Figure 1.23). Un des éléments importants à considérer est que dans tous les cas, ces modèles donnent une idée globale du repliement du ribozyme mais n'offrent aucune information sur le détail structural des jonctions et des interactions tertiaires (*kissing-loop* et site actif) qui permettent la reconnaissance du substrat et son clivage.

Tout d'abord, le modèle biochimique provient des données de protection aux radicaux hydroxyl (Figure 1.23A; [224]). Selon ces données, le ribozyme maintient un nombre important d'interactions, même en absence du substrat, ce qui diffère de la plupart des autres ribozymes. En présence du substrat, l'hélice I se positionne entre les hélices II et VI ce qui permet de former le site actif entre la boucle interne du substrat et la boucle A730. Ce modèle présente également l'interaction *kissing-loop* I/V dans un arrangement coaxial où les hélices I et V forment un empilement continu.

Le modèle de FRET a été construit à partir des études de FRET réalisées sur les jonctions II-III-VI et III-IV-V (Figure 1.23B; [243, 249]). Ce modèle présente le domaine catalytique du ribozyme VS où les hélices IV, III et VI forment un empilement continu. Les hélices II et V définissent des angles avec le segment principal (Figure 1.23B). Selon des études supplémentaires de mutagenèse et de FRET, l'angle entre les hélices II et V est estimé à environ 70° [243]. La modélisation d'un substrat en *trans* dans ce modèle permet de rationaliser la formation du site actif (Figure 1.23B). Cependant, cette modélisation incorpore une interaction *kissing-loop* entre les hélices I et V qui n'est basée sur aucune donnée expérimentale et qui diffère de la structure RMN de cette interaction.

Finalement, des expériences de SAXS ont également été effectuées sur différentes constructions du ribozyme VS (Figure 1.23C; [260]). Ces constructions incluent les deux jonctions triples, le domaine catalytique du ribozyme VS ainsi que la forme *cis* à laquelle s'ajoute l'hélice VII, qui favorise l'équilibre vers la ligation. Pour les deux constructions du ribozyme VS, le rayon de giration diminue avec l'ajout de Mg²⁺, suggérant un repliement plus compact de ces ARN. Les modèles de FRET ont par la suite été utilisés pour remplir les enveloppes de SAXS et générer la structure globale du ribozyme VS (Figure 1.23C). Dans le modèle de SAXS, les hélices III, IV et VI forment un empilement coaxial similaire

au modèle de FRET. L'hélice V forme un angle aigu d'environ 40° avec l'hélice III tandis que l'hélice VI forme un empilement coaxial avec l'hélice II (Figure 1.23C). Selon ce modèle structural, l'interaction *kissing-loop* I/V définit un angle interhélicale similaire à ce qui est retrouvé dans la structure RMN [260]. Finalement, ce modèle est en accord avec la formation d'un site actif entre la boucle interne de SLI et la boucle A730.

Cependant, bien que ces modèles fournissent une structure globale pour le ribozyme VS, aucun de ces modèles à basse résolution ne permet d'expliquer d'un point de vue atomique les interactions prenant place dans le ribozyme VS. Ainsi, il est nécessaire d'obtenir les structures à haute résolution des sous-domaines importants manquants, soit la boucle A730 et les jonctions II-III-VI et III-IV-V afin d'élucider la contribution structurale des résidus composant ces domaines. Une fois ces structures obtenues, il serait possible de générer un modèle à haute résolution du ribozyme VS complet qui permettrait de mieux comprendre les mécanismes de reconnaissance du substrat et de la formation du site actif du ribozyme VS.

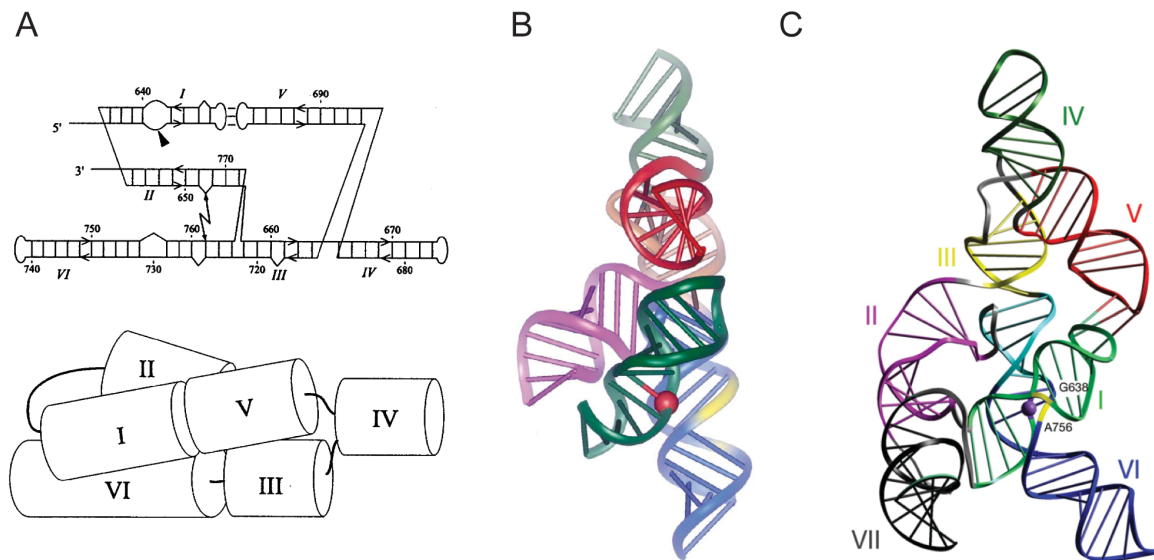


Figure 1.23. Modèles à basse résolution du ribozyme VS de *Neurospora*. (A) Modèle biochimique, (B) modèle de FRET et (C) modèle de SAXS. En (B) et (C), le site de clivage est représenté en rouge et en bleu, respectivement, tandis que la boucle A730 est représentée en jaune. Tiré de [219, 224, 259, 260].

1.6. La spectroscopie de résonance magnétique nucléaire

Deux techniques principales sont utilisées pour déterminer une structure à haute résolution de molécules d'ARN : la cristallographie par diffraction de rayons X et la spectroscopie de RMN. En date du mois de janvier 2015, environ 55% des structures d'ARN déposées sur RCSB, qui est la base de données de structures de macromolécules (www.rcsb.org), ont été obtenues par cristallographie et 45% par RMN, ce qui indique que la RMN est une technique essentielle pour la détermination de structure d'ARN. L'avantage de la RMN est que les expériences se déroulent en solution et qu'en plus de permettre d'obtenir une structure à haute résolution, la RMN permet de faire la caractérisation des systèmes étudiés comme par exemple en déterminant le dynamisme, le pK_a et les sites de liaison de métaux.

La RMN se base sur le magnétisme inhérent à chaque noyau atomique, que l'on appelle le *spin* nucléaire [261]. Lorsqu'un atome se trouve dans un champ magnétique stable, son *spin* interagit avec le champ magnétique, ce qui permet de lever la dégénérescence des niveaux d'énergie du *spin* nucléaire, ne laissant que deux états énergétiques possibles. C'est ce qu'on appelle l'effet Zeeman nucléaire. Le *spin* peut donc s'aligner avec ou contre le champ magnétique et l'état du *spin* le plus stable est légèrement plus peuplé que l'autre, en accord avec la loi de Boltzmann, ce qui crée un *spin* macroscopique non nul. Par l'utilisation d'impulsions électromagnétiques, il est possible de stimuler les noyaux à leur fréquence de résonance pour que leur état de *spin* passe d'un état de basse énergie à un état de haute énergie. Lors du retour à l'équilibre des populations de *spin*, la fréquence de résonance des noyaux peut être détectée. Cette fréquence est mesurée et donne lieu à des signaux, ou résonances, qui sont observés dans les spectres RMN [261]. La fréquence d'un noyau est hautement dépendante de son environnement chimique, ce qui est informatif sur la structure de la molécule étudiée [261]. Lorsque la fréquence (Hz) est convertie en partie par million (ppm) on parle de déplacement chimique (δ , *chemical shift*; [261]). L'utilité des mesures en ppm est qu'il s'agit d'une unité de mesure qui est indépendante des champs magnétiques, permettant la comparaison entre des spectres collectés sur des spectromètres différents. Les atomes les plus utilisés en RMN de l'ARN sont les ^1H , ^{13}C , ^{15}N et ^{31}P . Comme les atomes ^{13}C et ^{15}N ont une abondance naturelle peu élevée, leur utilisation nécessite la production d'ARN enrichis par ces isotopes, permettant de faire de la RMN hétéronucléaire.

La collecte de spectres RMN se fait à l'aide d'échantillons d'ARN concentrés (généralement de 0.5 à 2 mM), marqués en ^{15}N ou $^{13}\text{C}/^{15}\text{N}$. Plusieurs types de spectres peuvent être collectés et procurent des informations complémentaires sur la structure des molécules en solution. Il faut donc généralement un ensemble de spectres diversifiés pour permettre l'attribution des résonances (^1H , ^{13}C , ^{15}N et ^{31}P) d'une molécule et l'obtention de sa structure tridimensionnelle. Les spectres 1D du ^1H permettent généralement d'observer tous les protons d'une molécule [261]. Ces spectres donnent une idée du repliement d'un ARN ou encore des modifications observables suite au changement des conditions expérimentales. Chez l'ARN, les protons imino sont retrouvés entre 9 et 15 ppm, les protons amino entre 6 et 9 ppm et les protons non-échangeables entre 4 et 9 ppm [262]. D'ordre général, les expériences 1D du ^1H n'offrent pas une bonne résolution de signaux puisque ceux-ci sont empilés les uns sur les autres. Les spectres 2D hétéronucléaires offrent quant à eux une bien meilleure résolution puisque les signaux sont dispersés sur deux dimensions [261]. L'utilisation d'échantillons marqués permet d'ajouter une dimension ^{15}N (pour les protons imino et amino) ou ^{13}C (pour les protons non-échangeables). Ils impliquent un transfert de magnétisation entre le proton et l'hétéroatome. Les spectres 2D hétéronucléaires les plus couramment utilisés sont les HSQC (*heteronuclear single quantum coherence*; [263, 264]) et les HMQC (*heteronuclear multiple quantum correlation*; [265]). Ils permettent de relier le ^1H à l'hétéroatome auquel il est attaché, facilitant l'attribution des signaux puisque les déplacements chimiques des hétéroatomes sont caractéristiques selon le groupement chimique (par exemple, le groupement N1 de G se distingue du groupement N3 de U). Finalement, les expériences 3D ajoutent une 3^e dimension qui permet d'obtenir des informations additionnelles. Ces spectres sont directement utilisés pour l'attribution des résonances des molécules d'ARN. Deux types d'expériences 3D différentes sont couramment utilisées : les expériences 3D de corrélation et les expériences 3D de type NOESY. Les expériences de corrélation se basent sur le transfert de magnétisation entre les atomes d'un même système de *spin*, permettant ainsi d'attribuer les signaux provenant du même système [261]. Dans l'ARN, ces expériences sont particulièrement utiles pour l'attribution des protons du ribose (H1', H2', H3', H4' et H5'/H5'') et l'attribution des protons H5 et H6 des pyrimidines. Ces expériences incluent entre autre les 3D HCCH-COSY (*correlation spectroscopy*; [266]) et 3D HCCH-TOCSY (*total correlation spectroscopy*; [266]). Les expériences de type NOESY (*nuclear*

overhauser effect spectroscopy) se basent sur le transfert de magnétisation entre des protons qui ne sont pas reliés directement par des liens chimiques mais qui sont à proximité dans l'espace (généralement moins de 5 Å; [267]). Les signaux observés, nommés NOE, ont une intensité inversement proportionnelle à la distance les séparant, à la puissance 6 ($1/r^6$ où r est la distance). Les expériences de type NOESY sont utilisées directement pour l'attribution des signaux d'une molécule mais aussi pour l'obtention de contraintes structurales qui seront utilisées dans le calcul de structure [265, 268, 269].

Généralement, les premiers signaux à être attribués sont les protons échangeables, particulièrement les protons imino de G et U. Ils sont révélateurs du repliement global d'un ARN puisque dans les conditions expérimentales généralement utilisées, les protons imino doivent former des ponts hydrogènes pour être stabilisés et donc observés dans les spectres RMN [270]. Les protons imino des G et des U sont donc informatifs de la formation de paire de bases canoniques et non-canoniques ou de certains autres types d'interactions. Ainsi, les paires de bases canoniques sont confirmées par des NOE de type G NH – C NH₂ et U NH – A H₂ tandis que d'autres types de NOE permettent de confirmer la formation de paires de bases non-canoniques. De plus, comme des NOE sont observables entre les protons imino de paires de bases adjacentes, la formation de domaines hélicaux peut être rapidement confirmée par RMN [270].

L'attribution des protons non échangeables est habituellement la partie la plus complexe de l'attribution d'une molécule d'ARN. Ces protons incluent les protons aromatiques (H2 de A, H8 de purines et H5/H6 de pyrimidines) et les protons du ribose (H1', H2', H3', H4' et H5'/H5''); [270]). L'approche classique pour l'attribution des protons non échangeables est la marche séquentielle où un H1'_(n) offre un NOE au proton H6/H8_(n) de son propre résidu et au proton H6/H8_(n+1) du résidu consécutif en 3'. Cette approche permet d'attribuer les protons H1' et H6/H8 de plusieurs résidus consécutifs assez rapidement. La spectroscopie de corrélation (3D HCCH-COSY et HCCH-TOCSY) est aussi très utile dans ces cas, puisqu'à partir de l'attribution du H1' d'un résidu, l'ensemble du système de *spin* (C1'-H1', C2'-H2', C3'-H3', C4'-H4' et C5'-H5'/H5'') du ribose peut être attribué.

Le but ultime de l'attribution des résonances d'une molécule d'ARN est de pouvoir extraire un maximum de contraintes structurales à partir des signaux NOE. Puisque l'intensité d'un NOE est reliée à la distance séparant deux protons, les signaux NOE sont

classifiées en différentes catégories basées selon leur intensité [270]. Des intervalles de distances leur sont ensuite associées qui permettent de définir les contraintes structurales. Ces contraintes forment généralement la majorité des contraintes RMN utilisées dans un calcul de structure. Cependant, des contraintes supplémentaires sont souvent ajoutées qui permettent de définir les paires de bases formées, les conformations des riboses et les angles glycosidiques. L'utilisation de RDC (*residual dipolar couplings*) permet aussi d'ajouter des contraintes à longue distance qui sont utiles pour le raffinement [271].

Les contraintes structurales sont ensuite utilisées dans un calcul de structure. Les calculs de structure d'ARN utilisent principalement l'approche de *simulated annealing* (SA) qui permet de replier un ARN à partir d'un état non-replié sur lequel s'applique ensuite l'ensemble des contraintes structurales expérimentales dans un champ de force donné [272]. Au début de la simulation, une énergie est attribuée à la molécule dans un contexte de température élevée. Durant la simulation, la température diminue (c'est l'étape d'*annealing*) et le système tente de satisfaire toutes les contraintes structurales afin de faire diminuer son énergie au minimum. Ainsi, plus le nombre de contraintes est élevé, plus la structure obtenue est bien définie. Un ensemble des structures de plus basses énergies est ensuite utilisé pour représenter la structure d'une molécule.

L'étude de molécules d'ARN de large taille (plus de 10-15 kDa ou 30-45 nucléotides) par spectroscopie de RMN peut être difficile pour plusieurs raisons [262]. Tout d'abord, un des éléments essentiels pour déterminer la structure d'un ARN est l'attribution des signaux. Cette étape peut être difficile si les signaux sont superposés dû à un nombre élevé de résonances observées. De ce fait, le manque de résolution des signaux peut complexifier l'obtention de contraintes structurales et diminuer la définition de la structure finale. Aussi, plus la taille d'une molécule augmente, plus les largeurs de raie sont importantes dues à une relaxation rapide des résonances [273]. Ceci réduit l'intensité des signaux observés et diminue leur résolution, compliquant donc l'attribution des spectres et l'obtention de contraintes structurales. Afin de faciliter les études RMN des ARN de plus grandes tailles, certaines approches sont possibles. Tout d'abord, l'utilisation de stratégies de marquage isotopique sélectif permet d'incorporer des isotopes spécifiques (^2H , ^{13}C ou ^{15}N) qui permet de réduire le nombre de signaux observés dans un spectre et du même coup, de réduire la superposition de signaux [274-276]. Il est aussi possible de faire du marquage isotopique segmentaire en ciblant un segment spécifique d'ARN, permettant de faciliter

l'attribution d'une molécule [277]. Aussi, d'autres études utilisent une approche de *divide-and-conquer* où les sous-domaines importants d'un ARN sont étudiés séparément et de façon isolée. Ces domaines isolés sont plus facilement étudiables et permettent de comprendre la structure et la contribution de ces domaines dans le cadre de la molécule entière. L'approche de *divide-and-conquer* a été largement utilisée pour l'étude de la structure des protéines et de l'ARN par RMN [278-281].

En plus de la détermination de structures d'ARN à haute résolution, la RMN permet aussi de caractériser plusieurs aspects d'une molécule d'ARN [270]. Comme les molécules d'ARN peuvent faire des mouvements moléculaires importants, la RMN peut caractériser la dynamique de certains systèmes qui se produisent sur différentes échelles de temps (10^{-12} à 10^3 secondes; [282]). De plus, l'utilisation de la RMN peut mener à la détermination des paramètres thermodynamiques reliés à l'interaction entre des molécules, tel que la constante de dissociation (K_d ; [270]). Il est aussi possible de déterminer la valeur du pK_a de certains résidus critiques pour l'adoption d'une structure ou qui participent à un mécanisme enzymatique [283]. Aussi, comme le rôle des ions divalents est particulièrement important dans l'adoption de la structure tridimensionnelle des ARN, la spectroscopie de RMN peut être utilisée pour localiser les sites de liaison de métaux [270].

Dans cette optique, plusieurs approches RMN sont disponibles pour localiser les ions divalents qui possèdent une géométrie similaire à celle des ions Mg^{2+} dans une structure d'ARN. Parmi celles-ci, l'utilisation de cobalt hexamine ($Co(NH_3)_6^{3+}$) permet de détecter des signaux NOE entre les protons du $Co(NH_3)_6^{3+}$ et les nucléotides qui forment les sites de liaison de métaux mais ne permet pas d'identifier les sites de liaison *inner-sphere* [284]. Une autre approche utilise le Mn^{2+} qui partage plusieurs propriétés avec le Mg^{2+} : ils possèdent une géométrie octaédrale avec six ligands et le Mn^{2+} se lie à des sites de liaison similaires. Cependant, le Mn^{2+} est un ion paramagnétique, ce qui lui confère un avantage spectroscopique en permettant de définir les sites de liaison de métaux en augmentant la relaxation des résonances se trouvant à proximité de cet ion [231, 285]. D'autres approches peuvent être utilisées mais requièrent l'utilisation d'ARN modifiés. Une de celles-ci consiste à utiliser des nucléotides modifiés incorporant du fluore 19 (^{19}F) qui permettent de localiser les sites de liaison de métaux en se basant sur les modifications de déplacement chimique du ^{19}F qui est très sensible à son environnement chimique [286]. Finalement la perturbation de déplacement chimique induite par le Cd^{2+} sur des ARN contenant des

modifications phosphorothioates peut être utilisée afin d'identifier les sites de liaison de métaux d'un ARN. L'utilisation du Cd^{2+} , qui est plus thiophilique que le Mg^{2+} , ainsi que des modifications phosphorothioates permettent d'évaluer la contribution des groupements phosphates et le mode de coordination (*outer* et *inner-sphere*) des métaux à ces sites [287, 288]. Dans tous les cas, ces techniques permettent d'identifier les ligands potentiels formant les sites de liaison de métaux. Cependant, malgré que le rôle des ions Mg^{2+} dans l'adoption des structures tridimensionnelles d'ARN soit important, la localisation des ions divalents dans les structures RMN de molécules d'ARN n'est pas systématique, limitant la compréhension des interactions ARN-métal en solution.

1.7. Objectifs de la thèse

L'obtention de la structure complète d'un ribozyme permet d'expliquer les données biochimiques propres à ce ribozyme, de comprendre les principes du repliement et de la stabilisation de cet ARN, en plus d'élucider les détails associés à la formation du site actif et au mécanisme enzymatique. Le ribozyme VS est le seul ribozyme de la classe des petits ribozymes nucléolytiques pour lequel une structure complète à haute résolution n'est pas encore publiée.

Afin d'obtenir la structure complète du ribozyme VS, des études de RMN ont été initiées dans le laboratoire. Comme le ribozyme VS fait environ 150 nucléotides, les études de RMN sont basées sur une approche de *divide-and-conquer* où les sous-domaines importants sont étudiés indépendamment du ribozyme complet afin de rendre possible l'analyse des spectres. L'objectif ultime de cette approche est de générer un modèle RMN à haute résolution du ribozyme VS complet lorsque la structure de tous les domaines importants sera disponible. Jusqu'à présent, les structures à haute résolution de SLI, de SLV et de l'interaction *kissing-loop* I/V ont été obtenues. Dans tous ces cas, les structures sont en accord avec les données biochimiques, permettant ainsi d'expliquer le rôle de chacun des résidus importants de ces domaines.

Cependant, aucune structure à haute résolution n'est disponible pour les deux jonctions à trois voies du domaine catalytique (jonctions II-III-VI et III-IV-V). De plus, la structure RMN disponible de la boucle A730 est en désaccord avec l'ensemble des données biochimiques concernant ce domaine, suggérant que cette structure correspond à une conformation inactive de cette boucle. L'obtention de structures à haute résolution pour ces

domaines est importante afin de comprendre la formation du site actif du ribozyme VS ainsi que les détails architecturaux permettant l'adoption de la structure active du domaine catalytique. J'ai donc défini avec mes collaborateurs les objectifs de travail ci-dessous.

1. Au Chapitre 2, nous nous sommes intéressés à la structure de la boucle A730. En effet, la structure RMN disponible de la boucle A730 ne concorde pas avec les données biochimiques et ne permet donc pas d'évaluer la contribution de cette boucle à la formation du site actif du ribozyme VS. Notre hypothèse de départ est que la boucle A730 doit adopter une structure précise en accord avec les données biochimiques et qui permet de positionner le A₇₅₆ pour qu'il puisse participer au mécanisme enzymatique. Nous avons utilisé la spectroscopie de RMN afin de définir un domaine minimal qui représente bien la boucle A730, de déterminer la structure RMN de cette boucle ainsi que le pK_a des adénines qui s'y retrouvent. Cette étude a mené à l'obtention d'un modèle du site actif du ribozyme VS qui est en accord avec les données biochimiques et mécanistiques et qui révèle de nouveaux éléments d'homologie entre les ribozymes *hairpin* et VS. Ainsi, cette étude nous permet de mieux comprendre l'organisation du site actif du ribozyme VS.

2. Au Chapitre 3, notre objectif était de localiser les ions Mg²⁺ dans la structure de la boucle A730. Comme la boucle A730 requiert la présence d'ions Mg²⁺ pour se replier de façon stable, nous avons émis l'hypothèse que la boucle A730 forme un ou plusieurs site(s) de liaison de métaux divalents. Afin de tester cette hypothèse, nous avons mis au point une nouvelle stratégie RMN pour la localisation de métaux divalents qui combine la relaxation paramagnétique induite par le Mn²⁺ ainsi que la perturbation de déplacement chimique induite par le Cd²⁺ sur des ARN contenant des modifications phosphorothioates. Cette étude a permis d'identifier que des ions Mg²⁺ interagissent et stabilisent la boucle A730 et qu'un ion Mg²⁺ pourrait participer à l'activité enzymatique en étant coordonné au phosphate du A₇₅₆. Cette étude a donc permis de mieux comprendre le rôle des ions Mg²⁺ dans le site actif du ribozyme VS tout en approfondissant notre compréhension des interactions ARN-métal.

3. Au Chapitre 4, notre objectif était de déterminer la structure de la jonction III-IV-V. La jonction III-IV-V est un domaine architectural important pour la formation de l'interaction *kissing-loop IV* et du site actif. L'absence de structure tridimensionnelle à haute résolution de ce domaine limite la compréhension de la contribution de ce domaine à l'architecture du domaine catalytique du ribozyme VS. Nous avons émis l'hypothèse que la jonction III-IV-V adopte sous une forme isolée une structure compacte et stable qui est représentative de sa

structure dans le ribozyme VS complet. Par spectroscopie de RMN, nous avons déterminé la structure de la jonction III-IV-V en plus de localiser les ions Mg^{2+} par l'approche de relaxation paramagnétique induite par le Mn^{2+} . Cette étude a révélé que cette jonction adopte une structure compacte qui est stabilisée par de nombreux éléments structuraux. La jonction III-IV-V forme aussi une charnière dynamique qui facilite la formation de l'interaction *kissing-loop* I/V ainsi que la catalyse. Cette étude contribue donc à la compréhension de l'organisation du domaine catalytique et à la formation des interactions tertiaires importantes pour l'activité du ribozyme VS.

4. Finalement, au Chapitre 5 notre objectif était de déterminer et de caractériser la structure de la jonction II-III-VI. La jonction II-III-VI est un domaine essentiel et, similairement à la jonction III-IV-V, l'obtention de sa structure est nécessaire pour mieux comprendre l'architecture du ribozyme VS. Ainsi, nous avons tout d'abord émis l'hypothèse que la jonction II-III-VI adopte sous une forme isolée une structure stable qui est représentative de sa structure dans le cadre du ribozyme VS complet. Une autre hypothèse de travail est que les renflements des hélices II, III et VI peuvent être importants pour le repliement de la jonction II-III-VI. Nous avons premièrement évalué par spectroscopie RMN la dépendance aux ions Mg^{2+} de la jonction II-III-VI en plus de faire des études de mutagenèse sur le renflement de l'hélice II pour déterminer sa contribution à l'adoption de la structure de la jonction. Nous avons ensuite déterminé la structure RMN de la jonction II-III-VI en plus de localiser les ions Mg^{2+} par l'approche de relaxation paramagnétique induite par le Mn^{2+} . La structure de la jonction II-III-VI a révélé la formation d'une interaction tertiaire entre les renflements des tiges II et VI qui est critique pour l'adoption de la structure tridimensionnelle de cette jonction. Cette étude contribue donc à la compréhension de la structure du domaine catalytique ainsi qu'à la formation des interactions tertiaires importantes à l'activité du ribozyme VS.

Les données structurales obtenues sur les domaines de la boucle A730 et sur les jonctions II-III-VI et III-IV-V apportent de nouvelles données essentielles à la compréhension de la reconnaissance du substrat et de la formation du site actif. Entre autre, ces données permettront de bâtir un modèle RMN à haute résolution du ribozyme VS et faciliteront les études d'ingénierie rationnelle du ribozyme VS. Ces études s'inscrivent également dans une contribution générale à la compréhension du repliement de l'ARN ainsi qu'au développement d'approches RMN pour la localisation des ions Mg^{2+} .

**2. Détermination de la structure RMN de la boucle
A730 et modélisation du site actif du ribozyme VS de
*Neurospora***

Mise en contexte

La formation du site actif du ribozyme VS requiert l'association de deux boucles internes, soit la boucle interne de l'hélice I et la boucle A730 de l'hélice VI. Ces deux boucles internes doivent positionner les nucléotides catalytiques dans une position permettant au mécanisme enzymatique de prendre place. Précédemment, la structure RMN de la boucle interne de l'hélice I sous sa forme active a été obtenue, révélant que le G₆₃₈ est à proximité du groupement 2'-OH du G₆₂₀, en accord avec son rôle de base générale dans le mécanisme enzymatique proposé du ribozyme VS. De plus, la structure de la boucle A730 a été obtenue précédemment par RMN. Cette structure a révélé un changement conformationnel dans l'hélice VI, montrant que la boucle A730 est formée de trois paires de bases non-canoniques, en désaccord avec la structure secondaire proposée. Cette structure ne permet donc pas d'expliquer les données biochimiques et de plus, le A₇₅₆ n'est pas disponible pour jouer son rôle d'acide général dans le mécanisme enzymatique [258]. Ainsi, nous avons émis l'hypothèse que la boucle A730 doit adopter une structure précise lui permettrait de positionner le résidu A₇₅₆ de sorte qu'il puisse participer au mécanisme enzymatique.

Dans cette première étude, nous avons déterminé la structure de la boucle A730 par spectroscopie RMN. Nous avons démontré que l'ARN SLVI, formé de 26 nucléotides, requiert la présence d'ions Mg²⁺ pour se replier de façon stable et compacte. En plus d'être complètement en accord avec les données biochimiques, la structure RMN de la boucle A730 révèle la formation de deux motifs structuraux qui n'étaient pas prédits par les données biochimiques. Premièrement, les résidus A₇₃₀ et G₇₅₇ forment une paire de bases *cis* WC/WC G-A qui permet de stabiliser le repliement de la boucle interne. De plus, plusieurs résidus sont impliqués dans la formation d'un motif *S-turn*. Le repliement de la boucle A730 mène à l'exclusion de deux résidus dans un sillon mineur élargi. Ces résidus sont le C₇₅₅ et le A₇₅₆, signifiant que le repliement de la boucle A730 permet de rendre le A₇₅₆ disponible pour participer au mécanisme enzymatique. De plus, la détermination du pK_a des adénines de la boucle a démontré que le pK_a des résidus A₇₃₀ et A₇₅₆ sont altérés, ce qui concorde avec le rôle catalytique du A₇₅₆. Finalement, la modélisation du site actif du ribozyme VS a pu être réalisée en utilisant les structures RMN de la boucle interne de SLI sous sa conformation active et la boucle A730 [253]. La structure RMN de la boucle A730

a donc contribué à une meilleure compréhension de l'architecture du site actif du ribozyme VS.

Dans cette étude, la préparation des échantillons a été réalisée par Geneviève Desjardins, alors étudiante au M.Sc en biochimie. Elle a aussi également contribué avec le Dr. Pascale Legault à la collecte des données RMN. Conjointement avec Geneviève Desjardins et le Dr. Pascale Legault, j'ai contribué à l'analyse des données RMN ainsi qu'au calcul et à l'analyse des structures. Jérôme Boisbouvier a contribué à la collecte des données de RDC et Nicolas Girard a réalisé la modélisation du site actif du ribozyme VS conjointement avec le Dr. Pascale Legault. Finalement, la rédaction de l'article a été effectuée par le Dr. Pascale Legault.

NMR Structure of the A730 loop of the *Neurospora* VS Ribozyme: Insights into the Formation of the Active Site

*Geneviève Desjardins^{1,5}, Eric Bonneau^{1,5}, Nicolas Girard¹, Jérôme Boisbouvier^{2,3,4}, and
Pascale Legault^{1,*}*

¹Département de Biochimie, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, Québec, H3C 3J7, Canada.

²Commissariat à l'Énergie Atomique, Institut de Biologie Structurale Jean-Pierre Ebel, 38027 Grenoble Cedex 1, France.

³Centre National de la Recherche Scientifique, Institut de Biologie Structurale Jean-Pierre Ebel, Grenoble, France.

⁴Université Joseph-Fourier, Institut de Biologie Structurale Jean-Pierre Ebel, Grenoble, France.

⁵The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

This work was originally published in Nucleic Acids Research:

Nucleic Acids Res. 2011, May;39(10):4427-37. doi: 10.1093/nar/gkq1244. Epub 2011 Jan 25.

ABSTRACT

The *Neurospora* VS ribozyme is a small nucleolytic ribozyme with unique primary, secondary and global tertiary structures, but which displays mechanistic similarities to the hairpin ribozyme. Here, we determined the high-resolution NMR structure of a stem-loop VI fragment containing the A730 internal loop, which forms part of the active site. In the presence of magnesium ions, the A730 loop adopts a structure that is consistent with existing biochemical data and most likely reflects its conformation in the VS ribozyme prior to docking with the cleavage site internal loop. Interestingly, the A730 loop adopts an S-turn motif that is also present in loop B within the hairpin ribozyme active site. The S-turn appears necessary to expose the Watson-Crick edge of a catalytically important residue (A756) so that it can fulfill its role in catalysis. The A730 loop and the cleavage site loop of the VS ribozyme display structural similarities to internal loops found in the active site of the hairpin ribozyme. These similarities provided a rationale to build a model of the VS ribozyme active site based on the crystal structure of the hairpin ribozyme.

Keywords: *Neurospora* VS ribozyme; A730 loop; active site; NMR structure; S-turn motif; hairpin ribozyme

INTRODUCTION

The *Neurospora* Varkud Satellite (VS) ribozyme is part of a family of small nucleolytic RNA enzymes identified from biological sources that also includes the hammerhead, hairpin, hepatitis delta virus (HDV) and *glmS* ribozymes (for recent reviews see [1-4]). With exception of the *glmS* ribozyme, these ribozymes control the replication cycle of their parental genome via self-cleavage and the reverse ligation reaction. They all catalyze scission of the RNA backbone by a transesterification reaction, which involves the initial attack of a specific 2'-oxygen on the adjacent 3'-phosphorus and yields two products with 2'-3' cyclic phosphate and 5'-OH termini. Their common chemical mechanism is equivalent to that of RNase A, and, although they use different tertiary architectures and chemical groups, several share basic mechanistic strategies: activation of the nucleophile, stabilization of an in-line geometry between reactive groups and leaving group stabilization [3]. Furthermore, although ribozymes were initially viewed as strict metalloenzymes [5], there is now evidence that several ribozymes employ an acid-base mechanism for the cleavage reaction, and use nucleobases as proton donors and/or acceptors [3].

The *Neurospora* VS ribozyme was first identified as a satellite RNA in the Varkud-1c strain and a few other isolates of *Neurospora* [6]. A contiguous sequence of 154 nucleotides was shown to be minimally required for self-cleavage *in vitro* [7]. Its secondary structure contains six helical domains: stem-loop I (SLI) forms the substrate and stem-loops II-VI (SLII-SLVI) define the catalytic domain (Figure 2.1a; [8]). The VS ribozyme is most active in the presence of divalent ions [9, 10]. Mg^{2+} ions allow tertiary structure formation [8, 11, 12] and partly contribute to the chemistry of the cleavage reaction [10]. Monovalent cations also support cleavage, albeit at a lower rate [10]. The self-cleavage reaction takes place when the SLI substrate is located either at the 5'-end or 3'-end of the catalytic domain [7, 13], and, alternatively, *trans* cleavage occurs when the substrate is synthesized separately from the catalytic domain [14]. Substrate recognition involves formation of a kissing-loop interaction between loop I of the substrate and loop V of the catalytic domain [11]. Formation of this I/V kissing-loop interaction is accompanied by a conformational change in SLI from an unshifted inactive conformation to a shifted active conformation [15]. Our present understanding is that, in order for cleavage to occur, the cleavage site

internal loop of SLI must dock in a cleft formed by SLII and SLVI to allow its interaction with the A730 loop of SLVI [4, 12, 16-25].

To date, the VS ribozyme is the only nucleolytic ribozyme for which there is no high-resolution structure of the complete RNA. High-resolution NMR structures of stem-loops I and V derived from the VS ribozyme have been obtained [26-31] and provide insights about substrate recognition and activation. The NMR structure of an SLVI fragment containing A730 loop residues was also previously determined, but this structure was found incompatible with previous biochemical data and thus, likely represents an inactive conformation of the A730 loop [29]. In addition, low-resolution models have been derived from biochemical data [12], fluorescence resonance energy transfer [32, 33] and small-angle X-ray scattering in solution [34]. These models help define a global architecture for the VS ribozyme that is unique among nucleolytic ribozymes, but provide no detailed insights into the active site.

Despite the lack of high-resolution structural information, several mechanistic insights have been uncovered for the VS ribozyme [10, 21-25] that reveal striking similarities with the hairpin ribozyme (Figure 2.1b; [4]). Although the VS and hairpin ribozymes adopt significantly different secondary and tertiary structures, in both cases formation of the active site results from the association of two internal loops (Figure 2.1b) and similar general acid-base cleavage mechanisms have been proposed [4]. Both ribozymes cleave at a purine-purine step within a purine-rich internal loop, between G620-A621 in the VS ribozyme and A-1 and G+1 in the hairpin ribozyme (Figure 2.1b). The cleavage site loops both contain a catalytically important residue proposed to act as the general base, G638 in the SLI internal loop of the VS ribozyme [22, 24, 25] and G8 in loop A of the hairpin ribozyme [4]. Interestingly, additional sequence similarities at the cleavage site internal loops have been reported [35], including a common GCR (R = A or G) sequence just downstream of the cleavage site and a common CGA(A)GCGG sequence [the (A) indicate that an A is found only in the hairpin ribozyme] on the opposite strand that comprises the proposed general base (Figure 2.1a). The other internal loops forming the active site contain a catalytically important residue proposed to act as the general acid, A756 in the A730 loop of the VS ribozyme [21, 23, 25] and A38 in loop B of the hairpin ribozyme [4]. Association of these two internal loops likely occurs with similar overall topology in both systems [4]. In summary, there are considerable similarities between the VS and hairpin

ribozymes, which suggest comparable active site architectures to achieve cleavage chemistry.

In order to gain insights into the formation of the VS ribozyme active site, we determined the high-resolution NMR structure of an SLVI fragment containing the A730 loop domain (Figure 2.1c). The three-dimensional structure of this A730 internal loop is consistent with a large number of biochemical data and likely represents a ground-state structure prior to its docking with the cleavage site loop. Based on this NMR structure and the previously determined NMR structure of the active conformation of the SLI internal loop, we identify additional similarities between the hairpin and VS ribozymes. We also discuss the structural changes that would be necessary in the SLI and SLVI internal loops for formation of a VS ribozyme active site that mimics the hairpin ribozyme.

MATERIALS AND METHODS

RNA sample preparation. Unlabeled, ^{15}N - and $^{13}\text{C}/^{15}\text{N}$ -labeled SLVI RNAs were synthesized *in vitro* with T7 RNA polymerase and purified as described previously [30]. The purified RNAs were first exchanged in a low salt buffer (10 mM sodium cacodylate pH 6.5, 50 mM KCl, 0.05 mM NaN_3 in 90% $\text{H}_2\text{O}/10\%$ D_2O) with Centricon-3 ultracentrifugation devices (Millipore, MA). They were heated to 95°C for 2 min and then cooled in iced water for 5 min before transferring to the standard NMR buffer [low salt buffer containing 5 mM MgCl_2 99.995% (Sigma-Aldrich, MO)] or other MgCl_2 -supplemented buffers (low salt buffer containing either 2.5 mM or 10 mM MgCl_2 99.995%) using ultracentrifugation devices. For NMR studies in D_2O , the samples were lyophilized several times and resuspended in 99.996% D_2O . The concentration of SLVI RNA for NMR studies ranged between 0.4 and 1.5 mM.

Determination of adenine pK_a 's in the SLVI fragment. All pH values were measured in D_2O with a glass electrode (Radiometer America Inc., OH) placed directly in the NMR sample. No correction was made for the D_2O present in the sample. The pH was adjusted with 0.01-0.1 M DCl and NaOD. The reported pH values are the average of the pH values measured before and after the NMR experiments. At each different pH value, 1D ^1H and 2D ^1H - ^{13}C HMQC spectra were collected at 25°C . The 1D ^1H spectra were used to verify

the integrity of the sample and the 2D ^1H - ^{13}C HMQC spectra were analyzed to determine the values of adenine C2 chemical shifts. In order to determine adenine pK_a of A8 and A20, the log of $[(\Delta_T - \Delta)/\Delta]$ was plotted as a function of pH, and the pK_a values were derived by linear regression based on the Henderson-Hasselbalch-type equation $\text{pH} = \log[(\Delta_T - \Delta)/\Delta] + \text{pK}_a$ [36], where Δ is the change in C2 chemical shift at a given pH relative to the unprotonated state (pH 8.6) and Δ_T is the total change in chemical shift between the protonated and unprotonated states. Given that full protonation of the A8 (A756) and A20 (A730) residues could not be observed within the pH range studied (pH 4.7-8.6), the values of Δ_T were therefore estimated between 6 and 8 ppm in agreement with reported values for folded RNAs [27, 29, 36-39]. Values of pK_a derived with $\Delta_T = 6, 7$ and 8 ppm were used to calculate an average pK_a value and the error was derived from the maximum difference with the mean.

NMR spectroscopy. All NMR data were acquired with 500 MHz, 600 MHz, and 800 MHz Varian INOVA spectrometers equipped with a pulsed field gradient unit and an actively shielded z gradient probe, either a $^1\text{H}\{^{13}\text{C}/^{15}\text{N}\}$ triple-resonance probe (standard or cold probe) or a $^1\text{H}\{^{15}\text{N}-^{31}\text{P}\}$ indirect detection probe. Exchangeable protons and their attached nitrogens were assigned at 15°C , whereas non-exchangeable protons and their attached carbons were obtained at 25°C in D_2O , as previously described [30]. In addition, 2D ^1H - ^{15}N HMQC optimized for transfers via $J=7.0$ Hz and $J=21$ Hz [40] were collected at 25°C in D_2O for assignment of uridine N1, pyrimidine N3, as well as N7 and N9 of purines. ^1H , ^{13}C and ^{15}N chemical shifts were referenced at 25°C to an external standard of 2,2-dimethyl-2-silapentane-5-sulfonic acid at 0.00 ppm. NMR data were processed using the NMRPipe/NMRDraw package [41] and analyzed with NMRView [42].

Structural restraints. An HNN-COSY spectrum was collected in H_2O at 15°C to detect $^2J_{\text{NN}}$ coupling across Watson-Crick base pairs [43]. Distance restraints for exchangeable protons were extracted from 2D flip-back watergate NOESY (mixing time (τ_m) = 160 ms; [44, 45]), 3D amino-optimized ^{15}N -edited NOESY-HSQC (τ_m of 80 and 160 ms; [46]) and 2D ^1H - ^{15}N CPMG-NOESY (τ_m of 80 and 160 ms; [47]) spectra collected at 15°C in H_2O . Distance restraints for non-exchangeable protons were extracted from 3D ^{13}C -edited HMQC-NOESY spectra (τ_m of 80 and 160 ms; [48]) collected at 25°C in D_2O . The NOE-

derived distance restraints were separated in 4 classes: strong (1.8-3.3 Å), medium (1.8-4.2 Å), weak (1.8-5.5 Å) and very weak (2.8-7.5 Å) based on NOE crosspeak intensities. Based on NMR evidence for specific base pairing obtained from NOESY and 2D HNN-COSY spectra, canonical distance restraints were employed to define the hydrogen-bond pattern and planarity of the following base pairs: G1-C26, A2-U25, G3-C24, C4-G23, U5-G22, G6-C21, C10-G19, A11-U18, C12-G17, and G13-A16. For the G9-A20 base pair, only two hydrogen-bond distance restraints were employed (G9 H1 - A20 N1 and G9 O6 - A20 H6). Dihedral angle restraints for the sugar puckers (δ) were obtained from analysis of 2D DQF-COSY [49] and 3D HCCH-E.COSY [50] spectra. The backbone dihedral angles α , γ , χ , and ζ were restrained using distance restraints derived from comparative NOE analyses [51]. Based on NMR evidence, backbone torsion angles of residues in helical regions (1-5, 11-12, 17-18, 22-26) were restrained to A-form values ($\pm 15^\circ$).

Liquid crystal NMR spectroscopy. For liquid crystal NMR studies, ^{15}N -labeled and $^{13}\text{C}/^{15}\text{N}$ -labeled SLVI RNAs (at 0.4-0.5 mM) were aligned by adding a concentrated Pf1 filamentous phage solution (ASLA biotech.) at a final concentration of ~ 17 mg/mL [52]. Prior to addition of the Pf1 phages, these were exchanged at least two times with the NMR buffer containing 5 mM MgCl_2 . The ^2H NMR splitting for D_2O observed at 25°C in these samples was 15.53 Hz for the ^{15}N -labeled SLVI and 15.24 Hz for the $^{13}\text{C}/^{15}\text{N}$ -labeled SLVI. 2D ^1H - ^{15}N HSQC spectra collected with active J_{HN} coupling in t_1 were used to extract imino ^1H - ^{15}N couplings with the isotropic and Pf1-aligned ^{15}N -labeled SLVI samples. Spin-state selective experiments [53, 54] were used to extract one-bond ^1H - ^{13}C couplings (C1'-H1', C2-H2, C5-H5, C6-H6 and C8-H8) in isotropic and Pf1-aligned $^{13}\text{C}/^{15}\text{N}$ -labeled SLVI samples. One-bond ^1H - ^{15}N and ^1H - ^{13}C couplings were measured from fitted peaks using NMRWish in the NMRPipe/NMRDraw package. Extracted residual dipolar coupling (RDC) values were scaled linearly with respect to the observed D_2O splitting, to take into account the small difference in the magnitude of the alignment from the different samples. The RDC restraints were separated in two classes to account for overlap of the peaks in the spectra, with error bars estimated at 1 Hz (well resolved) and 2 Hz (partial overlap). RDC restraints for A730 loop residues 6-9 and 20-21 were not included in the structure calculation because of evidence for local dynamics in the internal loop.

Structure calculation. Three-dimensional structures were calculated with restrained molecular dynamics and simulated annealing in X-PLOR-NIH version 2.1.9 [55] using a two-stage protocol. At stage one, an initial set of 50 structures was calculated from structures with randomized backbone angles, as previously described [31]. Structures obtained at stage one satisfy all distance and dihedral experimental restraints (no distance violation of $> 0.2 \text{ \AA}$ and no torsion angle violation of $> 5^\circ$). At stage two, these structures were refined with the same set of restraints supplemented by RDC restraints and using a single alignment tensor [56]. The magnitude of the axial and rhombic components of the alignment tensor were initially obtained using MODULE 1.0 [57], but were subsequently refined by using a grid search to obtain reliable D_a and R values [58]. The grid search procedure was repeated, using RDCs from the 5'-3' stem (residues 2-5 and 22-25) and the hairpin stem (residues 11-18), independently. All grid search calculations converged on similar values of D_a and R [values of $D_a(\text{CH}) = 16.6 \text{ Hz}$ and $R = 0.40$ were used], suggesting that the relative orientation of the two stems is well defined despite evidence of dynamic residues in the A730 loop. For RDC refinement, restrained molecular dynamics and simulated annealing were used followed by a final energy minimization. During simulated annealing, the force constants for RDC, NOE, impropers, angles and Van der Waals restraints were slowly increased. A final set of 500 structures was calculated, from which the 20 lowest-energy structures that satisfied the experimental restraints (no distance violation $> 0.2 \text{ \AA}$, no torsion angle violation $> 5^\circ$ and no RDC violation $> 5 \text{ Hz}$) were selected for analysis. These 20 lowest-energy structures were used to calculate an average structure that was minimized against NOE and dihedral restraints. All structures were visualized with PyMol (Schrödinger) and analyzed with PyMol and Curves+ [59].

Active site modeling. A subset of three residues (G8, A12 and G13) was extracted from the X-ray structure of the hairpin ribozyme determined with vanadate (pdb entry code 1M50; [60]). Residues A12 and G13 in this subset were modified using Coot [61] to match the base type of their equivalents in the VS ribozyme (G620 and A621). The bases of residues G8 and G12 of this modified model were then aligned on the model of the active conformation of SLI of the VS ribozyme [28]. A hybrid SLI model was generated in which residues G620, A621, G638 of SLI were replaced by their equivalents in the modified hairpin model. Using Coot, residue A756 of the SLVI structure was moved by hand to

allow stacking with A621. Finally, the model of the SLI/SLVI complex was energy minimized with Amber [62].

RESULTS

Magnesium-dependent stabilization of the A730 internal loop. To explore the three-dimensional structure of the A730 loop and gain insights into the active site of the VS ribozyme, the 26-nucleotide SLVI RNA shown in Figure 2.1c was investigated by NMR spectroscopy. This SLVI RNA is designed to reproduce the secondary structure context of the A730 internal loop from the VS ribozyme [8], including two sets of two adjacent closing base pairs on each side of this loop (Figure 2.1). A GAAA tetraloop was also included in the SLVI RNA to help stabilize the hairpin conformation and prevent possible duplex formation at the high RNA concentrations used for NMR studies. We initially verified that the SLVI RNA adopts a single hairpin conformation by native gel electrophoresis (not shown), and this was later confirmed from the ensemble of NMR data.

Given that the VS ribozyme is active in the presence of Mg^{2+} ions [9] and that Mg^{2+} -ion binding sites have been identified in the A730 loop domain [63], we first examined the effect of $MgCl_2$ concentration on the folding of the SLVI RNA by 1D imino 1H NMR spectra. The 1D imino 1H NMR spectrum of the SLVI RNA recorded in the absence of Mg^{2+} ions shows signals of variable intensities characteristic of folded RNAs containing dynamic regions (Figure 2.2, bottom spectrum). These imino proton signals were assigned from a 2D NOESY spectrum collected under the same conditions. It was found that all imino protons of the SLVI RNA give an intense signals in this 1D spectrum, except for the G6 and G9 imino protons, which are not observed, and the U18 and G19 imino protons, which yield signals of weak intensities (Figure 2.2, bottom spectrum). These absent and weak imino proton signals indicate that the A730 loop domain is unstable in the absence of Mg^{2+} ions. Upon addition of Mg^{2+} ions, the imino signals of U18 and G19 become significantly more intense (Figure 2.2, top spectra) and a weak imino proton signal from G6 appears in the spectrum. Analysis of 2D NOESY and 2D HNN-COSY spectra collected at 5 mM $MgCl_2$ confirms formation of all base pairs from the predicted secondary structure of the SLVI RNA (Figure 2.1c). Although the imino NMR data do not provide evidence for stabilization of non-canonical base pairs within the A730 loop, they indicate that the A730

loop domain is stabilized by Mg^{2+} ions, in agreement with previous biochemical studies. A MgCl_2 concentration of 5 mM was found sufficient to produce the characteristic spectral changes (Figure 2.2) and, therefore, was selected for future NMR studies.

The overall three-dimensional structure of the SLVI RNA. The structure of the SLVI RNA was obtained by multidimensional heteronuclear NMR methods. Complete resonance assignments were obtained for all observable ^1H , ^{13}C and ^{15}N atoms of SLVI (Supplementary Table 2.1). The structure determination included NOE-derived distance restraints, dihedral angle restraints and RDC restraints (Table 2.1). The SLVI structure is well defined by the NMR data with a heavy atom rmsd of $0.67 \pm 0.17 \text{ \AA}$ for the 20 lowest-energy structures (residues 2-25; Table 2.1 and Figure 2.3a). The SLVI RNA forms a hairpin structure containing a 5'-3' stem (residues 2-5 and 22-25; rmsd of $0.53 \pm 0.07 \text{ \AA}$), a hairpin stem (residues 11-18; rmsd of $0.31 \pm 0.01 \text{ \AA}$), and the A730 loop (residues 6-10 and 19-21; rmsd of $0.65 \pm 0.24 \text{ \AA}$). Stem regions form regular A-form helices and the GAAA tetraloop adopts the typical GNRA fold with its characteristic sheared G-A base pair and 3'-purine stack [64, 65]. The A730 loop domain imparts a $\sim 150^\circ$ interhelical angle between the two stems of SLVI and contains two structural characteristics that were not previously identified: a *cis* WC/WC G9-A20 base pair and an S-turn motif that protrudes C7 and A8 in the minor groove.

Formation of a *cis* WC/WC G-A base pair in the A730 loop. The evidence for a *cis* WC/WC G-A base pair in the A730 loop was obtained from 2D ^1H - ^{15}N CMPG-NOESY and 3D ^{13}C -edited HMQC-NOESY spectra. A strong NOE signal was observed between the A20 H2 and the G9 NH2 protons, which is typical of the *cis* WC/WC G-A base pair geometry (Figure 2.4a). Several NOE signals were also observed that indicate stacking of the G9-A20 base pair on the C10-G19 base pair, including NOEs between A20 H2 and G19 NH and between A20 H1' and G19 NH (Figure 2.4a). Initial structural calculations performed without specific hydrogen-bond restraints for the G9-A20 base pair revealed a *cis* WC/WC G-A base pair geometry, thus, hydrogen bond restraints defining this geometry were included in subsequent rounds of structural calculations. The superposition of the 20 lowest-energy structures show that this G-A base pair is well defined by the NMR data, although two conformations with different propeller twists and buckles are observed in the

ensemble of structure (Figure 2.4b). These two alternative conformations observed in the NMR structures may reflect insufficient NMR restraints and/or conformational dynamics for this base pair. The absence of a detectable imino proton signal for G9 (Figure 2.2) supports conformational dynamics for the G9-A20 base pair.

An S-turn motif in the A730 loop domain. The structure of the A730 loop domain is defined by a large number of NOEs, including several unusual sequential NOEs in the G6-G9 stretch and non-sequential NOEs between nucleotides G6 and G9 (Figure 2.5a). These NOEs are consistent with the unusual ribose-phosphate backbone of the A730 loop domain, which adopts an S-turn between nucleotides G6 and C10. The S-turn is a common RNA motif, first structurally identified in the loop E of eukaryotic 5S rRNA [66] and the sarcin-ricin loop of 28S rRNA [67], but since found in other structural contexts [68-72]. In the A730 loop, the S-turn is created by ribose reversal at A8, with its 2'-OH group pointing in a direction opposite to the 2'-OH groups of adjacent nucleotides (Figure 2.5b). In the majority of the lowest-energy structures (17/20), the ribose of A8 adopts a 2'-endo conformation, a characteristic of an S-turn, which is in agreement with the intense H1'-H2' signal in the DQF-COSY spectrum (not shown). The S-turn of the A730 loop leads to bulging out of both the C7 and A8 residues with their Watson-Crick edges exposed in the minor groove. The adjacent G9-A20 base pair possesses a larger C1'-C1' distance than standard Watson-Crick base pairs that likely helps stabilize the S-turn. Three hydrogen bonds involving G9 and the protruded C7 and A8 bases are found in the ensemble of 20 lowest-energy structures and connect: 1) G9 O4' and A8 O2' (2.6 ± 0.4 Å); 2) G9 O2' and A8 N3 (3.3 ± 1.0 Å) and 3) G9 N3 and C7 N4 (3.4 ± 0.8 Å; Figure 2.5c). Interestingly, an S-turn motif has also been previously found in loop B of the hairpin ribozyme [70], suggesting that it may be important for catalysis by the VS ribozyme (see discussion).

Shifted pK_a values for adenines of the A730 loop. Given the predicted role of A756 as a general acid in the cleavage reaction, we were interested to determine if the structure of the A730 loop imparts a shifted pK_a value for A756. Hence, we determined the pK_a of adenines in the A730 loop of SLVI by ^{13}C NMR methods. It has been previously shown that the C2 chemical shift of AMP undergoes an 8-ppm upfield displacement upon protonation at its N1 position [73], and pH-dependent change in C2 chemical shifts have

been used to determine adenine pK_a 's in folded RNAs [27, 29, 36-39, 73]. In SLVI, only two C2 resonances are significantly affected by pH; the A8 and A20 C2 resonances are shifted upfield by 2.4 and 3.3 ppm, respectively, when the pH is decreased from 8.6 to 4.7 (Figure 2.6a). For both A8 and A20, a single C2-H2 crosspeak is observed in this pH range, indicating fast exchange dynamics on the NMR chemical shift timescale. In such cases, the change in C2 chemical shift at a given pH value can be used to derive pK_a values. We obtained a pK_a of 4.44 ± 0.10 for A8 and a pK_a of 4.74 ± 0.12 for A20 (Figure 2.6b and Supplementary Figure 2.1). These pK_a values are slightly higher than the pK_a value of adenine in single-stranded RNA (~ 3.7 ; [74]) and are compatible with the accessibility of the N1 positions of A8 and A20 in the SLVI structure. Local electronic effects, such as sequence context, base stacking and Mg^{2+} binding could be responsible for these small pK_a shifts. However, there is a discrepancy of about one pH unit between the pK_a of A8 and the catalytic pK_a (between 5.2-5.8; [23, 25]). Such discrepancy is not surprising, since it is expected that formation of the active site will change the chemical environment of the A756 base and thus, affect its pK_a .

DISCUSSION

We are using a modular approach to structurally characterize the *Neurospora* VS ribozyme, which essentially consists of determining the NMR structure of small fragments of the VS ribozyme that are relevant to its function. We previously determined structures of an active SLI internal loop and of an SLV fragment by NMR spectroscopy to provide novel information on substrate recognition and activation in the VS ribozyme [28, 30, 31]. Here, we performed NMR studies of an SLVI fragment containing the A730 loop in order to gain structural insights into the formation of the VS ribozyme active site. Below, we demonstrate that our structure is compatible with a large amount of biochemical data on the VS ribozyme and likely represents the conformation of the A730 loop in the folded ribozyme prior to docking with the cleavage site internal loop. We also discuss the resemblance of the A730 loop with loop B of the hairpin ribozyme and its relevance to formation of the active site.

The A730 loop adopts a undocked conformation compatible with chemical probing data. Chemical probing data were previously obtained for a self-cleaving ribozyme in both the absence and presence of Mg^{2+} ions [8]. As observed in our NMR studies of SLVI, the chemical probing data demonstrated that addition of Mg^{2+} ions stabilize the structure of the A730 loop domain, including several loop closing base pairs that were unstable in the absence of Mg^{2+} ions [8]. The Watson-Crick edges of C755 and A756 were still strongly modified under native conditions [8] in agreement with the structure of SLVI in which the Watson-Crick edges of these two bases are exposed in the minor groove. This suggests that under the conditions used for chemical probing, the A730 loop is not stably docked with the cleavage site internal loop, but can also exist in an undocked state that is compatible with the NMR structure.

Compatibility of the A730 loop structure with mutational and chemical modification data. Thorough mutational analyses of the A730 loop domain were previously carried out and revealed the importance of this loop for catalysis [16, 18]. In agreement with our structure of the A730 loop, the reversal of the four Watson-Crick base pairs within this domain (U753-G732, G754-C731, C758-G729 and A759-U728) did not significantly affect cleavage activity, but mutations of loop residues had more important effects [16]. Although all base substitutions at G757 (G9) and A730 (A20) led to reduced activity, replacing the G9-A20 combination by a G-C, a G-U or an A-A was less detrimental than other substitutions [16, 18]. Thus, these mutational data indicate the importance of a purine at position 757 and of *cis* WC/WC base pairing [75] between residues 757 and 730. This is in agreement with the role of the G9-A20 base pair in stabilizing the S-turn motif. Mutation of A756 (A8) by U, C or G caused the largest reduction in cleavage activity (~400-800 fold; [16, 18]), which is consistent with its proposed mechanistic role as a general acid [21, 23, 25]. Functional group modifications of A756 revealed that the Watson-Crick edge of A756 is particularly important for cleavage [17]. In the structure of SLVI, the Watson-Crick edge of A8 is exposed in the minor groove and accessible for docking with the cleavage site internal loop and performing its catalytic role.

Mg²⁺ binding to the S-turn. The backbone conformation of the S-turn creates two phosphate clusters, suggesting that specific binding of Mg²⁺ ions may be associated with at least one of these clusters. There is also evidence that other S-turn motifs contain a divalent metal binding site [71, 76-79]. Evidence for Mg²⁺ binding to the S-turn of the A730 loop is available from phosphorothioate substitution interference, which was previously employed to identify pro-Rp oxygens that are important for activity. Three inhibitory phosphorothioates, those of C755 (C7), G757 (G9) and C758 (C10), were found in the A730 loop domain (Figure 2.1c; [20, 63]). Phosphorothioate inhibition at G757 and C758 was suppressed in the presence of thiophilic manganese ions, indicating that the pro-Rp oxygens of G757 and C758 directly coordinate a divalent metal ion (Figure 2.1c; [63]). In the SLVI structure, the pro-Rp oxygens of C7 (C755), G9 (G757) and C10 (C758) are part of the S-turn motif, and the proximity of the C7 and G9 pro-Rp oxygens suggest that they may coordinate the same Mg²⁺ ion (Figure 2.5b).

Importance of the S-turn motif for catalysis. The A730 loop adopts an S-turn motif that exposes the Watson-Crick edge of A8 (A756), the proposed general acid of the VS cleavage reaction, in an enlarged minor groove. Interestingly, in the X-ray crystal structures of the hairpin ribozyme [60, 78], an S-turn was also found in loop B where it allows A38, the proposed general acid of the hairpin cleavage reaction, to protrude in a broad minor groove and make several hydrogen bonds with either a transition state analog or the 2'-cyclic phosphate and 5'-OH of the cleaved form [60, 78]. Thus, the structure of SLVI reveals additional similarities between the VS and hairpin ribozymes. These structural similarities between the A730 loop of the VS ribozyme and loop B of the hairpin ribozyme could not be predicted from the sequence of these internal loops, which are very different.

Certain non-VS sequences were previously found to partially substitute for the natural A730 loop in a cleavage assay [18]. In concordance with protrusion of A756 in the S-turn, it was concluded that a minimal requirement for self-cleavage is the presence of unpaired or non Watson-Crick paired nucleotides at the location of the A730 loop [18]. The most active substitution mutant contained the A-rich bulge of the *Tetrahymena* group I intron [80]. Interestingly, this bulge adopts a structure in which the phosphate backbone makes a corkscrew turn that exposes the Watson-Crick faces of several adenines [80]. Similarly, in the VS and hairpin ribozymes, one essential role of the S-turn motif may be to protrude a

catalytically important adenine in the active site. In the hairpin ribozyme, however, the catalytically-important adenine (A38), is protruded in the docked state [78], but not in the undocked state [70]. This may explain why substitution of the A730 loop domain by loop B of the hairpin ribozyme did not produce a functional VS ribozyme [18].

Insights into the formation of the active site. There is substantial evidence that the cleavage site loop and A730 loop of the VS ribozyme intimately associate to form the active site [4, 12, 16-25]. Given that the NMR structures of both the active conformation of the cleavage site loop [28] and the A730 loop are available, it is interesting to consider how these two loops might interact to form the active site. As described below, these two internal loops of the VS ribozyme display structural similarities to the two internal loops found in the active site of the hairpin ribozyme, and such similarities can be used to build a simple model of the VS ribozyme active site.

In the crystal structure of a transition state mimic of the hairpin ribozyme, the N1 of G8 is within hydrogen-bonding distance of the 2'-OH nucleophile, and its exocyclic amine forms a hydrogen bond with the scissile phosphate [60]. A comparison between this X-ray structure of the hairpin ribozyme and the NMR structure of the cleavage site loop of the VS ribozyme [28] reveals that the G residue proposed to serve as the general base (G638 in the VS ribozyme and G8 in the hairpin ribozyme) and the N-1 residue (G620 in the VS ribozyme and A-1 in the hairpin ribozyme) form similar cross-strand purine stacks. The heavy atom superposition between these two sets of bases illustrates the similarity of these purine stacks (Figure 2.7a; rmsd of 0.8 Å). Thus, only small changes in the relative positioning of G638 with respect to G620 would be necessary to bring G638 in the proper orientation with respect to the cleavage site. For modeling purposes, we modified the positions of the G620 and G638 bases in the cleavage site internal loop structure to mimic the position of functionally equivalent residues in the hairpin ribozyme (Figure 2.7b).

The superposition of Figure 2.7a highlights an important difference between the structures of the cleavage site loops from the hairpin and VS ribozymes; whereas the N-1 and N+1 nucleotide of the hairpin ribozyme (A-1 and G+1) adopt a splayed conformation, those of the VS ribozyme (G620 and A621) do not. The splayed conformation at the cleavage site is characteristic of nucleolytic ribozymes and likely important for stabilizing the in-line geometry between the 2'-OH nucleophile, the scissile phosphate and the 5'-

oxygen leaving group that is typical of S_N2 reactions [3]. Thus, it is highly likely that A621 must undergo a substantial conformation change to adopt the characteristic splayed conformation at the active site. For modeling purposes, the conformation of A621 was modified such that it mimics the position of the functionally equivalent G+1 in the hairpin ribozyme (Figure 2.7b). This brings A621 out of the helix and leaves G-A tandem base pairs in the cleavage site loop, which are known to stabilize internal loops when flanked with G-C base pairs, as found here [81].

Different strategies are adopted by nucleolytic ribozymes to stabilize the splayed conformation of the 5' and 3' nucleotides flanking the scissile phosphate, but they all rely on hydrogen bond and stacking interactions [3]. In the hairpin ribozyme, an important base stacking interaction involves G+1 of loop A and A38 of loop B, which orients this catalytically important base near the 5'-oxygen of the scissile phosphate (Figure 2.7a). As a last step in this modeling exercise, we attempted to reproduce this latter interaction by stacking A756 (A8) of SLVI on the protruded A621 of our SLI model. Such stacking is compatible with the production of a UV-dependent crosslink between a 4-thio-uridine at position 621 and A756, as previously observed under conditions that are compatible with catalytic activity [19]. Interestingly, given the protruded conformation of A756, it was possible to stack A756 on A621 and prevent atomic clashes between the two internal loops with only small displacement of the A756 base (Figure 2.7b). This stacking also brings the Watson-Crick edge of A756 near the negatively charged scissile phosphate. Such environment would likely stabilize the protonated form of A756 and shifts its pK_a towards neutrality [23, 25]. Furthermore, in the resulting model the minor grooves of the two internal loops interact with one another with a similar overall topology as in the hairpin ribozyme (not shown). Thus, by using the NMR structures of the A730 loop and the active conformation of the SLI internal loop, it is straightforward to model the active site of the VS ribozyme by homology with the active site of the hairpin ribozyme. The resulting model preserves the relative orientation of N-1, N+1 and the two key catalytic residues (G620, A621, A756 and G638 in the VS ribozyme) while maintaining the overall topology between the two helical domains. The coherence of the model with existing structural data clearly reinforces the idea that the VS and hairpin ribozymes may adopt similar active site architectures. Nevertheless, a high-resolution structure of the full VS ribozyme is needed to

reveal structural details of the active site that are beyond the specific findings of the present study.

Accession Codes. The NMR chemical shifts, NMR restraints, and atomic coordinates of the SLVI RNA have been deposited to the RCSB Protein Data Bank with BMRB entry 17292, RCSB entry rcsb101999, and PDB entry 2L5Z, respectively.

Acknowledgments

We thank R. Richter for computer support and D.L. Bryce for providing his structure refinement protocol with RDC restraints. We also thank R.A. Collins and J.G. Omichinski for critical reading of the manuscript.

Funding

Canadian Institutes for Health Research (CIHR) to P.L. (MOP-86502). Fonds Québécois de la Recherche sur la Nature et les Technologies, Luigi-Liberatore Foundation and Université de Montréal (UdeM) M.Sc. scholarships to G.D. UdeM M.Sc. scholarship to E.B. P.L. holds a Canada Research Chair in Structural Biology and Engineering of RNA. Funding for open access charges: CIHR.

REFERENCES

1. Collins, R.A., *The Neurospora Varkud satellite ribozyme*. Bioch. Soc. Trans. Rev., 2002. **30**: p. 1122-1126.
2. Lilley, D.M., *The Varkud satellite ribozyme*. RNA, 2004. **10**: p. 151-158.
3. Cochrane, J.C. and S.A. Strobel, *Catalytic strategies of self-cleaving ribozymes*. Acc. Chem. Res., 2008. **41**(8): p. 1027-35.
4. Lilley, D.M.J., *The hairpin and Varkud satellite ribozymes*, in *Ribozymes and RNA catalysis*, D.M.J. Lilley and F. Eckstein, Editors. 2008, Royal Society of Chemistry: Cambridge. p. 66-91.
5. Pyle, A.M., *Ribozymes: a distinct class of metalloenzymes*. Science, 1993. **261**: p. 709-714.
6. Saville, B.J. and R.A. Collins, *A site-specific self-cleavage reaction performed by a novel RNA in Neurospora mitochondria*. Cell, 1990. **61**: p. 685-696.
7. Guo, H.C.T., et al., *Nucleotide sequence requirements for self-cleavage of Neurospora VS RNA*. J. Mol. Biol., 1993. **232**: p. 351-361.
8. Beattie, T.L., J.E. Olive, and R.A. Collins, *A secondary-structure model for the self-cleaving region of Neurospora VS RNA*. Proc. Natl. Acad. Sci. U. S. A., 1995. **92**: p. 4686-4690.
9. Collins, R.A. and J.E. Olive, *Reaction conditions and kinetics of self-cleavage of a ribozyme derived from Neurospora VS RNA*. Biochemistry, 1993. **32**: p. 2795-2799.
10. Smith, M.D., et al., *The ionic environment determines ribozyme cleavage rate by modulation of nucleobase pKa*. RNA, 2008. **14**(9): p. 1942-9.
11. Rastogi, T., et al., *A long-range pseudoknot is required for activity of the Neurospora VS ribozyme*. EMBO J., 1996. **15**(11): p. 2820-2825.
12. Hiley, S.L. and R.A. Collins, *Rapid formation of a solvent-inaccessible core in the Neurospora Varkud satellite ribozyme*. EMBO J., 2001. **20**(19): p. 5461-5469.
13. Rastogi, T. and R.A. Collins, *Smaller, faster ribozymes reveal the catalytic core of Neurospora VS RNA*. J. Mol. Biol., 1998. **277**: p. 215-224.
14. Guo, H.C.T. and R.A. Collins, *Efficient trans-cleavage of a stem-loop RNA substrate by a ribozyme derived from Neurospora VS RNA*. EMBO J., 1995. **14**(12): p. 368-376.

15. Andersen, A. and R.A. Collins, *Rearrangement of a stable RNA secondary structure during VS ribozyme catalysis*. Mol. Cell, 2000. **5**: p. 469-478.
16. Lafontaine, D.A., et al., *The A730 loop is an important component of the active site of the VS ribozyme*. J. Mol. Biol., 2001. **312**: p. 663-674.
17. Lafontaine, D.A., et al., *Functional group requirements in the probable active site of the VS ribozyme*. J. Mol. Biol., 2002. **323**: p. 23-34.
18. Sood, V.D. and R.A. Collins, *Identification of the catalytic subdomain of the VS ribozyme and evidence for remarkable sequence tolerance in the active site loop*. J. Mol. Biol., 2002. **320**: p. 443-454.
19. Hiley, S.L., et al., *4-thio-U cross-linking identifies the active site of the VS ribozyme*. EMBO J., 2002. **21**(17): p. 4691-4698.
20. Jones, F.D. and S.A. Strobel, *Ionization of a critical adenosine residue in the Neurospora Varkud satellite ribozyme active site*. Biochemistry, 2003. **42**: p. 4265-4276.
21. Zhao, Z.Y., et al., *Nucleobase participation in ribozyme catalysis*. J. Am. Chem. Soc., 2005. **127**(14): p. 5026-7.
22. Wilson, T.J., A.C. McLeod, and D.M. Lilley, *A guanine nucleobase important for catalysis by the VS ribozyme*. EMBO J., 2007. **26**(10): p. 2489-500.
23. Smith, M.D. and R.A. Collins, *Evidence for proton transfer in the rate-limiting step of a fast-cleaving Varkud satellite ribozyme*. Proc. Natl. Acad. Sci. U. S. A., 2007. **104**(14): p. 5818-23.
24. Jaikaran, D., et al., *An important role of G638 in the cis-cleavage reaction of the Neurospora VS ribozyme revealed by a novel nucleotide analog incorporation method*. RNA, 2008. **14**(5): p. 938-49.
25. Wilson, T.J., et al., *Nucleobase-mediated general acid-base catalysis in the Varkud satellite ribozyme*. Proc. Natl. Acad. Sci. U. S. A., 2010. **107**(26): p. 11751-6.
26. Michiels, P.J.A., et al., *Structure of the ribozyme substrate hairpin of Neurospora VS RNA: A close look at the cleavage site*. RNA, 2000. **6**: p. 1821-1832.
27. Flinders, J. and T. Dieckmann, *A pH controlled conformational switch in the cleavage site of the VS ribozyme substrate RNA*. J. Mol. Biol., 2001. **308**: p. 665-679.

28. Hoffmann, B., et al., *NMR structure of the active conformation of the Varkud satellite ribozyme cleavage site*. Proc. Natl. Acad. Sci. U. S. A., 2003. **100**(12): p. 7003-7008.
29. Flinders, J. and T. Dieckmann, *The solution structure of the VS ribozyme active site loop reveals a dynamic "hot-spot"*. J. Mol. Biol., 2004. **341**: p. 935-949.
30. Campbell, D.O. and P. Legault, *NMR structure of the Varkud satellite ribozyme stem-loop V RNA and magnesium-ion binding from chemical-shift mapping*. Biochemistry, 2005. **44**: p. 4157-4170.
31. Campbell, D.O., et al., *NMR structure of Varkud satellite ribozyme stem-loop V in the presence of magnesium ions and localization of metal-binding sites*. Biochemistry, 2006. **45**(35): p. 10591-605.
32. Lafontaine, D.A., D.G. Norman, and D.M. Lilley, *Structure, folding and activity of the VS ribozyme: importance of the 2-3-6 helical junction*. EMBO J., 2001. **20**(6): p. 1415-1424.
33. Lafontaine, D.A., D.G. Norman, and D.M. Lilley, *The global structure of the VS ribozyme*. EMBO J., 2002. **21**(10): p. 2461-2471.
34. Lipfert, J., et al., *The complete VS ribozyme in solution studied by small-angle X-ray scattering*. Structure, 2008. **16**(9): p. 1357-67.
35. Elder, D. and R.J. Harris, *Ribozymes: the hairpin and Varkud ribozymes are related*. Riv. Biol., 2003. **96**(3): p. 433-9.
36. Legault, P. and A. Pardi, *Unusual dynamics and pKa shifts at the active site of a lead-dependent ribozyme*. J. Am. Chem. Soc., 1997. **119**: p. 6621-6628.
37. Cai, Z. and I.J. Tinoco, *Solution structure of loop A from the hairpin ribozyme from Tobacco ringspot virus satellite*. Biochemistry, 1996. **35**: p. 6026-6036.
38. Smith, J.S. and E.P. Nikonowicz, *NMR structure and dynamics of an RNA motif common to the spliceosome branch-point helix and the RNA-binding site for phage GA coat protein*. Biochemistry, 1998. **37**(39): p. 13486-98.
39. Ravindranathan, S., S.E. Butcher, and J. Feigon, *Adenine protonation in domain B of the hairpin ribozyme*. Biochemistry, 2000. **39**(51): p. 16026-32.
40. Legault, P., *Structural studies of ribozymes by heteronuclear NMR spectroscopy*, 1995, University of Colorado at Boulder: Boulder.

41. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes*. J. Biomol. NMR, 1995. **6**: p. 277-293.
42. Johnson, B.A. and R.A. Blevins, *NMRView: a computer program for the visualization and analysis of NMR data*. J. Biomol. NMR, 1994. **4**: p. 603-614.
43. Dingley, A.J. and S. Grzesiek, *Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide $^2J_{NN}$ couplings*. J. Am. Chem. Soc., 1998. **120**: p. 8293-8297.
44. Piotto, M., V. Saudek, and V. Skleňár, *Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions*. J. Biomol. NMR, 1992. **2**: p. 661-665.
45. Grzesiek, S. and A. Bax, *The importance of not saturating water in protein NMR. Application to sensitivity enhancement of NOE measurements*. J. Am. Chem. Soc., 1993. **115**: p. 12593-12594.
46. Zhang, O., et al., *Backbone 1H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques*. J. Biomol. NMR, 1994. **4**(6): p. 845-58.
47. Mueller, L., P. Legault, and A. Pardi, *Improved RNA structure determination by detection of NOE contacts to exchange-broadened amino groups*. J. Am. Chem. Soc., 1995. **117**: p. 11043-11048.
48. Ikura, M., et al., *Three-dimensional NOESY-HMQC spectroscopy of a ^{13}C -labeled protein*. J. Magn. Reson., 1990. **86**: p. 204-209.
49. Rance, M., et al., *Improved spectral resolution in cosy 1H NMR spectra of proteins via double quantum filtering*. Biochem. Biophys. Res. Commun, 1983. **117**(2): p. 479-85.
50. Schwalbe, H., et al., *Determination of a complete set of coupling constants in ^{13}C -labeled oligonucleotides*. J. Biomol. NMR, 1994. **4**: p. 631-644.
51. Wijmenga, S.S., M.M.W. Mooren, and C.W. Hilbers, *NMR of nucleic acids; from spectrum to structure*, in *NMR of macromolecules: a practical approach.*, G.C.K. Roberts, Editor. 1993, Oxford University Press: New York. p. 217-288.
52. Hansen, M.R., L. Mueller, and A. Pardi, *Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions*. Nature Struct Biol, 1998. **5**: p. 1064-1074.

53. Brutscher, B., et al., *Improved sensitivity and resolution in H-1-C-13 NMR experiments of RNA*. J. Am. Chem. Soc., 1998. **120**: p. 11845-11851.
54. Boisbouvier, J., et al., *Resolution-optimized NMR measurement of (1)D(CH), (1)D(CC) and (2)D(CH) residual dipolar couplings in nucleic acid bases*. J. Biomol. NMR, 2004. **30**(3): p. 287-301.
55. Schwieters, C.D., et al., *The Xplor-NIH NMR molecular structure determination package*. J. Magn. Reson., 2003. **160**: p. 66-74.
56. Van Melckebeke, H., et al., *Liquid-crystal NMR structure of HIV TAR RNA bound to its SELEX RNA aptamer reveals the origins of the high stability of the complex*. Proc. Natl. Acad. Sci. USA, 2008. **105**(27): p. 9210-5.
57. Dosset, P., et al., *A novel interactive tool for rigid-body modeling of multi-domain macromolecules using residual dipolar couplings*. J. Biomol. NMR, 2001. **20**(3): p. 223-31.
58. Clore, G.M., A.M. Gronenborn, and N. Tjandra, *Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude*. J. Magn. Reson., 1998. **131**(1): p. 159-62.
59. Lavery, R., et al., *Conformational analysis of nucleic acids revisited: Curves+*. Nucleic Acids Res., 2009. **37**(17): p. 5917-29.
60. Rupert, P.B., et al., *Transition state stabilization by a catalytic RNA*. Science, 2002. **298**(5597): p. 1421-4.
61. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics*. Acta Crystallogr. D Biol. Crystallogr., 2004. **60**(Pt 12 Pt 1): p. 2126-32.
62. Pearlman, D.A., et al., *Amber, a package of computer-programs for applying molecular mechanics, normal-mode analysis, molecular-dynamics and free-energy calculations to simulate the structural and energetic properties of molecules*. Comput. Phys. Commun., 1995. **91**(1-3): p. 1-41.
63. Sood, V.D., T.L. Beattie, and R.A. Collins, *Identification of phosphate groups involved in metal binding and tertiary interactions in the core of the Neurospora VS ribozyme*. J. Mol. Biol., 1998. **282**: p. 741-750.
64. Heus, H.A. and A. Pardi, *Structural features that give rise to the unusual stability of RNA hairpins containing GNRA tetraloops*. Science, 1991. **253**(5016): p. 191-194.

65. Jucker, F.M., et al., *A network of heterogeneous hydrogen bonds in GNRA tetraloops*. J. Mol. Biol., 1996. **264**(5): p. 968-80.
66. Wimberly, B., G. Varani, and I. Tinoco, Jr., *The conformation of loop E of eukaryotic 5S ribosomal RNA*. Biochemistry, 1993. **32**(4): p. 1078-87.
67. Szewczak, A.A., et al., *The conformation of the sarcin/ricin loop from 28S ribosomal RNA*. Proc. Natl. Acad. Sci. U. S. A., 1993. **90**(20): p. 9581-5.
68. Correll, C.C., et al., *Crystal structure of the ribosomal RNA domain essential for binding elongation factors*. Proc. Natl. Acad. Sci. U. S. A., 1998. **95**(23): p. 13436-41.
69. Leontis, N.B. and E. Westhof, *A common motif organizes the structure of multi-helix loops in 16 S and 23 S ribosomal RNAs*. J. Mol. Biol., 1998. **283**(3): p. 571-83.
70. Butcher, S.E., F.H. Allain, and J. Feigon, *Solution structure of the loop B domain from the hairpin ribozyme*. Nat. Struct. Biol., 1999. **6**(3): p. 212-6.
71. Zimmermann, G.R., et al., *Molecular interactions and metal binding in the theophylline-binding core of an RNA aptamer*. RNA, 2000. **6**: p. 659-667.
72. Allain, F.H., et al., *Solution structure of the two N-terminal RNA-binding domains of nucleolin and NMR study of the interaction with its RNA target*. J. Mol. Biol., 2000. **303**(2): p. 227-41.
73. Legault, P. and A. Pardi, *In situ probing of adenine protonation in RNA by ¹³C NMR*. J. Am. Chem. Soc., 1994. **116**: p. 8390-8391.
74. Moody, E.M., J.T. Lecomte, and P.C. Bevilacqua, *Linkage between proton binding and folding in RNA: a thermodynamic framework and its experimental application for investigating pKa shifting*. RNA, 2005. **11**(2): p. 157-72.
75. Leontis, N.B., J. Stombaugh, and E. Westhof, *The non-Watson-Crick base pairs and their associated isosteric matrices*. Nucleic Acids Res., 2002. **30**(16): p. 3497-531.
76. Hermann, T. and E. Westhof, *Exploration of metal ion binding sites in RNA folds by Brownian-dynamics simulations*. Structure, 1998. **6**(10): p. 1303-14.
77. Butcher, S.E., F.H.-T. Allain, and J. Feigon, *Determination of metal ion binding sites within the hairpin ribozyme domains by NMR*. Biochemistry, 2000. **39**: p. 2174-2182.

78. Rupert, P.B. and A.R. Ferré-D'Amaré, *Crystal structure of a haripin ribozyme-inhibitor complex with implications for catalysis*. *Nature*, 2001. **410**: p. 780-786.
79. Alam, S., et al., *Conformational heterogeneity at position U37 of an all-RNA hairpin ribozyme with implications for metal binding and the catalytic structure of the S-turn*. *Biochemistry*, 2005. **44**(44): p. 14396-408.
80. Cate, J.H., et al., *Crystal structure of a group I ribozyme domain: principles of RNA packing*. *Science*, 1996. **273**: p. 1678-1685.
81. Walter, A.E., M. Wu, and D.H. Turner, *The stability and structure of tandem GA mismatches in RNA depend on closing base pairs*. *Biochemistry*, 1994. **33**(37): p. 11349-54.

TABLE

Table 2.1. Structural statistics of the SLVI RNA.

Distance restraints	1017
Number of NOE-derived distance restraints	965
Inter-nucleotide	591
Intra-nucleotide	355
Ambiguous	19
Hydrogen-bond restraints	52
Dihedral angle restraints	88
Residual dipolar coupling restraints	30
Total number of restraints	1135
Rmsd from experimental restraints	
NOE (Å) (none > 0.2)	0.0075 ± 0.004
Dihedral (°) (none > 5°)	0.089 ± 0.009
Residual dipolar couplings (none > 5 Hz)	0.24 ± 0.02
Rmsd from idealized geometry	
Bonds (Å)	0.00555 ± 0.00002
Angles (°)	1.1968 ± 0.0006
Improper (°)	0.422 ± 0.005
Heavy-atom rmsd to the minimized average structure (Å)	
Overall (residues 2-25)	0.67 ± 0.17
5'-3' stem (residues 2-5 and 22-25)	0.53 ± 0.07
Hairpin stem (residues 11-18)	0.31 ± 0.01
A730 loop (residues 6-10 and 19-21)	0.65 ± 0.24

FIGURES

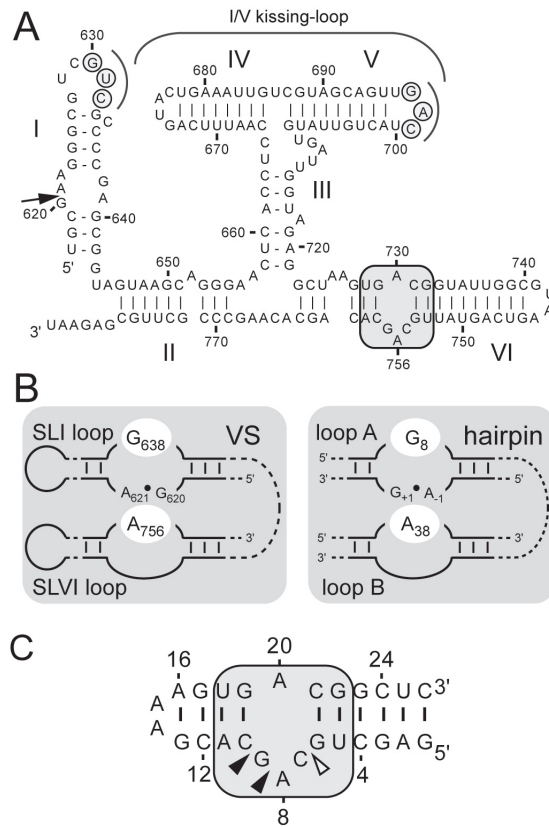


Figure 2.1. The *Neurospora* VS ribozyme and its A730 internal loop. (a) Primary and secondary structures of the VS ribozyme (wild-type sequence nucleotides 617-783). The site of self-cleavage is indicated by an arrow, and circled nucleotides in loops I and V form the I/V kissing-loop interaction. (b) Schematic of the VS ribozyme (left) and hairpin ribozyme (right) illustrating similarities at the active site (see text). The residues highlighted with white circles are key players of proposed general acid-based mechanisms [4]. (c) Primary and predicted secondary structures of the 26-nt SLVI RNA fragment, which includes the A730 loop domain (grey box). Phosphate groups that display inhibitory effect when substituted by R_p phosphorothioate are indicated by an arrow [20, 63], and the arrow is filled in those cases where the inhibition could be suppressed by addition of manganese ions [63].

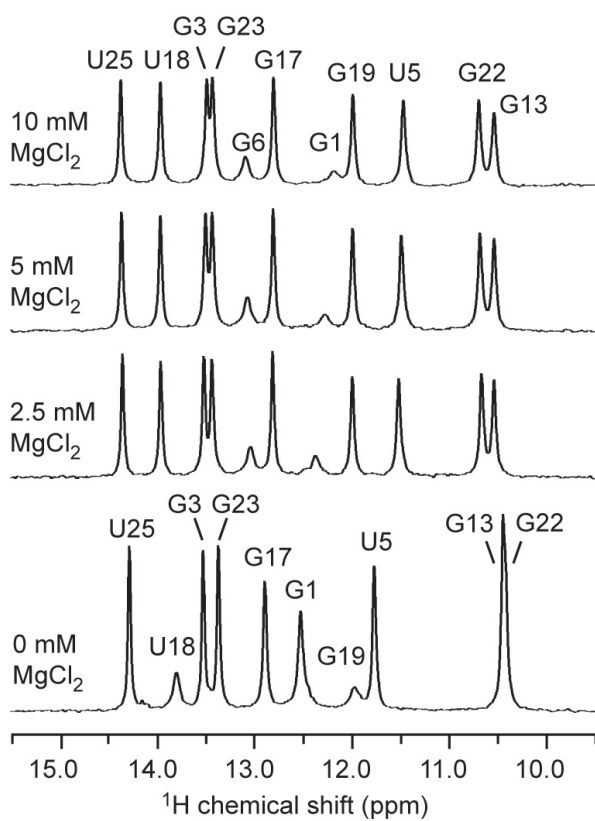


Figure 2.2. Stabilization of the A730 loop by Mg^{2+} ions. Imino regions of 1D flip-back watergate [44, 45] ^1H spectra of SLVI collected at 15°C in NMR buffer containing different concentrations of free MgCl_2 . Imino proton assignments were derived from 2D NOESY spectra collected in NMR buffer at 0, 5, and 10 mM MgCl_2 .

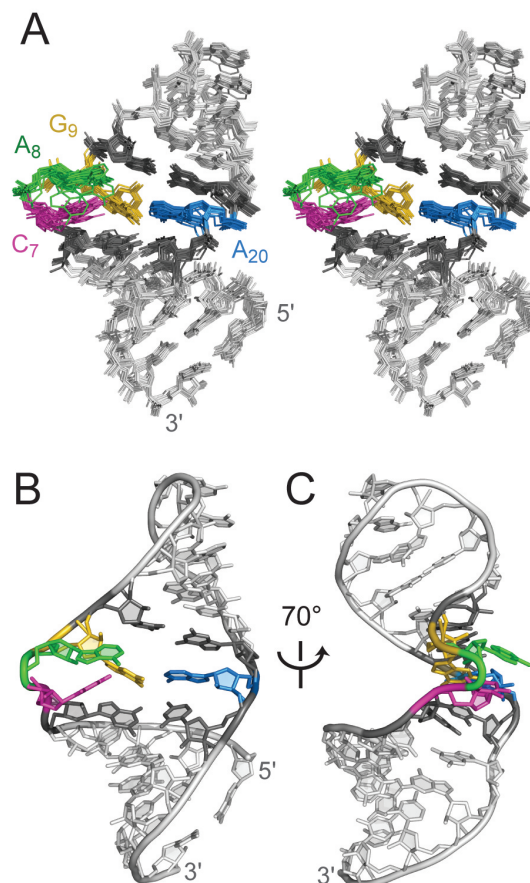


Figure 2.3. NMR solution structure of the SLVI RNA fragment. (a) Stereoview of the 20 lowest-energy structures. The superposition was made on the minimized average structure (not shown) over heavy atoms of residues 2-25. The view is into the minor groove of the A730 active site internal loop. (b) and (c) Stick representations of the lowest-energy structure of SLVI. For simplicity only heavy atoms are shown and the ribbon replacing the phosphorus and non-bonded oxygen atoms is used to indicate the backbone topology. SLVI nucleotides are color-coded: the loop closing base pairs (G6-C21 and C10-G19) are dark gray, C7 (C755) is magenta, A8 (A756) is green, G9 (G757) is gold and A20 (A730) is blue.

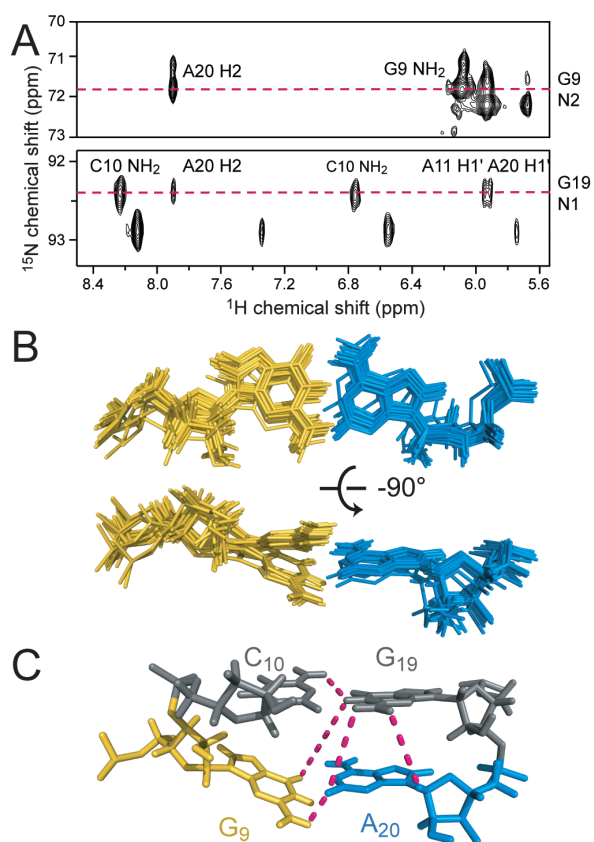


Figure 2.4. Formation of a *cis* WC/WC G9-A20 base pair in the A730 loop. (a) Selected regions from a 2D ^1H - ^{15}N CPMG-NOESY spectrum showing NOEs that define the geometry of the G9-A20 base pair. The spectrum was collected at 15°C with a mixing time of 160 ms. (b) The G9-A20 base pair in the 20 lowest-energy structures. The superposition is from Figure 2.3a. (c) Stacking of the G9-A20 base pair onto the C10-G19 base pair in the lowest-energy structure of SLVI. Dashed lines connect protons for which a NOE is observed in (a).

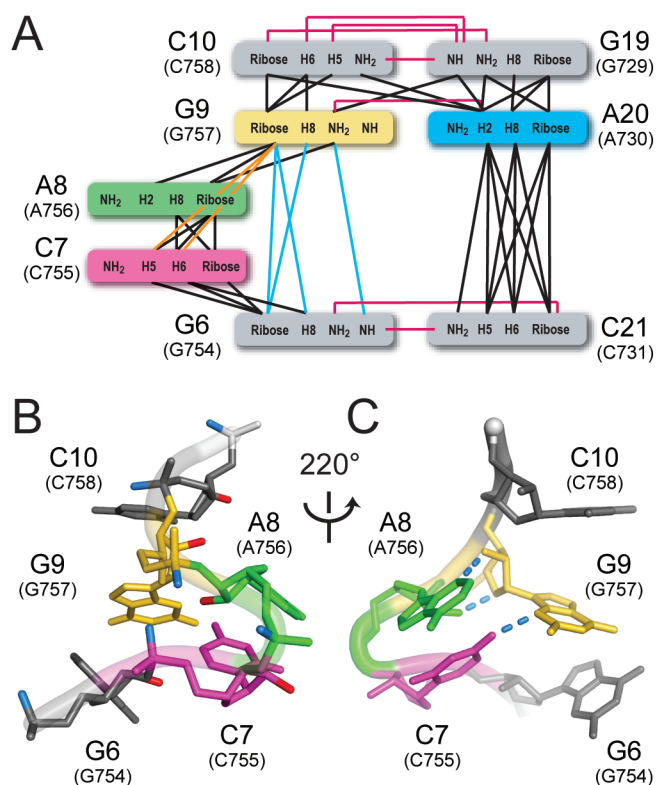


Figure 2.5. S-turn motif in the A730 loop of the SLVI RNA. (a) Schematic summarizing the inter-residue NOEs for the A730 internal loop of the SLVI RNA. Black lines indicate NOEs between nucleotides that are adjacent in the sequence, pink lines indicate NOEs between base-pairing residues, blue lines indicate NOEs between G6 and G9 and orange lines refer to NOEs between C7 and G9. For simplicity, all ribose protons (H1', H2', H3', H4', H5' and H5'') were grouped under the ribose denomination. (b) and (c) Close up views of the S-turn motif in the lowest-energy structure showing (b) the ribose reversal at A8 and nearby phosphates and (c) stacking of C7 and A8 in the minor groove and stabilizing hydrogen bonds. In (b) the pro-Rp oxygens are shown in blue and the 2'-oxygens in red. In (c) three hydrogen bonds are shown (A8 N3 : G9 2'-OH, C7 NH₂ : G9 N3 and A8 2'-OH : G9 O4') that likely stabilize the S-turn motif. For simplicity only heavy atoms are shown and the ribbon replacing the phosphorus and non-bonded oxygen atoms is used to indicate the backbone topology.

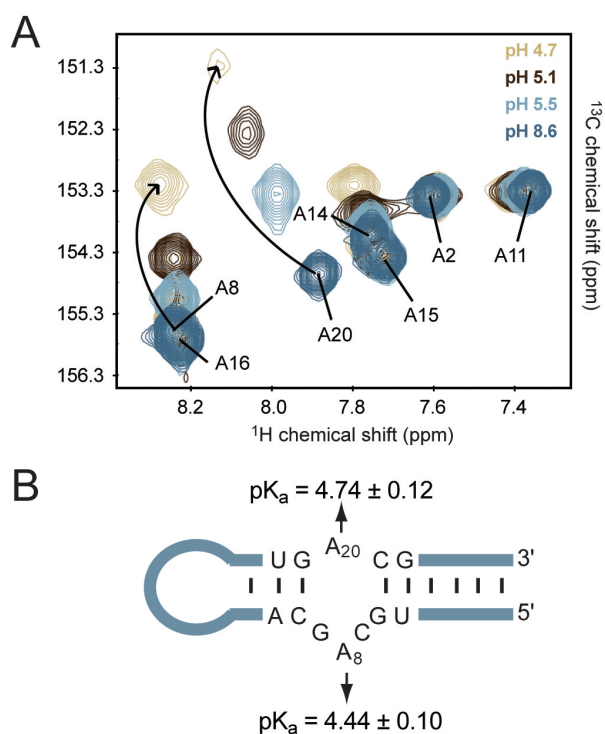


Figure 2.6. Determination of adenine pK_a 's in SLVI. (a) Superposition of the aromatic C2-H2 regions of 2D ^1H - ^{13}C HMQC spectra collected at 25°C and at pH 4.7 (beige), pH 5.1 (black), pH 5.5 (pale blue) and pH 8.6 (grayish blue). Arrows point to significant pH-dependent changes in ^{13}C chemical shift for A8 and A20. (b) Summary of the adenine pK_a values in the A730 internal loop.

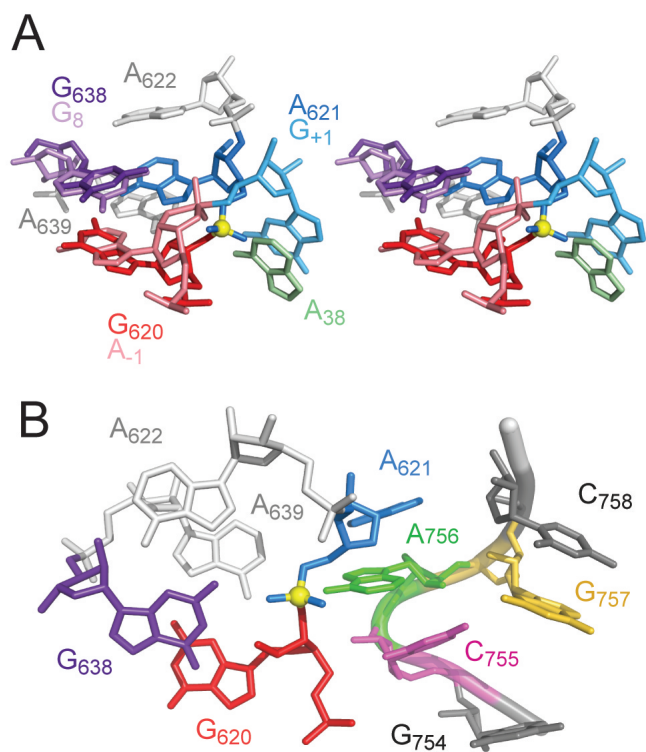


Figure 2.7. Homology modeling of the VS ribozyme active site. (a) Heavy atom superposition of the G638 and G620 nucleotides of the VS ribozyme with the G8 and A-1 nucleotides of the hairpin ribozyme (pdb entry 1M50; [60]). (b) Modeling of the active site by association of the substrate internal loop and the A730 internal loop (see text). For simplicity only heavy atoms are shown and the ribbon replacing the phosphorus and non-bonded oxygen atoms is used to indicate the backbone topology. The yellow sphere represents the scissile phosphate.

SUPPORTING INFORMATION

Supplementary Table 2.1. Resonance assignment of SLVI. ^a

NON-EXCHANGEABLE PROTON CHEMICAL SHIFTS (ppm)										
(ppm +/- 0.01)										
	H1'	H2'	H3'	H4'	H5'	H5''	H2	H5	H6	H8
G1	5.61	4.76	4.64	4.30	4.02	3.91				7.99
A2	6.06	4.70	4.81	4.58	4.63	4.25	7.59			8.09
G3	5.72	4.46	4.50	4.50	4.53	4.13				7.38
C4	5.54	4.62	4.31	4.45	4.55	4.10		5.16	7.56	
U5	5.65	4.15	4.56	4.41	4.55	4.11		5.71	7.83	
G6	5.68	4.15	4.85	4.38	4.52	4.18				8.02
C7	5.94	4.47	4.65	4.46	4.33	4.25		5.68	7.83	
A8	6.18	4.74	4.98	4.65	4.37	4.32	8.24			8.58
G9	5.84	4.83	4.62	4.59	4.52	4.37				8.05
C10	5.53	4.45	4.72	4.54	4.55	4.25		5.25	7.67	
A11	5.95	4.44	4.81	4.55	4.63	4.26	7.34			8.14
C12	5.54	4.30	4.47	4.53	4.55	4.13		5.10	7.38	
G13	5.74	4.33	4.70	4.40	4.44	4.16				7.44
A14	5.67	4.77	4.44	4.22	4.30	3.95	7.74			8.33
A15	5.39	4.37	4.56	4.35	4.26	3.96	7.70			7.89
A16	5.99	4.55	4.90	4.45	4.64	4.26	8.21			8.14
G17	3.57	4.21	4.09	4.21	4.31	4.21				7.98
U18	5.54	4.63	4.31	4.45	4.57	4.11		5.04	7.80	
G19	5.71	4.54	4.69	4.52	4.58	4.19				7.71
A20	5.92	4.66	4.57	4.56	4.63	4.23	7.88			7.94
C21	5.36	4.31	4.53	4.41	4.56	4.15		5.44	7.64	
G22	5.75	4.64	4.58	4.52	4.49	4.15				7.50
G23	5.65	4.45	4.57	4.44	4.58	4.13				7.33
C24	5.55	4.39	4.48	4.44	4.58	4.10		5.25	7.72	
U25	5.62	4.39	4.55	4.44	4.58	4.12		5.46	7.97	
C26	5.85	4.04	4.21	4.22	4.56	4.11		5.73	7.81	
CARBON CHEMICAL SHIFTS (ppm)										
(ppm +/- 0.4)										
	C1'	C2'	C3'	C4'	C5'	C2	C5	C6	C8	
G1	92.6	75.1	75.5	84.6	62.2				138.8	
A2	92.8	75.6	72.7	82.1	64.8	153.4			140.0	
G3	92.5	75.4	72.7	81.8	65.3				135.8	
C4	93.7	75.3	72.0	82.0	64.3		97.1	140.6		
U5	94.7	75.5	73.0	82.8	65.2		105.9	141.3		
G6	90.2	77.1	75.5	83.6	65.9				138.7	
C7	92.4	76.2	74.5	83.1	66.0		98.3	142.7		
A8	89.9	77.6	77.4	86.1	67.3	155.6			141.8	
G9	93.1	75.4	74.0	83.4	66.5				139.2	
C10	94.4	75.2	72.0	82.3	64.7		97.2	141.0		
A11	92.7	75.6	72.2	81.9	64.7	153.3			139.6	
C12	93.0	75.9	71.9	82.3	64.3		97.5	139.4		
G13	93.7	75.8	71.7	81.9	64.3				136.1	
A14	93.1	75.5	72.5	83.0	63.7	154.0			142.1	
A15	93.1	75.6	73.0	82.8	64.4	154.4			139.9	

A16	91.8	76.3	72.8	82.1	65.0	155.8			140.0	
G17	93.4	74.5	74.6	82.7	69.4				137.7	
U18	93.4	75.1	72.1	81.9	64.5		102.4	141.6		
G19	92.5	75.5	73.1	82.0	65.1				136.6	
A20	92.7	75.5	72.7	82.1	65.0	154.6			139.0	
C21	93.1	75.7	72.1	81.8	64.3		98.6	139.7		
G22	92.8	75.3	72.8	82.0	65.3				136.2	
G23	92.8	75.3	72.5	82.2	64.5				136.0	
C24	93.9	75.4	71.9	81.7	64.3		97.1	141.0		
U25	93.9	75.4	72.1	81.8	64.3		103.3	142.2		
C26	92.9	77.5	69.7	83.3	65.2		98.0	142.1		

BASE EXCHANGEABLE PROTON CHEMICAL SHIFTS (ppm)								
	H1 ^b	H21 ^b	H22 ^b	H3 ^b	H41 ^b	H42 ^b	H61 ^b	H62 ^b
G1	12.19							
A2								
G3	13.51							
C4					8.35	6.73		
U5				11.51				
G6	13.08							
C7								
A8							6.80	6.80
G9		6.07	6.07					
C10					8.22	6.75		
A11								
C12					8.11	6.54		
G13	10.54							
A14							6.37	6.37
A15							6.35	6.35
A16								
G17	12.83							
U18				13.97				
G19	12.01							
A20								
C21					8.50	6.81		
G22	10.70	5.91	5.91					
G23	13.44							
C24					8.54	6.86		
U25				14.38				
C26					8.22	6.95		
BASE NITROGEN CHEMICAL SHIFTS (ppm)								
	N1 ^c	N2 ^d	N3 ^e	N4 ^d	N6 ^f	N7 ^f	N9 ^f	
G1	145.9					231.3	169.7	
A2	221.4		213.1		82.4	228.7	171.2	
G3	148.2	75.6				233.7	169.7	
C4	151.7		196.1	99.0				
U5	146.1		158.0					
G6	148.1					235.7	170.0	
C7	152.6							
A8	226.0		217.1		79.1	232.7	168.9	
G9		71.8				233.1	169.8	
C10	151.6		198.1	97.0				
A11	220.9		213.3		83.8	230.3	171.3	
C12	150.7		195.4	99.2				
G13	145.5	75.8				230.1		
A14	224.5		215.9		78.8	228.4	169.9	
A15	227.3		214.7		80.4	225.9	170.4	
A16	226.8		215.6		80.9	231.6	173.0	
G17	147.2	75.3				232.0	169.5	
U18	146.5		162.3					
G19	146.8	74.7				234.2		

A20	222.8		213.1		84.9	230.3	171.1	
C21	150.8		197.7	99.6				
G22	142.9	72.2				234.5	170.4	
G23	148.5	76.5				233.6	169.9	
C24	151.2		197.1	99.7				
U25	146.9		162.7					
C26	152.9		197.4	97.3				

- a. Note that there is ambiguity on the assignment of certain pairs of protons (H5' and H5'', H21 and H22, H41 and H42, H61 and H62)
- b. +/- 0.01 ppm.
- c. +/- 0.1 ppm for uridines, guanines and cytosines, 0.2 ppm for adenines.
- d. +/- 0.1 ppm.
- e. +/- 0.1 ppm for uridines, 0.2 ppm for adenines and 0.5 ppm for cytosines.
- f. +/- 0.2 ppm.

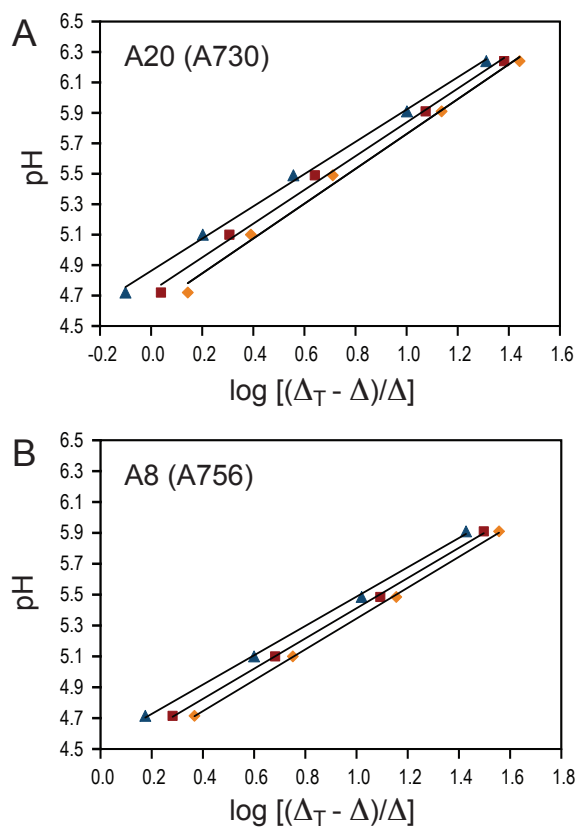


Figure S2.1. Hill plot of the pH titration data for (a) A20 and (b) A8. The pK_a values were derived by linear regression based on the equation $pH = \log[(\Delta_T - \Delta)/\Delta] + pK_a$ [36], where Δ is the change in C2 chemical shift at a given pH relative to the unprotonated state (pH 8.6) and Δ_T is the total change in chemical shift between the protonated and unprotonated states. Using values of Δ_T estimated at 6 ppm (blue triangles), 7 ppm (burgundy squares) and 8 ppm (mustard diamonds), pK_a values of 4.86, 4.73 and 4.62 were respectively obtained for A20 ($pK_a = 4.74 \pm 0.12$) and pK_a values of 4.54, 4.43 and 4.35 were respectively obtained for A8 ($pK_a = 4.44 \pm 0.10$).

3. Localisation des ions Mg^{2+} dans la structure RMN de la boucle A730

Mise en contexte

L'étude précédente nous a permis d'approfondir notre compréhension du repliement de la boucle A730 en identifiant les motifs structuraux définissant la structure de ce sous-domaine important. Ces motifs sont une paire de bases *cis* WC/WC A₇₃₀-G₇₅₇ et un motif *S-turn* (chapitre 2). L'obtention de cette structure RMN était particulièrement intéressante pour permettre de comprendre comment la boucle A730 arrive à positionner le A₇₅₆ pour l'activité catalytique. Nous avons noté lors de cette première étude que la boucle A730 adopte une structure stable uniquement en présence d'ions Mg²⁺ (Chapitre 2). Étant donné que dans le motif *S-turn*, trois groupements phosphates se retrouvent à proximité les uns des autres (C₇₅₅, A₇₅₆ et G₇₅₇), il est possible que ce motif forme un site de liaison de métaux. Afin de tester cette hypothèse, nous avons décidé de localiser les ions Mg²⁺ dans la structure RMN de la boucle A730.

Cette deuxième étude a été accomplie en combinant deux approches différentes de spectroscopie RMN. Premièrement, la relaxation paramagnétique induite par le manganèse (*Mn²⁺-induced paramagnetic relaxation enhancement*) nous a fourni de nombreuses contraintes entre l'ARN et les métaux se liant à la boucle A730. Ensuite, nous avons évalué la contribution de certains groupements phosphates à la liaison des ions Mg²⁺ en utilisant l'approche de perturbation de déplacement chimique induite par le cadmium (*Cd²⁺-induced chemical shift perturbation*) sur des ARN ayant des modifications phosphorothioates. Cette étude nous a permis d'identifier la liaison de cinq ions Mg²⁺ dans l'ARN SLVI. Plus particulièrement, deux ions Mg²⁺ sont coordonnés par la boucle A730. Le premier (site 2) est coordonné de façon indirecte (*outer-sphere*) par le groupement 5'-phosphate du C₇₅₈. Le deuxième ion Mg²⁺ (site 3) est coordonné de façon indirecte par les groupements 5'-phosphates des résidus C₇₅₅ et A₇₅₆ et de façon directe par le groupement 5'-phosphate du G₇₅₇. La liaison de ce métal semble contribuer à la stabilisation structurale et expliquer la dépendance aux ions divalents de la boucle A730. De plus, les trois autres ions Mg²⁺ se liant à l'ARN SLVI permettent de mieux comprendre les interactions ARN-métal. Nous avons aussi pu comparer la localisation des ions Mg²⁺ dans ces cinq sites avec des sites de liaison similaires obtenus à partir des bases de données.

Dans ce deuxième article, j'ai contribué avec le Dr. Pascale Legault à la conception du projet. J'ai également réalisé la préparation des échantillons en plus de collecter et d'analyser les expériences de spectroscopie UV. De plus, la collecte et l'analyse des

données de RMN, le calcul et l'analyse des structures ainsi que la rédaction de l'article a été réalisé par le Dr. Pascale Legault et moi-même.

NMR Localization of Divalent Cations at the Active Site of the
Neurospora VS Ribozyme Provides Insights into RNA—Metal—Ion
Interactions

Eric Bonneau and Pascale Legault*

Département de Biochimie et Médecine Moléculaire, Université de Montréal, C.P. 6128,
Succursale Centre-Ville, Montréal, Québec H3C 3J7, Canada

This research was originally published in Biochemistry:

Biochemistry. 2014 Jan 28;53(3):579-90. doi: 10.1021/bi401484a. Epub 2014 Jan 10

ABSTRACT

Metal cations represent key elements of RNA structure and function. In the *Neurospora* VS ribozyme, metal cations play diverse roles; they are important for substrate recognition, formation of the active site, and shifting the pK_a 's of two key nucleobases that contribute to the general acid-base mechanism. Recently, we determined the NMR structure of the A730 loop of the VS ribozyme active site (SLVI) that contributes the general acid (A_{756}) in the enzymatic mechanism of the cleavage reaction. Our studies showed that magnesium (Mg^{2+}) ions are essential to stabilize the formation of the S-turn motif within the A730 loop that exposes the A_{756} nucleobase for catalysis. In this article, we extend these NMR investigations by precisely mapping the Mg^{2+} -ion binding sites using manganese-induced paramagnetic relaxation enhancement and cadmium-induced chemical-shift perturbation of phosphorothioate RNAs. These experiments identify five Mg^{2+} -ion binding sites within SLVI. Four Mg^{2+} ions in SLVI are associated with known RNA structural motifs, including the G-U wobble pair and the GNRA tetraloop, and our studies reveal novel insights about Mg^{2+} ion binding to these RNA motifs. Interestingly, one Mg^{2+} ion is specifically associated with the S-turn motif, confirming its structural role in the folding of the A730 loop. This Mg^{2+} ion is likely important for formation of the active site and may play an indirect role in catalysis.

INTRODUCTION

Metal cations play several essential functions in the biochemistry of RNA. Although Na^+ , K^+ , Mg^{2+} , and Ca^{2+} ions are the predominant cations in cells, Mg^{2+} is considered to be the most important in RNA folding [1-4]. Mg^{2+} can mediate helical packing [5], promote long-range interactions [2, 6-9], organize multiway junctions [10], stabilize RNA motifs [11-13], and reduce conformational fluctuations [8, 9, 14]. Mg^{2+} is particularly adept at stabilizing tightly packed phosphate oxygens found in several RNA conformations because of its small ionic radius, high charge density, and well-defined octahedral geometry that accommodates six inner-sphere ligands. In addition, Mg^{2+} can either form outer-sphere interactions with ligands through its first-hydration shell or specific inner-sphere interactions by displacing a water molecule and interacting directly with a ligand.

Following the discovery of the catalytic properties of ribozymes, the role of Mg^{2+} ions in catalysis has been intensively investigated. Mg^{2+} ions can play either a direct or an indirect role in the enzymatic mechanism [3, 15-17]. Direct roles have been demonstrated including nucleophile activation [18, 19], coordination of nonbridging oxygens of the scissile phosphate [20], promoting the optimal geometry of the reactive species [21], and stabilization of the leaving group [22]. Potential indirect roles have been proposed involving long-range electrostatic stabilization [16] and/or alteration of pK_a 's to facilitate a general acid-base mechanism [16, 23, 24].

The *Neurospora* Varkud satellite (VS) ribozyme is a member of the small nucleolytic ribozyme family that also includes the hammerhead, hairpin, hepatitis delta virus (HDV), and *glms* ribozymes [25-30]. The VS ribozyme self-cleaves at the phosphodiester bond between G620 and A621 to generate products with 5'-OH and 2'-3'-cyclic phosphate termini. Its secondary structure contains six helical domains: the stem-loop I (SLI) contains the cleavage site and is considered as the substrate domain, whereas stem-loops II–VI (SLII–SLVI) form the catalytic domain (Figure 3.1A; [31]). Substrate recognition is defined by a high-affinity kissing-loop interaction between the SLI and SLV terminal loops (Figure 3.1A) that induces a conformational change within the SLI substrate that is necessary for catalysis [6, 32-34]. The proposed general acid–base mechanism involves the close association of the SLI and SLVI internal loops to form the active site [35-39], where two nucleobases, G₆₃₈ of SLI and A₇₅₆ of the A730 loop, act as the general base and acid in the cleavage reaction [24, 40-46]. Although the secondary and tertiary structures of

the VS ribozyme are unique among the nucleolytic ribozymes, its cleavage mechanism is similar to that of the hairpin ribozyme in which the active site is also formed by the docking of two internal loops, and a nucleobase from each loop participates in the general acid–base mechanism [30].

The VS ribozyme is dependent on the presence of divalent metal ions (Mg^{2+} , Mn^{2+} , or Ca^{2+}) or high concentration of monovalent salt for its cleavage activity [47, 48]. Mg^{2+} ions have been found to be very important for several aspects of folding and function in the VS ribozyme. They are critical for substrate recognition, both for proper formation of the I/V kissing-loop interaction and for the conformational change in SLI [6, 31, 35]. In addition, they are required for folding of the two three-way junctions [49] and for formation of the active site [35, 37]. Metal cations do not play a direct role in the general acid-base mechanism but appear to play an indirect role by altering the $\text{p}K_a$ of the two key nucleobases [24, 50].

Phosphorothioate interference, chemical modifications, nucleotide analogue interference mapping, and NMR studies have been previously conducted in attempts to identify precisely the metal-binding sites in the VS ribozyme [12, 40, 51–54]. In the phosphorothioate interference-rescue experiments, the effect of *pro-R_p* phosphorothioate substitutions in the VS ribozyme were monitored on the Mg^{2+} -dependent cleavage, and several substitutions were shown to reduce the cleavage activity significantly [51]. Interestingly, the phosphorothioate inhibition of cleavage for select residues of the SLI loop (U_{631}), the SLV loop (C_{699}), and the A730 loop (G_{757} and C_{758}) could be rescued by thiophilic Mn^{2+} ions, suggesting that these phosphates are involved in inner-sphere metal coordination [51].

Recently, we performed NMR studies with a small stem-loop containing the A730 internal loop, termed SLVI (Figure 3.1B), in order to gain structural insights into the formation of the VS ribozyme active site [50]. We established that Mg^{2+} ions are required to stabilize the structure of the A730 loop, and we determined the NMR structure of SLVI in the presence of Mg^{2+} ions. The NMR structure reveals a *cis*-WC/WC G-A base pair and an S-turn motif in the A730 loop that protrude the A_{756} into an enlarged minor groove such that it becomes available to participate in catalysis. Here, we extend these investigations by defining the precise locations of associated Mg^{2+} ions by combining Mn^{2+} -induced paramagnetic relaxation enhancement (PRE) with Cd^{2+} -induced chemical-shift perturbation

(CSP) of phosphorothioate RNAs. These data are used to refine the NMR structure of SLVI to include five Mg^{2+} -ion binding sites. Four of the binding sites are associated with known RNA structural motifs, including the GNRA tetraloop and the G-U wobble pair. A comparison with similar structures from the PDB reveals new insights into the role of cation-binding sites in RNA structure. In agreement with the prior phosphorothioate interference-rescue experiments with the VS ribozyme [51], one Mg^{2+} ion is associated with the S-turn motif, revealing its role in the folding of the A730 loop as well as a possible indirect role in catalysis.

EXPERIMENTAL PROCEDURES

Sample Preparation. Unlabeled, ^{15}N -labeled, and $^{13}\text{C}/^{15}\text{N}$ -labeled SLVI RNAs (Figure 3.1B) were synthesized in vitro using T7 RNA polymerase (prepared in-house) and purified as described previously [55]. The SLVI RNAs containing phosphorothioate modifications at nonbridging phosphate oxygens (phosphorothioate RNAs; Figure 3.1B) were purchased as diastereomeric mixtures (Thermo Fisher Scientific, MA). These mixtures were first purified by denaturing polyacrylamide gel electrophoresis (20% acrylamide/bisacrylamide 19:1 and 7 M urea). The R_p and S_p isomers were subsequently separated by reverse-phase HPLC on a μ Bondapak C-18 column (10 μm and 7.8 x 300 mm; Waters, MA) using a multilinear gradient of HPLC-A (0.1 M ammonium acetate) and HPLC-B (0.1 M ammonium acetate and 50% acetonitrile) buffers at 1 mL/min: (a) 0–10% HPLC-B over 180 mL, (b) 10–13% HPLC-B over 120 mL, (c) 13–100% HPLC-B over 5 mL, and (d) 100–0% over 5 mL. The precise determination of configuration was made on the basis of enzymatic digestions with snake venom phosphodiesterase and Nuclease P1 [56].

The purified RNA samples were exchanged in NMR buffer A (10 mM sodium cacodylate pH 6.5, 50 mM KCl, and 0.05 mM NaN_3 in 90% $\text{H}_2\text{O}/10\%$ D_2O) with Amicon Ultra-4 centrifugation filter devices (Millipore, MA). The RNAs were then heated at 95°C for 2 min and then cooled in ice water for 5 min before changing to the final NMR buffer [NMR buffer A with either 5 mM MgCl_2 99.995% (Sigma-Aldrich, MO) or a 5-mM mixture of varying concentrations of MgCl_2 99.995% and CdCl_2 99.999% (Sigma-Aldrich, MO)]. For NMR studies in D_2O , the samples were obtained by multiple cycles of lyophilization and resuspension in 99.996% D_2O .

NMR Spectroscopy. NMR experiments were conducted at either 15 or 25°C on Varian UnityINOVA 500 and 600 MHz spectrometers equipped with a pulse-field gradient unit and an actively—shielded z-gradient probe (either a $^1\text{H}/^{13}\text{C}/^{15}\text{N}$ triple resonance probe or a $^1\text{H}\{^{15}\text{N}-^{31}\text{P}\}$ indirect detection probe). The NMR spectra were processed using the NMRPipe/NMRDraw package [57] and analyzed with NMRView [58]. ^1H , ^{13}C and ^{15}N chemical shifts were referenced to an external standard of 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS) [59] at 0.00 ppm, and ^{31}P chemical shifts were referenced to an external standard of 85% phosphoric acid at 0.00 ppm.

Metal-Ion Binding Studies. Manganese (Mn^{2+}) titrations were performed by adding small volumes of a concentrated solution (0.5 M) of 99.99% MnCl_2 (Sigma-Aldrich, MO) directly to the RNA sample to achieve final concentrations of 5, 10, 20, 40, and 80 μM MnCl_2 . The first Mn^{2+} titration was carried out with a 0.9 mM $^{13}\text{C}/^{15}\text{N}$ -labeled SLVI sample in NMR buffer A with 5 mM MgCl_2 and 100% D_2O . The paramagnetic effect was monitored by collecting 2D ^1H - ^{13}C CT-HSQC [60, 61] and 2D ^1H - ^{13}C HMQC [62] spectra at 25°C. A second Mn^{2+} titration was carried out with a 1.0 mM ^{15}N -labeled SLVI sample in NMR buffer A with 5 mM MgCl_2 and 90%:10% $\text{H}_2\text{O}/\text{D}_2\text{O}$. The paramagnetic effect was monitored by 1D ^1H watergate [63] and imino- and amino-optimized 2D ^1H - ^{15}N HSQC [64] spectra at 15°C. The third Mn^{2+} titration was carried out with a 1.0 mM ^{15}N -labeled SLVI sample in NMR buffer A with 5 mM MgCl_2 and 100% D_2O . The paramagnetic effect was monitored by 2D long-range ^1H - ^{15}N HMQC [65] spectra at 25°C optimized for detection of adenine N1/ N3 atoms and purine N7/N9 atoms ($J_{\text{HN}} = 21$ Hz).

Cadmium (Cd^{2+}) titrations were conducted by exchanging the RNA samples with Amicon filters to NMR buffer A containing varying concentrations of CdCl_2 and MgCl_2 . The total divalent metal-ion concentration was kept constant at 5 mM under all conditions. A first Cd^{2+} titration was conducted with a 0.2 mM $^{13}\text{C}/^{15}\text{N}$ -labeled SLVI RNA sample in 100% D_2O and CdCl_2 concentrations of 0.0625, 0.125, 0.25, 0.5, 1.0, 1.5, 2.5 and 4.0 mM. Chemical shift changes (^1H and ^{13}C) were monitored for each CdCl_2 concentration from 2D ^1H - ^{13}C CT-HSQC [60, 61] spectra recorded at 25°C. Additional Cd^{2+} titrations were conducted for SLVI RNAs with phosphorothioate modifications (0.15-0.4 mM) in 90%:10% $\text{H}_2\text{O}/\text{D}_2\text{O}$ and CdCl_2 concentrations of 0.0625, 0.125, and 0.25 mM. Chemical-

shift changes (^1H and ^{31}P) were monitored for each CdCl_2 concentration from 1D ^1H watergate [63] and 1D ^{31}P spectra recorded at 25°C .

UV Spectroscopy and Determination of T_m Values. Thermal-stability studies of SLVI RNAs were conducted with a Cary 300 UV-vis spectrophotometer equipped with a Peltier temperature-control accessory. All samples contained $5\ \mu\text{M}$ SLVI RNA in NMR buffer A supplemented with divalent metal ions (either $5\ \text{mM}$ MgCl_2 or $4.75\ \text{mM}$ $\text{MgCl}_2/0.25\ \text{mM}$ CdCl_2). Control samples were also prepared that contained only the selected buffers. Samples were heated from 25 to 98°C at a rate of $1^\circ\text{C}/\text{min}$, and absorbance data were collected at $260\ \text{nm}$ after each 1°C temperature increment. Subtraction of the control samples from the RNA samples was processed automatically using a dual-beam mode. Melting temperatures (T_m) were determined from the second derivative of the absorbance versus temperature curve, as previously described [12]. Reported T_m are averaged values derived from two or more individual UV-denaturation profiles.

Native Gel Electrophoresis. Nondenaturing polyacrylamide gel electrophoresis was carried out on 7.5% polyacrylamide gels, acrylamide/bisacrylamide (37.5:1) in Tris-borate buffer ($50\ \text{mM}$ Tris-borate, pH 8.0) supplemented with $20\ \text{mM}$ magnesium acetate. The gels were prerun for $30\ \text{min}$ at $200\ \text{V}$, loaded with RNA samples ($2\ \mu\text{g}$ of RNA at a concentration of $20\ \mu\text{M}$), and then run for $2\ \text{h}$ at $250\ \text{V}$ and 4°C . The gels were stained with Stains All (Sigma-Aldrich, MO).

Structural Modeling of $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ -Binding Sites. Three-dimensional structures of SLVI bound to $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ complexes were determined by complementing previously-defined restraints [50] with distance restraints to position the five $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ complexes. These new restraints were derived from Mn^{2+} -induced paramagnetic relaxation enhancement (PRE) using the ratio of signal intensity (I_0/I_{Mn}) determined from 2D spectra collected at $0\ \mu\text{M}$ MnCl_2 (I_0) and $10\ \mu\text{M}$ MnCl_2 (I_{Mn}). For each signal in these spectra with $I_0/I_{\text{Mn}} \geq 8$ -fold and $4 \leq I_0/I_{\text{Mn}} < 8$ -fold, distance restraints were set to 1.8 - $7.0\ \text{\AA}$ and 1.8 - $8.0\ \text{\AA}$, respectively, between a defined Mg^{2+} center and the nuclei giving rise to the observed signal (Table S3.1 of Supporting Information). When the identity of the Mg^{2+} center could not be defined, ambiguous metal restraints were defined with bounds of 1.8 - $10.0\ \text{\AA}$ (Table

S3.1 of Supporting Information). Restraints were also derived from Cd²⁺-induced ³¹P CSP of the sulfur-bound phosphorus from phosphorothioate SLVI RNAs [$\Delta_p = (\delta_{Mg} - \delta_{Cd})$] calculated from 1D ³¹P spectra collected in 5 mM MgCl₂ (δ_{Mg}) and 0.25 mM CdCl₂/4.75 mM MgCl₂ (δ_{Cd}). Phosphorothioates with $\Delta_p \geq 1.0$ ppm were considered as inner-sphere ligands, and a distance constraint of 1.8-2.2 Å was defined between a Mg²⁺ center and the modified nonbridging phosphate oxygen. Each phosphorothioate with $0.2 < \Delta_p < 1.0$ ppm was considered as an outer-sphere ligand, and a distance restraint of 1.8-7.0 Å was defined between a Mg²⁺ center and the modified nonbridging phosphate oxygen. On the basis of a clustering analysis of the nuclei that were most affected by Mn²⁺-induced PRE ($I_0/I_{Mn} \geq 8$ fold) and Cd²⁺-induced CSP of phosphorothioates ($\Delta_p > 0.2$ ppm), we inferred five distinct Mg²⁺-binding sites for SLVI, termed sites 1–5, including two pentahydrated Mg²⁺-ion complexes [Mg(H₂O)₅²⁺; sites 1 and 3] and three fully hydrated Mg²⁺-ion complexes [Mg(H₂O)₆²⁺; sites 2, 4, and 5].

Three-dimensional structures of SLVI with bound Mg(H₂O)_n²⁺ complexes were calculated by restraining molecular dynamics and simulated annealing with X-PLOR-NIH version 2.1.9 [66] by adapting the two-stage protocol previously used for structure determination of free SLVI [50]. At stage one, an initial set of structures was calculated from RNA structures with randomized backbone angles to which Mg(H₂O)_n²⁺ coordinates and parameters derived from HIC-Up were added [67]. Several rounds of calculations were performed at this stage to allow stepwise incorporation of restraints to metal-ion complexes. At the end of stage one, 50 structures were obtained that satisfy all distance and dihedral experimental restraints (no distance violation of > 0.2 Å and no torsion angle violation of $> 5^\circ$), including all of those involving the five metal complexes. At stage two, these structures were refined with the same set of restraints but were supplemented with RDC restraints [50]. A final set of 500 structures was calculated, from which the 20 lowest-energy structures that satisfied the experimental restraints (no distance violation > 0.2 Å, no torsion angle violation $> 5^\circ$, and no RDC violation > 5 Hz) were selected for analysis. These 20 lowest-energy structures were used to calculate an average structure that was minimized against NOE and dihedral restraints. All structures were visualized and analyzed with PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC.

Pattern Search of RNA Metal-Binding Motifs in the Protein Data Bank (PDB).

Metal-binding motifs similar to those found in SLVI were searched in all available X-ray, cryo-EM, and NMR structures of the PDB using WebFR3D [68]. The search patterns were defined according to their sequential and structural context in SLVI using the FR3D symbolic search algorithm. For each search, the results were extracted from the FR3D html file (raw hits) and filtered to remove the PDB files that did not contain metal ions. For each resulting hit (hit with a cation), a heavy-atom superposition of the nucleotides defining the search pattern was obtained between the PDB coordinates of the hit and the SLVI average structure, and this superposition was used to extract a pairwise heavy-atom rmsd. In cases where residues used for the superposition differ, the base of these residues was omitted for the superposition and rmsd calculation. The superposition was also used to extract the distance between the divalent metal ions of the two structures. For each FR3D search, those structures with metal distances smaller than 4.5 Å were termed positive hits and kept for further analysis.

RESULTS

Probing Divalent Metal-Ion-Binding Sites in SLVI through Mn²⁺-Induced Paramagnetic Relaxation Enhancement (PRE). To identify metal-ion-binding sites in the SLVI RNA (Figure 3.1B), we performed several Mn²⁺ titrations and monitored them by NMR spectroscopy. Mn²⁺ is a paramagnetic metal ion that specifically enhances the relaxation of nuclei located within a short distance range (~10 Å) [69]. The PRE is proportional to r^{-6} and can be used to derive distance (r) restraints between the Mn²⁺ and the observed nuclei [69]. Thus, Mn²⁺-induced PRE represents a common approach for identifying divalent metal-ion-binding sites in RNA [12, 70-76].

Most of the previously assigned ¹H, ¹³C, and ¹⁵N signals in SLVI were probed using 2D ¹H-¹³C and ¹H-¹⁵N correlation experiments recorded during the titration of 0-80 μM MnCl₂ to an SLVI RNA folded in 5 mM MgCl₂. The high concentration of Mg²⁺ is necessary for proper RNA folding, whereas a ~1000-fold lower concentration of Mn²⁺ is sufficient for specific binding of the paramagnetic ion in the fast-exchange regime [70, 71, 77]. At 10 μM MnCl₂, the relaxation of several nuclei in SLVI is specifically affected, as shown for aromatic signals of the 2D ¹H-¹³C CT-HSQC spectrum (Figure 3.2A,B) that either become significantly less intense (G₁₃) or completely disappear (G₁, G₃, G₆, G₉, C₁₂, G₁₉, G₂₂, and

G₂₃). For each well-resolved signal in this and other 2D spectra, the signal intensity was measured in the absence (I_0) and presence of 10 μM Mn^{2+} (I_{Mn}) to calculate a disappearance ratio (I_0/I_{Mn}). Interestingly, the signals that undergo the strongest decrease in intensity ($I_0/I_{\text{Mn}} \geq 8$) correspond to ^1H , ^{13}C , and ^{15}N nuclei that are dispersed throughout the RNA (Figure 3.2C and Table S3.1 of the Supporting Information). Thus, the results from Mn^{2+} -induced PRE are compatible with several divalent-metal-binding sites in SLVI. The precise localization of these binding sites is, however, not possible at this stage because specific phosphates involved in metal binding could not be identified from Mn^{2+} -induced PRE as a result of the severe ^{31}P spectral overlap and the lack of specific ^{31}P assignment.

Identification of Phosphate-Metal Interactions by Cd^{2+} -Induced ^{31}P Chemical-Shift Perturbation (CSP) of Phosphorothioate RNAs. A phosphorothioate RNA contains a modification that substitutes either the *pro*-R_p or the *pro*-S_p nonbonded phosphate oxygen (*pro*-R_p/S_p OP) by a sulfur atom and significantly alters the chemical shift of the associated ^{31}P signal such that it can be easily identified in a 1D ^{31}P NMR spectrum [78-80]. Because phosphorothioates prefer thiophilic metal ions, like Cd^{2+} , over Mg^{2+} ions, formation of phosphorothioate-metal complexes can be monitored by recording 1D ^{31}P NMR spectra during Cd^{2+} titration [79, 80]. Chemical-shift perturbations (CSP) of the sulfur-bound phosphorus have been associated with coordination of Cd^{2+} ions, with large upfield shifts being characteristic of inner-sphere coordination [79, 80].

To investigate the contribution of phosphates to metal coordination, several SLVI RNAs containing single phosphorothioate modifications were prepared (Figure 3.1B). Phosphorothioate modifications of C₇, A₈, G₉, and C₁₀ are of particular interest to examine metal binding in the S-turn of SLVI. In addition, modification of A₁₅ provides a positive control because an inner-sphere metal contact has been previously reported at this position of a GAAA tetraloop [79], and modification of C₂₄ provides a negative control because this residue is not predicted to interact with a metal ion. The phosphorothioate RNAs were prepared by chemical synthesis as mixtures of *pro*-R_p and *pro*-S_p isomers, and reverse-phase HPLC was used to separate the individual isomers. SLVI RNAs with a 5'-phosphorothioate at a specific residue N are denoted as either N-R_p or N-S_p. With the exception of the C24-R_p/S_p mixture, all of the individual isomers were separated from their mixtures in high purity.

The Cd²⁺-titration conditions were carefully defined to prevent Cd²⁺-induced structural changes in SLVI and/or RNA aggregation that may result from an excess of Cd²⁺ ions [81, 82]. We first collected 2D ¹H-¹³C CT-HSQC spectra of a 0.2 mM ¹³C/¹⁵N-labeled SLVI RNA sample in the presence of increasing Cd²⁺ concentrations. Only small changes in chemical shifts were observed between 0 and 0.25 mM of Cd²⁺, suggesting that the SLVI structure remains essentially intact. However, significant CSP are observed at Cd²⁺ concentrations ≥ 1.0 mM (Figure S3.1 of the Supporting Information), which indicates that the SLVI structure is significantly modified by high Cd²⁺:RNA ratios. In addition, high Cd²⁺ concentrations cause smearing of SLVI on native gels (at 4 mM Cd²⁺; not shown) and a change in the shape of the UV melting curves (at 1.5 mM Cd²⁺; not shown), in agreement with aggregation of SLVI under these conditions.

To preserve the structural integrity of SLVI, Cd²⁺ titrations of phosphorothioate RNAs were thus performed at low Cd²⁺:RNA ratios. These titrations were monitored by 1D ³¹P spectra (Figure 3.3), and CSP of the sulfur-bound phosphorus was determined for each phosphorothioate RNA using $\Delta_p = (\delta_{Mg} - \delta_{Cd})$, where δ_{Mg} = chemical shift (δ) at 5 mM MgCl₂ and δ_{Cd} = δ at 4.75 mM MgCl₂:0.25 mM CdCl₂ (Table 3.1). A large CSP ($\Delta_p \geq 1.0$ ppm) was detected for the A15-S_p phosphorothioate ($\Delta_p = 1.43$ ppm; Figure 3.3A and Table 3.1), as was previously observed for an equivalent phosphorothioate of a GAAA tetraloop involved in an inner-sphere metal coordination [79]. Similarly, the large CSP associated with the G9-R_p phosphorothioate ($\Delta_p = 1.62$ ppm; 3. 3.3B and Table 3.1) provides evidence for inner-sphere metal coordination at the nonbridging G₉ *pro*-R_p phosphate oxygen. Smaller but still significant CSP ($0.2 < \Delta_p < 1.0$ ppm) for seven other phosphorothioates (C7-R_p/S_p, A8-R_p/S_p, G9-S_p, C10-R_p, and A15-R_p; Figure 3.3, Figure S3.2 of the Supporting Information and Table 3.1) are compatible with outer-sphere metal coordination. Finally, the C10-S_p and C24-R_p/S_p phosphorothioates display small CSP ($\Delta_p < 0.2$ ppm; Figure S3.2 of the Supporting Information and Table 3.1), as is expected for those not involved in divalent metal coordination.

Previous reports have indicated that phosphorothioate modifications may affect RNA structure by inducing distortions in the sugar-phosphate backbone [83, 84], thus control experiments were performed to verify that they have a minimal effect on the structure of SLVI. First, 1D ¹H imino NMR spectra were collected as part of the Cd²⁺ titration of phosphorothioate SLVI RNAs, and these spectra are all very similar to that of the

unmodified SLVI (Figure S3.3 of the Supporting Information), suggesting that the phosphorothioate modifications do not significantly affect the global structure of SLVI. In addition, melting temperatures (T_m) were determined for each phosphorothioate RNA by UV-vis spectroscopy (Table 3.2). Generally, the observed T_m values of the phosphorothioate SLVI RNAs (Table 3.2) are very similar to that of the unmodified SLVI ($74.6 \pm 0.1^\circ\text{C}$), further supporting that these modifications do not alter the structural integrity of the RNA. However, significant differences in T_m values are noted for two phosphorothioate RNAs. The T_m value for the A15-R_p phosphorothioate ($82.5 \pm 0.1^\circ\text{C}$) is 8°C higher than that of the unmodified SLVI RNA, similar to what was previously observed for an equivalent residue in a GAAA hairpin loop [83]. In contrast, the T_m value for the C7-R_p phosphorothioate RNA ($64.6 \pm 1.1^\circ\text{C}$) is 10°C lower than that of the unmodified SLVI RNA, indicating that this particular modification destabilizes SLVI. Given the position of the C₇ 5'-PO₄⁻ in the S-turn [50], the destabilizing effect of the C7-R_p phosphorothioate modification could be due to a perturbed divalent metal coordination involving the modified atom. To investigate this possibility, T_m values were obtained in the presence of 0.25 mM Cd²⁺/4.75 mM Mg²⁺ (Table 3.2). Interestingly, although this Cd²⁺-containing buffer does not affect the stability of the unmodified SLVI RNA ($T_m = 74.6 \pm 0.2$), it restores the stability of the C7-R_p phosphorothioate RNA ($T_m = 73.5 \pm 0.2$), implying that the modified position is involved in metal coordination. Consequently, these controls support the validity of the Cd²⁺-titration results, at least in terms of inferring specific phosphate-metal interactions in SLVI.

Structural Modeling of SLVI with Mg(H₂O)_n²⁺ Complexes. Three-dimensional NMR structures of SLVI bound to Mg(H₂O)_n²⁺ complexes (SLVI^{Mg}) were determined as previously reported for free SLVI (SLVI^{free}; [50]) with the addition of Mg²⁺-RNA distance restraints derived from the Mn²⁺-induced PRE (Table S3.1 of the Supporting Information) and the Cd²⁺-induced CSP of phosphorothioate RNAs (Table 3.1). Five Mg(H₂O)_n²⁺ binding sites (sites 1-5) were defined by circumscribing into separate clusters atoms that were most affected by the Mn²⁺ and Cd²⁺ titrations. Metal complexes at sites 2, 4, and 5 were modeled as hexahydrated Mg²⁺ complexes [Mg(H₂O)₆²⁺], whereas those at sites 1 and 3 were modeled as pentahydrated Mg²⁺ complexes [Mg(H₂O)₅²⁺], given the evidence for inner-sphere metal coordination of the 5'-PO₄⁻ of A₁₅ and G₉ (Table 3.1). The resulting 20

lowest-energy structures of SLVI^{Mg} (Table 3.3) are compatible with all of the experimental restraints, indicating that the new Mg²⁺-RNA restraints are fully consistent with distance, dihedral, and RDC restraints defined for structure determination of the SLVI^{free} RNA [50]. The NMR structure of SLVI^{Mg} is very well-defined with an overall heavy-atom rmsd of 0.33 ± 0.14 Å (Table 3.3). Thus, the ensemble of SLVI^{Mg} structures is well-represented by the minimized average structure (Figure 3.4). By comparison, the NMR structure of SLVI^{free} was previously determined with an overall heavy atom rmsd of 0.67 ± 0.17 Å [50], indicating that the new Mg²⁺-RNA restraints help to define better the SLVI structure.

The structure of SLVI^{Mg} reveals five well-defined Mg²⁺-binding sites. The three outer-sphere metal-binding sites, sites 2, 4, and 5, are all located in the major groove of SLVI (Figure 3.5 and Table 3.4). Site 2 is located in the hairpin stem near the C₁₀ 5'-PO₄⁻ of the S-turn (Figure 3.5A), where G₁₇ O6, U₁₈ O4, and G₁₉ O6 act as potential outer-sphere ligands and C₁₀ 5'-PO₄⁻ likely contributes to electrostatic stabilization. Site 4 is associated with the U₅-G₂₂ wobble base within the 5'-3' stem (Figure 3.5B), with U₅ O4, G₂₂ O6/N7/5'-*pro*-R_p OP and G₂₃ O6/N7 acting as potential outer-sphere ligands and G₂₂ 5'-PO₄⁻ ensuring electrostatic stabilization. Interestingly, the Mg(H₂O)₆²⁺ complex at site 4 also interacts with the base of C₂₁, forming a cation- π interaction. The Mg(H₂O)₆²⁺ complex at site 5 is found at the extremity of the 5'-3' stem (Figure 3.5C), close enough to G₁ O5'/N7/O6, A₂ N7/5'-*pro*-R_p OP, and G₃ O6/N7 for outer-sphere coordination and to the 5'-PO₄⁻ of A₂ and G₃ for electrostatic stabilization. For site 5, the position of the Mg(H₂O)₆²⁺ complex is not as well-defined as for sites 2 and 4 by the NMR data (Figure 3.5), indicating that metal coordination may be undergoing dynamic exchange of ligands.

The two inner-sphere metal-binding sites, sites 1 and 3 (Figure 3.6), are formed by unique structural elements within SLVI that contain electronegative pockets rich in phosphate oxygens, with one of the phosphates involved in inner-sphere metal coordination (A₁₅ 5'-*pro*-S_p OP at site 1 and G₉ 5'-*pro*-R_p OP at site 3; Table 3.4). Site 1 within the G₁₃A₁₄A₁₅A₁₆ tetraloop also utilizes three potential outer-sphere ligands identified as A₁₅ O5'/5'-*pro*-R_p OP and A₁₆ 5'-*pro*-R_p OP (Table 3.4). Remarkably, the Mg(H₂O)₅²⁺ at site 1 is positioned on one face of the G₁₃ base, forming a cation- π interaction (Figure 3.6A) [85]. Site 3 also involves several potential outer-sphere ligands identified as C₇ O3'/O5'/5'-*pro*-R_p OP, A₈ O5'/5'-*pro*-R_p OP, and G₉ O5' (Table 3.4). Moreover, site 3 is unique in that the 5'-

PO_4^- of C_7 , A_8 , and G_9 within the S-turn form an electronegative cluster resembling a three-prong mount that holds the $\text{Mg}(\text{H}_2\text{O})_5^{2+}$ complex (Figure 3.6B).

Search for Metal-Binding Motifs in the PDB. To complement our NMR study of metal-binding sites in SLVI, we used the WebFR3D symbolic search algorithm [68] to find similar metal-binding sites in the PDB. The search patterns correspond to sites 1-5 (Figure 3.4), as summarized in Table S3.2 of the Supporting Information. Positive hits from this search fit the search pattern and also contain a metal ion that is within 4.5 Å of that found for the same motif in SLVI.

For the GNRA motif in SLVI (site 1), the search pattern was defined by the consecutive CGNRAG sequence (N is an undefined nucleotide and R a purine) with the first and last nucleotides forming a Watson-Crick base pair. The search resulted in three positive hits with pairwise heavy-atom rmsd values and metal distances to the SLVI average structure between 1.0-1.3 Å and 0.7-4.1 Å, respectively (Figure 3.7A). All of these motifs adopt a GNRA tetraloop fold with a metal ion in the minor groove interacting with the 5'- PO_4^- corresponding to that of A_{15} in SLVI, although this interaction does not involve inner-sphere coordination as observed for SLVI. Remarkably, in all cases, the metal ion forms a cation- π interaction with the conserved G nucleobase of the motif.

For site 2, the search pattern corresponds to a 3-bp helical domain that matches the exact sequence of the first three base pairs of the hairpin stem ($\text{C}_{10}\text{-C}_{12}/\text{G}_{17}\text{-G}_{19}$). The search resulted in eight positive hits, with pairwise heavy atom rmsd values and metal distances to the SLVI average structure between 0.6-1.1 Å and 1.6-3.4 Å, respectively (Figure 3.7B). From the superposition of the eight positive hits to the average SLVI structure, it is clear that the associated metal ions always lie in the major groove, although they occupy several different locations within this ensemble of structures.

For the S-turn motif (site 3), two search patterns were defined from the A730 loop sequence (Table S3.2 of the Supporting Information); however, only one raw hit (i.e., without metal) was obtained corresponding to the previously determined SLVI structure [50].

For the G-U base pair in SLVI (site 4), the search pattern was defined as a helical domain with three Watson-Crick base pairs corresponding to nucleotides $\text{C}_4\text{-G}_6/\text{C}_{21}\text{-G}_{23}$ in SLVI. This search resulted in three positive hits with pairwise heavy-atom rmsd values and

metal distances to the SLVI average structure between 1.0-1.9 Å and 1.5-2.0 Å, respectively (Figure 3.7C). In one case (PDB code 1FJG), the divalent metal ion forms a cation- π interaction with a cytidine [85].

For the purine stretch at the 5'-end of the helix (site 5), the search pattern corresponds to a 3-bp helical domain that matches the exact sequence of the three first base pairs in SLVI (G₁-G₃/C₂₄-C₂₆). The search resulted in 16 positive hits with pairwise heavy-atom rmsd values and metal distances to the SLVI average structure between 0.4-1.1 Å and 1.7-4.0 Å, respectively (Figure 3.7D). From the superposition of the 16 positive hits to the average SLVI structure, it is clear that the associated metal ions lie in the major groove, although they occupy a variety of different locations in these structures.

DISCUSSION

Powerful NMR Approach for Localization of Divalent Metal-Binding Sites in RNA.

The present work highlights the strength of combining Mn²⁺-induced PRE experiments and Cd²⁺-induced CSP of phosphorothioate RNAs to localize divalent-cation-binding sites in RNA. Mn²⁺-induced PRE experiments have been widely used to probe RNA structures and to define metal-binding sites [9, 12, 71, 74, 76] and can provide a large number of RNA-metal restraints, especially for ¹H, ¹³C, and ¹⁵N atoms of ¹³C/¹⁵N-labeled RNAs. However, Cd²⁺-induced CSP of phosphorothioate RNAs identifies specific metal-phosphate interactions and provides information on the coordination state of the associated metal ions [79, 80, 86-88]. Thus, these two approaches are complementary, and together they provide a large number of nonredundant experimental restraints to define precisely metal-ion binding in the SLVI RNA. An important consideration in using this dual approach is ensuring that the different metal ions and chemically-modified RNAs employed do not affect the structural integrity of the RNA. Importantly, the two metal ions used, Mn²⁺ and Cd²⁺, both support the catalytic activity of the VS ribozyme and thus likely the correct folding at the active site [47, 48, 51]. In addition, experimental controls were performed to ensure reliable measurement of distance restraints between the RNA and the associated divalent metals. Overall, the distance restraints derived from these data help to define better the structure of SLVI and to allow the localization of five well-defined divalent metal-binding sites in the major groove of the RNA. One of the metal-binding sites is specific to the S-turn of the A730 loop of SLVI, whereas the other four are commonly found in other

RNAs. Thus, the NMR structure of SLVI with bound divalent cations provides valuable insights into both our general understanding of RNA-metal interactions and the cation requirement for formation of the VS ribozyme active site.

New Insights about Known Cation-Binding Sites in RNA. According to our FR3D search, four of the five metal-binding sites in SLVI have been previously observed in similar RNA structures. Two of these sites (sites 2 and 5) bind a hexahydrated Mg^{2+} ion within the major groove of A-form helical stems formed by standard Watson-Crick base pairs. The site located at the 5'-end of SLVI (site 5) is formed by three consecutive purines. This site offers a favorable environment for cations because of the presence of several electronegative groups associated with the 5'-GAG-3' sequence. This metal-binding site is not as well-defined by the NMR data as other sites in SLVI, suggesting that more than one coordination mode are possible and that diffuse binding may occur along the purine stretch. Using FR3D, we identified 16 similar metal-binding sites in purine patches with the same sequence and structure context. The location of the metal is also not well-defined within these metal-binding sites, further supporting that the major groove of consecutive purines represents a common metal-binding site in double-stranded RNA that allows diffuse binding over the Hoogsteen edge of these purines [71, 74, 89]. In contrast, the binding of a Mg^{2+} ion within the hairpin stem at site 2 is better defined, with the base keto-oxygens of the 5'-GUG-3' sequence and the *pro*- R_p OP of the C_{10} 5'- PO_4^- as potential outer-sphere ligands. Using FR3D, we identified eight similar sites in the PDB, and the location of the metal ion is more clearly defined within these available structures than for the 5'-GAG-3' site, indicating fewer modes of metal coordination at the 5'-GUG-3' site. These results also point out sequence-dependent variations in cation binding within the major groove of A-form RNA helices.

G-U wobble pairs are well-known metal-binding sites that form depending on the sequence and structure context [5, 71, 73, 90-93]. In SLVI, a metal ion is associated with the G_{22} - U_5 base pair in the 5'-3' stem near the A730 loop (site 4). The Hoogsteen edges of G_{22} and G_{23} are mainly involved in the interaction with the hexahydrated Mg^{2+} ion, in agreement with previous studies. In the structure of SLVI, a cation- π interaction is also observed between the Mg^{2+} ion and the C_{21} nucleobase [85]. To our knowledge, this interaction has not been previously associated with a G-U wobble pair. Using FR3D, we

found three similar structures in the PDB. Interestingly, in one of these three structures (PDB entry 1FJG; [94]) the bound metal ion also forms a cation- π interaction with a cytidine. In this case, however, the metal ion is located further from the center of the base and has a larger angle with the centroid of the nucleobase, suggesting a weaker cation- π interaction. In SLVI, the nearby A730 loop may render C₂₁ more accessible for a cation- π interaction with the Mg²⁺ ion. However, given that the G-U pair is found in a different context in SLVI versus the full-length VS ribozyme, it is not clear if Mg²⁺ binding at this site is relevant to the activity of the ribozyme.

The GNRA tetraloop has been extensively characterized as a metal-binding site [72, 79, 95-97]. In the context of the present study, the GAAA loop of SLVI was used as a positive control because an inner-sphere divalent-metal interaction involving the 5'-PO₄⁻ of A₁₅ was previously identified in an equivalent position of a GAAA hairpin on the basis Cd²⁺-titration of phosphorothioates [79]. These NMR results were essentially reproduced here for the GNRA loop of SLVI. In addition, we identified a cation- π interaction in the SLVI structure between the Mg²⁺ ion and the pyrimidine ring of the first G nucleobase of the GNRA loop. The importance of this cation- π interaction in the GNRA fold has not been previously established, at least to our knowledge; therefore, it is interesting to examine cation binding to other GNRA loop structures in the PDB. Using FR3D, we identified three other similar GNRA loop structures associated with a cation. Remarkably, although the precise location of the cation differs between these structures, it always interacts with the R 5'-PO₄⁻ and forms a cation- π interaction with the conserved G nucleobase.

Cation-Binding Site in the A730 Loop Is Important for the Activity of the VS Ribozyme. From previous NMR studies, we established that the presence of Mg²⁺ ions is required for the formation of a stable A730 loop. It is now clear that binding of a specific Mg²⁺ ion at the S-turn (site 3) is central to the folding of the A730 loop [50]. This S-turn brings three phosphate groups in proximity to one another [5'-PO₄⁻ of C₇ (C₇₅₅), A₈ (A₇₅₆), and G₉ (G₇₅₇)] with particularly short OP-OP distances between G₉ and C₇ and between G₉ and A₈ (4.1 and 5.1 Å in the average structure, respectively). This tight packing of repulsive electronegative charges is dependent on the multidentate chelation of a Mg²⁺ ion, which involves an inner-shell ligand [5'-*pro*-R_p OP of G₉ (G₇₅₇)] and several potential outer-sphere ligands. Although our FR3D search indicates that there is no precedent for a S-turn

structure in a similar loop context, the S-turn motif is found in other structural contexts, both with or without a divalent cation [74, 98-102]. With the A730 loop being relatively small, this clearly imposes topological and electrostatic constraints on the phosphate backbone at the S-turn that prevents stable folding in the absence of metal ions.

In previous work with the VS ribozyme, phosphorothioate interference and manganese-rescue experiments suggested that the *pro-R_p* OP of G₇₅₇ (G₉) and C₇₅₈ (C₁₀) are involved in direct metal binding and that the phosphate group of C₇₅₅ (C₇) makes an outer-sphere contact to a metal ion or participates in hydrogen bonding [40, 51]. The NMR studies of SLVI reveal an outer-sphere metal contact for C₇₅₅ (C₇) *pro-R_p* OP and an inner-sphere metal contact for G₇₅₇ (G₉) *pro-R_p* OP, in agreement with the biochemical studies [51]. For C₇₅₈ (C₁₀), the NMR studies reveal an outer-sphere contact to a Mg²⁺ ion but not the predicted inner-sphere contact [51]. Within SLVI, the C₇₅₈ 5'-PO₄⁻ may partially contribute to stabilization of the S-turn via an outer-sphere interaction with a Mg²⁺ ion in the hairpin stem. However, within the context of the VS ribozyme, this metal interaction may be modified because of nearby elements of secondary and tertiary structures, including formation of the active site between stem-loops I and VI. In summary, our NMR studies of SLVI generally agree with phosphorothioate interference of VS ribozyme cleavage, which indicates that the Mg²⁺-dependent S-turn in the A730 loop is essential for catalysis by the VS ribozyme [50].

The Mg²⁺ ion at the S-turn plays a structural role in catalysis by the VS ribozyme by allowing the folding of the S-turn in the A730 loop and enabling A₇₅₆ to be extruded in the minor groove to participate as a general acid in catalysis [50]. In the hairpin ribozyme, which catalyzes phosphodiester-bond cleavage by a similar mechanism [30], the proposed general acid in the reaction (A₃₈) also protrudes from an S-turn structure that is stabilized by at least two divalent cations [103, 104]. Although these cations do not play a direct role in the cleavage mechanism, they may contribute either through electrostatic stabilization or by modulating the p*K_a* of functional residues [16, 23]. In the VS ribozyme, it is also unlikely that the Mg²⁺-ion at the S-turn plays a direct role in catalysis, given that its binding site is remote from the scissile phosphate in a model of the VS ribozyme active site [50]. Nevertheless, given that this Mg²⁺ ion is not too distant from the catalytic site, we can not rule out the possibility that it contributes to the chemistry of the reaction. It is known that the p*K_a* value of the A₇₅₆ and the G₆₃₈ nucleobases are shifted toward neutrality to facilitate

the cleavage reaction [40, 41, 43-46], and that the pK_a of these nucleobases is modulated by the identity of the metal cation [24]. Thus, the Mg^{2+} ion in the A730 loop clearly plays a structural role in defining the VS ribozyme active site, but it also likely contributes indirectly to the general acid-base chemistry.

ASSOCIATED CONTENT

Supporting Information

Effect of Cd^{2+} on the 2D 1H - ^{13}C HSQC spectrum of SLVI RNA; effect of Cd^{2+} on the 1D ^{31}P NMR spectrum and 1H imino NMR spectrum of SLVI RNAs containing a single phosphorothioate modification; distance restraints derived from Mn^{2+} -induced PRE and results from the WebFR3D motif search. This material is available free of charge via the Internet at <http://pubs.acs.org>. This work was supported by a Canadian Institutes of Health Research (CIHR) grant (MOP-86502) as well as graduate scholarships (E.B.) from CIHR and the Université de Montréal. P. L. holds a Canada Research Chair in Structural Biology and Engineering of RNA.

Accession Codes

The NMR chemical shifts, structural restraints, and structural coordinates of SLVI with bound $Mg(H_2O)_n^{2+}$ complexes have been deposited through the RCSB Protein Data Bank and given BMRB entry code 19692 and PDB entry code 2MIS.

Author Information

Notes

The authors declare no competing financial interest.

Acknowledgments

We are very grateful to Luke Ward and Victoria De Rose for useful discussions and to Nicolas Girard for help with motif search in WebFR3D. We also thank Patricia Bouchard and James G. Omichinski for their critical reading of the manuscript.

Abbreviations Used

2D, two-dimensional; 3D, three-dimensional; bp, base pair; Cd²⁺, cadmium; CSP, chemical-shift perturbation; CT, constant time; δ , chemical shift; Δ_p , ³¹P CSP of the phosphorothioate; HSQC, heteronuclear single quantum coherence; Mg²⁺, magnesium; Mn²⁺, manganese; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; OP, phosphate oxygen; PO₄⁻, phosphate; PRE, paramagnetic relaxation enhancement; PDB, protein data bank; rmsd, root-mean-square deviation; T_m , melting temperature.

REFERENCES

1. Wang, Y.X., S. Huang, and D.E. Draper, *Structure of a U.U pair within a conserved ribosomal RNA hairpin*. Nucleic Acids Res., 1996. **24**(14): p. 2666-72.
2. Nissen, P., et al., *RNA tertiary interactions in the large ribosomal subunit: the A-minor motif*. Proc. Natl. Acad. Sci. U. S. A., 2001. **98**(9): p. 4899-903.
3. Molinaro, M. and I. Tinoco, Jr., *Use of ultra stable UNCG tetraloop hairpins to fold RNA structures: thermodynamic and spectroscopic applications*. Nucleic Acids Res, 1995. **23**(15): p. 3056-63.
4. DeRose, V.J., *Characterization of nucleic-acid-metal ion binding by spectroscopic techniques*, in *Nucleic Acid-Metal Ion Interactions*, N.V. Hud, Editor. 2008, RSC Publishing: Cambridge. p. 447.
5. Cate, J.H. and J.A. Doudna, *Metal-binding sites in the major groove of a large ribozyme domain*. Structure, 1996. **4**: p. 1221-1229.
6. Rastogi, T., et al., *A long-range pseudoknot is required for activity of the Neurospora VS ribozyme*. EMBO J., 1996. **15**(11): p. 2820-2825.

7. Klein, D.J., P.B. Moore, and T.A. Steitz, *The contribution of metal ions to the structural stability of the large ribosomal subunit*. RNA, 2004. **10**(9): p. 1366-1379.
8. Conn, G.L., et al., *Crystal structure of a conserved ribosomal protein-RNA complex*. Science, 1999. **284**(5417): p. 1171-4.
9. Noeske, J., H. Schwalbe, and J. Wohnert, *Metal-ion binding and metal-ion induced folding of the adenine-sensing riboswitch aptamer domain*. Nucleic Acids Res, 2007. **35**(15): p. 5262-73.
10. Bouchard, P. and P. Legault, *Structural insights into substrate recognition by the Neurospora Varkud satellite ribozyme: importance of U-turns at the kissing-loop junction*. Biochemistry, 2014. **53**(1): p. 258-69.
11. Cabello-Villegas, J., I. Tworowska, and E.P. Nikonowicz, *Metal ion stabilization of the U-turn of the A₃₇ N⁶-dimethylallyl-modified anticodon stem-loop of Escherichia coli tRNA^{Phe}*. Biochemistry, 2004. **43**: p. 55-66.
12. Campbell, D.O., et al., *NMR structure of Varkud satellite ribozyme stem-loop V in the presence of magnesium ions and localization of metal-binding sites*. Biochemistry, 2006. **45**(35): p. 10591-605.
13. Sashital, D.G., et al., *U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction*. Nat. Struct. Mol. Biol., 2004. **11**(12): p. 1237-42.
14. Guthrie, C. and B. Patterson, *Spliceosomal snRNAs*. Annu. Rev. Genet., 1988. **22**: p. 387-419.
15. Fedor, M.J., *The role of metal ions in RNA catalysis*. Current Opin. in Struct. Biol., 2002. **12**: p. 289-295.
16. Sigel, R.K. and A.M. Pyle, *Alternative roles for metal ions in enzyme catalysis and the implications for ribozyme chemistry*. Chem Rev, 2007. **107**(1): p. 97-113.
17. Butcher, S.E. and A.M. Pyle, *The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks*. Acc. Chem. Res., 2011. **44**(12): p. 1302-11.
18. Wimberly, B.T., et al., *A detailed view of a ribosomal active site: the structure of the L11-RNA complex*. Cell, 1999. **97**(4): p. 491-502.
19. Leontis, N.B., W. Kwok, and J.S. Newman, *Stability and structure of three-way DNA junctions containing unpaired nucleotides*. Nucleic Acids Res., 1991. **19**(4): p. 759-66.

20. Wilson, T.J., et al., *Folding and catalysis of the hairpin ribozyme*. Biochem. Soc. Trans., 2005. **33**(Pt 3): p. 461-5.
21. Roth, A., et al., *A widespread self-cleaving ribozyme class is revealed by bioinformatics*. Nat. Chem. Biol., 2014. **10**(1): p. 56-60.
22. Lilley, D.M., et al., *A nomenclature of junctions and branchpoints in nucleic acids*. Nucleic Acids Res., 1995. **23**(17): p. 3363-4.
23. de la Pena, M., D. Dufour, and J. Gallego, *Three-way RNA junctions with remote tertiary contacts: a recurrent and highly versatile fold*. RNA, 2009. **15**(11): p. 1949-64.
24. Smith, M.D., et al., *The ionic environment determines ribozyme cleavage rate by modulation of nucleobase pKa*. RNA, 2008. **14**(9): p. 1942-9.
25. Saville, B.J. and R.A. Collins, *A site-specific self-cleavage reaction performed by a novel RNA in Neurospora mitochondria*. Cell, 1990. **61**: p. 685-696.
26. Collins, R.A., *The Neurospora Varkud satellite ribozyme*. Bioch. Soc. Trans. Rev., 2002. **30**: p. 1122-1126.
27. Lilley, D.M., *The Varkud satellite ribozyme*. RNA, 2004. **10**: p. 151-158.
28. Lilley, D.M.J., *The hairpin and Varkud satellite ribozymes*, in *Ribozymes and RNA catalysis*, D.M.J. Lilley and F. Eckstein, Editors. 2008, Royal Society of Chemistry: Cambridge. p. 66-91.
29. Cochrane, J.C. and S.A. Strobel, *Catalytic strategies of self-cleaving ribozymes*. Acc. Chem. Res., 2008. **41**(8): p. 1027-35.
30. Wilson, T.J. and D.M. Lilley, *Do the hairpin and VS ribozymes share a common catalytic mechanism based on general acid-base catalysis? A critical assessment of available experimental data*. RNA, 2011. **17**(2): p. 213-21.
31. Beattie, T.L., J.E. Olive, and R.A. Collins, *A secondary-structure model for the self-cleaving region of Neurospora VS RNA*. Proc. Natl. Acad. Sci. U. S. A., 1995. **92**: p. 4686-4690.
32. Andersen, A. and R.A. Collins, *Rearrangement of a stable RNA secondary structure during VS ribozyme catalysis*. Mol. Cell, 2000. **5**: p. 469-478.
33. Andersen, A.A. and R.A. Collins, *Intramolecular secondary structure rearrangement by the kissing interaction of the Neurospora VS ribozyme*. Proc. Natl. Acad. Sci U. S. A., 2001. **98**(14): p. 7730-7735.

34. Klostermeier, D. and D.P. Millar, *Tertiary structure stability of the hairpin ribozyme in its natural and minimal forms: different energetic contributions from a ribose zipper motif*. *Biochemistry*, 2001. **40**(37): p. 11211-8.
35. Hiley, S.L. and R.A. Collins, *Rapid formation of a solvent-inaccessible core in the Neurospora Varkud satellite ribozyme*. *EMBO J.*, 2001. **20**(19): p. 5461-5469.
36. Lafontaine, D.A., et al., *The A730 loop is an important component of the active site of the VS ribozyme*. *J. Mol. Biol.*, 2001. **312**: p. 663-674.
37. Hiley, S.L., et al., *4-thio-U cross-linking identifies the active site of the VS ribozyme*. *EMBO J.*, 2002. **21**(17): p. 4691-4698.
38. Lafontaine, D.A., et al., *Functional group requirements in the probable active site of the VS ribozyme*. *J. Mol. Biol.*, 2002. **323**: p. 23-34.
39. Sood, V.D. and R.A. Collins, *Identification of the catalytic subdomain of the VS ribozyme and evidence for remarkable sequence tolerance in the active site loop*. *J. Mol. Biol.*, 2002. **320**: p. 443-454.
40. Jones, F.D. and S.A. Strobel, *Ionization of a critical adenosine residue in the Neurospora Varkud satellite ribozyme active site*. *Biochemistry*, 2003. **42**: p. 4265-4276.
41. McLeod, A.C. and D.M. Lilley, *Efficient, pH-dependent RNA ligation by the VS ribozyme in trans*. *Biochemistry*, 2004. **43**(4): p. 1118-25.
42. Zhao, Z.Y., et al., *Nucleobase participation in ribozyme catalysis*. *J. Am. Chem. Soc.*, 2005. **127**(14): p. 5026-7.
43. Smith, M.D. and R.A. Collins, *Evidence for proton transfer in the rate-limiting step of a fast-cleaving Varkud satellite ribozyme*. *Proc. Natl. Acad. Sci. U. S. A.*, 2007. **104**(14): p. 5818-23.
44. Wilson, T.J., A.C. McLeod, and D.M. Lilley, *A guanine nucleobase important for catalysis by the VS ribozyme*. *EMBO J.*, 2007. **26**(10): p. 2489-500.
45. Jaikaran, D., et al., *An important role of G638 in the cis-cleavage reaction of the Neurospora VS ribozyme revealed by a novel nucleotide analog incorporation method*. *RNA*, 2008. **14**(5): p. 938-49.
46. Wilson, T.J., et al., *Nucleobase-mediated general acid-base catalysis in the Varkud satellite ribozyme*. *Proc. Natl. Acad. Sci. U. S. A.*, 2010. **107**(26): p. 11751-6.

47. Collins, R.A. and J.E. Olive, *Reaction conditions and kinetics of self-cleavage of a ribozyme derived from Neurospora VS RNA*. *Biochemistry*, 1993. **32**: p. 2795-2799.
48. Murray, J.B., et al., *The hammerhead, hairpin and VS ribozymes are catalytically proficient in monovalent cations alone*. *Chem. Biol.*, 1998. **5**: p. 587-595.
49. Lafontaine, D.A., D.G. Norman, and D.M. Lilley, *Structure, folding and activity of the VS ribozyme: importance of the 2-3-6 helical junction*. *EMBO J.*, 2001. **20**(6): p. 1415-1424.
50. Desjardins, G., et al., *NMR structure of the A730 loop of the Neurospora VS ribozyme: insights into the formation of the active site*. *Nucleic Acids Res.*, 2011. **39**(10): p. 4427-37.
51. Sood, V.D., T.L. Beattie, and R.A. Collins, *Identification of phosphate groups involved in metal binding and tertiary interactions in the core of the Neurospora VS ribozyme*. *J. Mol. Biol.*, 1998. **282**: p. 741-750.
52. Tzokov, S.B., I.A. Murray, and J.A. Grasby, *The role of magnesium ions and 2'-hydroxyl groups in the VS ribozyme-substrate interaction*. *J. Mol. Biol.*, 2002. **324**: p. 215-226.
53. Hoffmann, B., et al., *NMR structure of the active conformation of the Varkud satellite ribozyme cleavage site*. *Proc. Natl. Acad. Sci. U. S. A.*, 2003. **100**(12): p. 7003-7008.
54. Kovacheva, Y.S., et al., *The role of phosphate groups in the VS ribozyme-substrate interaction*. *Nucleic Acids Res.*, 2004. **32**(21): p. 6240-50.
55. Campbell, D.O. and P. Legault, *NMR structure of the Varkud satellite ribozyme stem-loop V RNA and magnesium-ion binding from chemical-shift mapping*. *Biochemistry*, 2005. **44**: p. 4157-4170.
56. Slim, G. and M.J. Gait, *Configurationaly defined phosphorothioate-containing oligoribonucleotides in the study of the mechanism of cleavage of hammerhead ribozymes*. *Nucleic Acids Res.*, 1991. **19**(6): p. 1183-8.
57. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes*. *J. Biomol. NMR*, 1995. **6**: p. 277-293.
58. Johnson, B.A. and R.A. Blevins, *NMRView: a computer program for the visualization and analysis of NMR data*. *J. Biomol. NMR*, 1994. **4**: p. 603-614.

59. Wishart, D.S., et al., *¹H, ¹³C, ¹⁵N chemical shift referencing in biomolecular NMR*. J. Biomol. NMR, 1995. **6**: p. 135-140.
60. Vuister, G.W. and A. Bax, *Resolution enhancement and spectral editing of uniformly ¹³C-enriched proteins by homonuclear broadband ¹³C decoupling*. J. Magn. Reson., 1992. **98**: p. 428-435.
61. Santoro, J. and G.C. King, *A constant-time 2D overbroadband experiment for inverse correlation of isotopically enriched species*. J. Magn. Reson., 1992. **97**: p. 202-207.
62. Ikura, M., et al., *Three-dimensional NOESY-HMQC spectroscopy of a ¹³C-labeled protein*. J. Magn. Reson., 1990. **86**: p. 204-209.
63. Piotto, M., V. Saudek, and V. Skleňár, *Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions*. J. Biomol. NMR, 1992. **2**: p. 661-665.
64. Kay, L.E., P. Keifer, and T. Saarinen, *Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity*. J. Am. Chem. Soc., 1992. **114**: p. 10663-10665.
65. Skleňár, V., et al., *Correlation of nucleotide base and sugar protons in a ¹⁵N-labeled HIV-1 RNA oligonucleotide by ¹H-¹⁵N HSQC experiments*. J. Biomol. NMR, 1994. **4**: p. 117-122.
66. Schwieters, C.D., et al., *The Xplor-NIH NMR molecular structure determination package*. J. Magn. Reson., 2003. **160**: p. 66-74.
67. Kleywegt, G.T. and T.A. Jones, *Databases in protein crystallography*. Acta Crystallogr. D, 1998. **54**: p. 1119-11131.
68. Jucker, F.M. and A. Pardi, *Solution structure of the CUUG hairpin loop: a novel RNA tetraloop motif*. Biochemistry, 1995. **34**(44): p. 14416-27.
69. Bertini, I. and C. Lychinat, *NMR of paramagnetic molecules in biological systems*. 1986, Menlo Park, CA: Benjamin/Cummings.
70. Ott, G., L. Arnold, and S. Limmer, *Proton NMR studies of manganese ion binding to tRNA-derived acceptor arm duplexes*. Nucleic Acids Res, 1993. **21**(25): p. 5859-64.

71. Allain, F.H. and G. Varani, *Divalent metal ion binding to a conserved wobble pair defining the upstream site of cleavage of group I self-splicing introns*. Nucleic Acids Res, 1995. **23**(3): p. 341-50.
72. Hansen, M.R., et al., *Identification and characterization of a novel high affinity metal-binding site in the hammerhead ribozyme*. RNA, 1999. **5**: p. 1099-1104.
73. Colmenarejo, G. and I. Tinoco, Jr., *Structure and thermodynamics of metal binding in the P5 helix of a group I intron ribozyme*. J. Mol. Biol., 1999. **290**: p. 119-135.
74. Butcher, S.E., F.H.-T. Allain, and J. Feigon, *Determination of metal ion binding sites within the hairpin ribozyme domains by NMR*. Biochemistry, 2000. **39**: p. 2174-2182.
75. Zuleeg, T., et al., *NMR spectroscopic evidence for Mn(2+)(Mg(2+)) binding to a precursor-tRNA microhelix near the potential RNase P cleavage site*. J Mol Biol, 2001. **305**(2): p. 181-9.
76. Davis, J.H., et al., *Role of metal ions in the tetraloop-receptor complex as analyzed by NMR*. RNA, 2007. **13**(1): p. 76-86.
77. Granot, J., J. Feigon, and D.R. Kearns, *Interactions of DNA with divalent ions. I. ³¹P-NMR studies*. Biopolymers, 1982. **21**: p. 181-201.
78. Baumruk, V., et al., *Comparison between CUUG and UUCG tetraloops: thermodynamic stability and structural features analyzed by UV absorption and vibrational spectroscopy*. Nucleic Acids Res., 2001. **29**(19): p. 4089-96.
79. Maderia, M., T.E. Horton, and V.J. DeRose, *Metal interactions with a GAAA RNA tetraloop characterized by ³¹P NMR and phosphorothioate substitution*. Biochemistry, 2000. **39**: p. 8193-8200.
80. Maderia, M., L.M. Hunsicker, and V.J. DeRose, *Metal-phosphate interactions in the hammerhead ribozyme observed by ³¹P NMR and phosphorothioate substitutions*. Biochemistry, 2000. **39**(40): p. 12113-20.
81. Duguid, J.G. and V.A. Bloomfield, *Aggregation of melted DNA by divalent metal ion-mediated cross-linking*. Biophys J, 1995. **69**(6): p. 2642-8.
82. Duguid, J.G., et al., *Raman spectroscopy of DNA-metal complexes. II. The thermal denaturation of DNA in the presence of Sr²⁺, Ba²⁺, Mg²⁺, Ca²⁺, Mn²⁺, Co²⁺, Ni²⁺, and Cd²⁺*. Biophys J, 1995. **69**(6): p. 2623-41.

83. Horton, T.E., M. Maderia, and V.J. DeRose, *Impact of phosphorothioate substitutions on the thermodynamic stability of an RNA GAAA tetraloop: an unexpected stabilization*. *Biochemistry*, 2000. **39**(28): p. 8201-7.
84. Smith, J.S. and E.P. Nikonowicz, *Phosphorothioate substitution can substantially alter RNA conformation*. *Biochemistry*, 2000. **39**(19): p. 5642-52.
85. McFail-Isom, L., X. Shui, and L.D. Williams, *Divalent cations stabilize unstacked conformations of DNA and RNA by interacting with base pi systems*. *Biochemistry*, 1998. **37**(49): p. 17105-11.
86. Huppler, A., et al., *Metal binding and base ionization in the U6 RNA intramolecular stem-loop structure*. *Nature Struct Biol*, 2002. **9**(6): p. 431-435.
87. Osborne, E.M., et al., *The identity of the nucleophile substitution may influence metal interactions with the cleavage site of the minimal hammerhead ribozyme*. *Biochemistry*, 2009. **48**(44): p. 10654-64.
88. Ward, W.L. and V.J. Derose, *Ground-state coordination of a catalytic metal to the scissile phosphate of a tertiary-stabilized Hammerhead ribozyme*. *RNA*, 2012. **18**(1): p. 16-23.
89. Strobel, S.A. and T.R. Cech, *Tertiary interactions with the internal guide sequence mediate docking of the P1 helix into the catalytic core of the Tetrahymena ribozyme*. *Biochemistry*, 1993. **32**(49): p. 13593-604.
90. Kieft, J.S. and I.J. Tinoco, *Solution structure of a metal-binding site in the major groove of RNA complexed with cobalt(III) hexamine*. *Current Biology*, 1997. **5**: p. 713-721.
91. Keel, A.Y., et al., *A general strategy to solve the phase problem in RNA crystallography*. *Structure*, 2007. **15**(7): p. 761-72.
92. Al-Sogair, F.M., et al., *Probing the metal-ion-binding strength of the hydroxyl group*. *Chem. Rev.*, 2011. **111**(8): p. 4964-5003.
93. Chen, J.H., et al., *A 1.9 Å crystal structure of the HDV ribozyme precleavage suggests both Lewis acid and general acid mechanisms contribute to phosphodiester cleavage*. *Biochemistry*, 2010. **49**(31): p. 6508-18.
94. Carter, A.P., et al., *Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics*. *Nature*, 2000. **407**: p. 340-348.

95. Rüdisser, S. and I.J. Tinoco *Solution structure of cobalt(III) hexamine complexed to the GAAA tetraloop, and metal-ion binding to G-A mismatches*. J. Mol. Biol., 2000. **295**: p. 1211-1223.
96. Salvail-Lacoste, A., et al., *Affinity purification of T7 RNA transcripts with homogeneous ends using ARiBo and CRISPR tags*. RNA, 2013. **19**(7): p. 1003-14.
97. Woese, C.R., S. Winker, and R.R. Gutell, *Architecture of ribosomal RNA: constraints on the sequence of "tetra-loops"*. Proc. Natl. Acad. Sci. U. S. A., 1990. **87**(21): p. 8467-71.
98. Szewczak, A.A., et al., *The conformation of the sarcin/ricin loop from 28S ribosomal RNA*. Proc. Natl. Acad. Sci. U. S. A., 1993. **90**(20): p. 9581-5.
99. Wimberly, B., G. Varani, and I. Tinoco, Jr., *The conformation of loop E of eukaryotic 5S ribosomal RNA*. Biochemistry, 1993. **32**(4): p. 1078-87.
100. Correll, C.C., et al., *Crystal structure of the ribosomal RNA domain essential for binding elongation factors*. Proc. Natl. Acad. Sci. U. S. A., 1998. **95**(23): p. 13436-41.
101. Butcher, S.E., F.H. Allain, and J. Feigon, *Solution structure of the loop B domain from the hairpin ribozyme*. Nat. Struct. Biol., 1999. **6**(3): p. 212-6.
102. Zimmermann, G.R., et al., *Molecular interactions and metal binding in the theophylline-binding core of an RNA aptamer*. RNA, 2000. **6**: p. 659-667.
103. Rupert, P.B. and A.R. Ferré-D'Amaré, *Crystal structure of a haripin ribozyme-inhibitor complex with implications for catalysis*. Nature, 2001. **410**: p. 780-786.
104. Alam, S., et al., *Conformational heterogeneity at position U37 of an all-RNA hairpin ribozyme with implications for metal binding and the catalytic structure of the S-turn*. Biochemistry, 2005. **44**(44): p. 14396-408.

TABLES

Table 3.1. ^{31}P NMR CSP (Δ_p) following Cd^{2+} titrations of phosphorothioate SLVI RNAs.

phosphorothioate RNA	Δ_p (ppm) ^a
C7-R _p	0.31
C7-S _p	0.35
A8-R _p	0.41
A8-S _p	0.23
G9-R _p	1.62
G9-S _p	0.39
C10-R _p	0.32
C10-S _p	0.09
A15-R _p	0.35
A15-S _p	1.43
C24-R _p /S _p ^b	0.05/0.01

^a $\Delta_p = (\delta_{\text{Mg}} - \delta_{\text{Cd}})$, where $\delta_{\text{Mg}} = \delta$ at 5 mM MgCl_2 and $\delta_{\text{Cd}} = \delta$ at 4.75 mM $\text{MgCl}_2/0.25$ mM CdCl_2 .

^bThis sample is a racemic mixture of R_p and S_p phosphorothioates.

Table 3.2. T_m values for the unmodified and phosphorothioate SLVI RNAs.

RNA	T_m (°C)
<i>5 mM MgCl₂</i>	
SLVI	74.6 ± 0.1
C7-R _p	64.6 ± 1.1
C7-S _p	74.4 ± 0.1
A8-R _p	74.3 ± 0.2
A8-S _p	74.5 ± 0.1
G9-R _p	74.2 ± 0.1
G9-S _p	73.8 ± 0.1
C10-R _p	73.7 ± 0.3
C10-S _p	73.4 ± 0.1
A15-R _p	82.5 ± 0.1
A15-S _p	75.0 ± 0.1
C24-R _p /S _p ^a	74.0 ± 0.2
<i>4.75 mM MgCl₂ + 0.25 mM CdCl₂</i>	
SLVI	74.6 ± 0.2
C7-R _p	73.5 ± 0.2

^aThis SLVI RNA is a racemic mixture of R_p and S_p phosphorothioates.

Table 3.3. Structural statistics of the SLVI RNA with $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ complexes.

distance restraints	1086
number of NOE-derived distance restraints	965
inter-nucleotide	591
intra-nucleotide	355
ambiguous	19
hydrogen-bond restraints	52
number of Mg^{2+} -RNA distance restraints	
from Mn^{2+} -induced PRE	60
from Cd^{2+} -induced CSP (Δ_p)	9
dihedral angle restraints	88
residual dipolar coupling restraints	30
total number of restraints	1204
rmsd from experimental restraints	
NOE (Å) (none > 0.2)	0.0123 ± 0.004
dihedral (°) (none > 5°)	0.11 ± 0.01
residual dipolar couplings (none > 5 Hz)	0.23 ± 0.01
dmsd from idealized geometry	
bonds (Å)	0.00547 ± 0.00002
angles (°)	1.1741 ± 0.0007
impropers (°)	0.441 ± 0.004
heavy-atom rmsd (Å)^a	
overall (residues 2-25)	0.33 ± 0.14
5'-3' stem (residues 2-5 and 22-25)	0.18 ± 0.08
hairpin stem (residues 11-18)	0.15 ± 0.07
A730 loop (residues 6-10 and 19-21)	0.23 ± 0.08
metal site 1 (residues 13-16)	0.087 ± 0.025
metal site 2 (residues 10-12 and 17-19)	0.14 ± 0.06
metal site 3 (residues 7-9)	0.23 ± 0.06
metal site 4 (residues 3-6 and 21-24)	0.20 ± 0.11
metal site 5 (residues 1-4 and 23-26)	0.17 ± 0.06

^aHeavy-atom rmsd to the minimized averaged structure.

Table 3.4. Description of $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ -binding sites in SLVI.

site	inner-sphere ligands ^a	outer-sphere ligands ^b	electrostatic interactions ^c	cation- π interactions ^d
site 1 (n=5)	A_{15} 5'- <i>pro</i> -S _p OP	A_{15} O5' A_{15} 5'- <i>pro</i> -R _p OP A_{16} 5'- <i>pro</i> -R _p OP	A_{15} 5'-PO ₄ ⁻ A_{16} 5'-PO ₄ ⁻	G ₁₃
site 2 (n=6)		G ₁₇ O6 U ₁₈ O4 G ₁₉ O6	C ₁₀ 5'-PO ₄ ⁻	
site 3 (n=5)	G ₉ 5'- <i>pro</i> -R _p OP	C ₇ O3' C ₇ O5' C ₇ 5'- <i>pro</i> -R _p OP A ₈ O5' A ₈ 5'- <i>pro</i> -R _p OP G ₉ O5'	C ₇ 5'-PO ₄ ⁻ A ₈ 5'-PO ₄ ⁻ G ₉ 5'-PO ₄ ⁻	
site 4 (n=6)		U ₅ O4 G ₂₂ O6 G ₂₂ N7 G ₂₂ 5'- <i>pro</i> -R _p OP G ₂₃ O6 G ₂₃ N7	G ₂₂ 5'-PO ₄ ⁻	C ₂₁
site 5 (n=6)		G ₁ O5' G ₁ N7 G ₁ O6 A ₂ N7 A ₂ 5'- <i>pro</i> -R _p OP G ₃ O6 G ₃ N7	A ₂ 5'-PO ₄ ⁻ G ₃ 5'-PO ₄ ⁻	

^aInner-sphere ligands were determined from Cd²⁺-induced CSP ($\Delta_p > 1.0$ ppm).

^bOuter-sphere ligands were determined by selecting all hydrogen-bond acceptors within 4.0 Å of a $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ oxygen that were in a favorable orientation for outer-sphere interaction in at least 5 of the 20 lowest-energy structures.

^cElectrostatic interactions were determined by selecting 5'-PO₄⁻ within 8.0 Å of a Mg²⁺ ion.

^dCation- π interactions were determined for Mg²⁺ ions within 5.2 Å and at an angle smaller than 52° from the ring centroid [85].

FIGURES

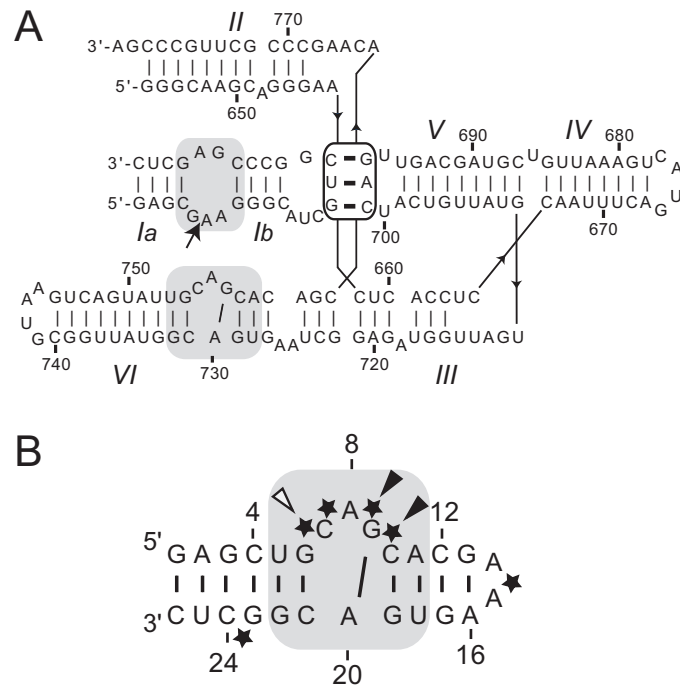


Figure 3.1. Primary and secondary structures of the *Neurospora* VS ribozyme and the SLVI RNA. (A) *trans*-Cleaving VS ribozyme (containing wild-type residues 648-777) with an SLVI substrate. The cleavage site is shown by an arrow. Both the cleavage site internal loop and the active site internal loop are shaded in gray and the I/V kissing-loop interaction is boxed. (B) SLVI RNA, which includes the A730 loop of the VS ribozyme (gray shading). Phosphate groups that display inhibitory effects on cleavage when substituted by phosphorothioates are indicated by an arrow, and the arrow is filled for those cases where the activity can be rescued by addition of Mn^{2+} ions.[51] For this study, single phosphorothioate SLVI RNAs were prepared that contain either an S_p or R_p phosphorothioate at one of the phosphates indicated by black stars.

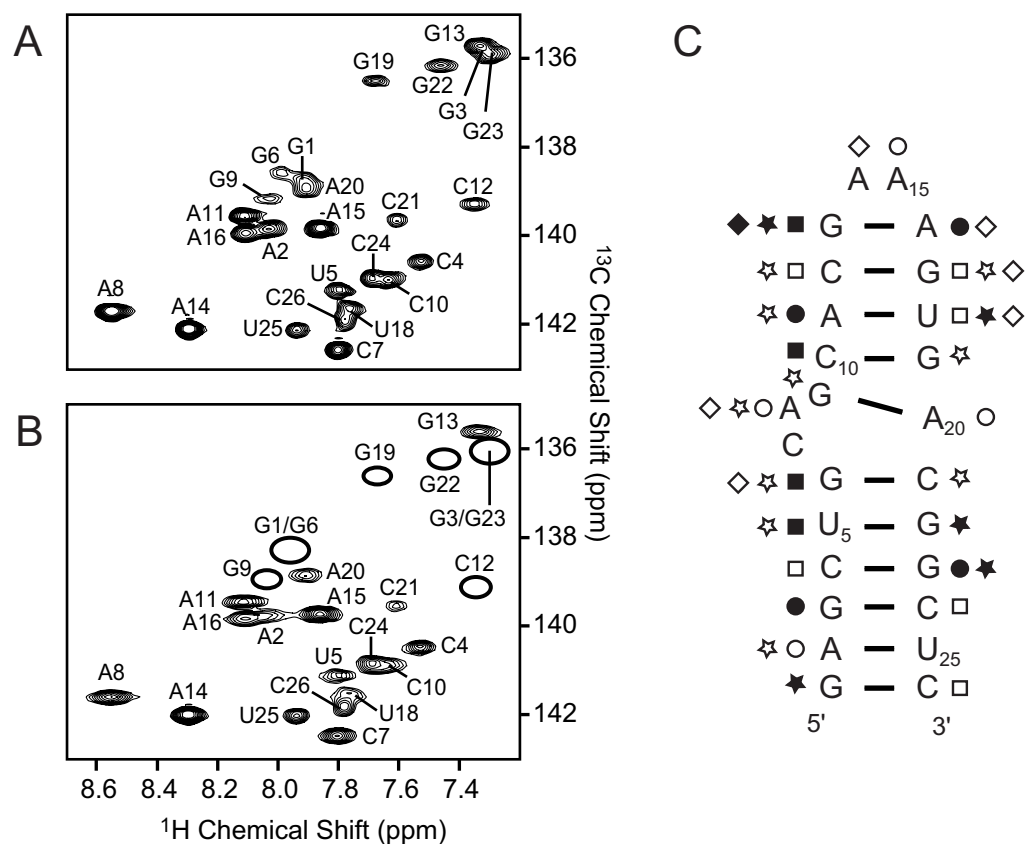


Figure 3.2. Probing divalent-metal-ion-binding sites in SLVI through Mn^{2+} -induced PRE. The effect of Mn^{2+} ions on the relaxation of SLVI nuclei is shown for aromatic (C6-H6 and C8-H8) signals by comparing the 2D 1H - ^{13}C HSQC spectra of SLVI RNA collected in (A) the absence and (B) the presence of 10 μM $MnCl_2$. Peaks that disappeared below the noise level are denoted with empty ovals. (C) Summary of the effect of Mn^{2+} ions on the relaxation of SLVI nuclei mapped on the secondary structure of SLVI. Filled and open symbols represent large ($I_0/I_{Mn} \geq 8$ -fold) and medium ($4 \leq I_0/I_{Mn} < 8$ -fold) PRE. Different symbols are used to distinguish PRE of imino and amino N-H signals (squares), N7, N9, and N3 signals (circles), aromatic C-H signals (stars), and ribose C-H signals (diamonds).

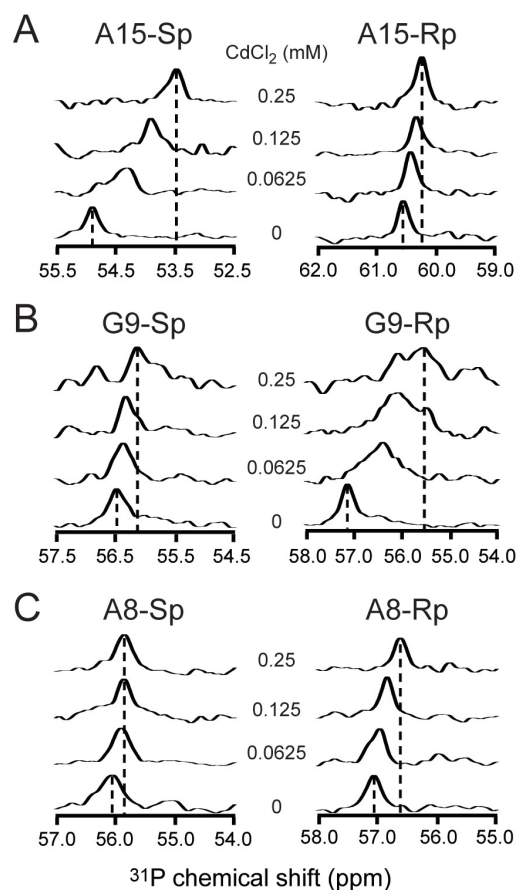


Figure 3.3. Effect of Cd^{2+} ions on the 1D ^{31}P NMR spectrum of selected SLVI RNAs containing single phosphorothioate modifications. The 1D ^{31}P spectra of (A) A15-S_p and A15-R_p, (B) G9-S_p and G9-R_p, and (C) A8-S_p and A8-R_p were collected in buffer containing $\text{CdCl}_2/\text{MgCl}_2$ concentrations of 0.0:5.0, 0.0625:4.9375, 0.125:4.875, and 0.25:4.75 mM. Only the most downfield region of the spectra containing the signal from the phosphorothioate is shown.

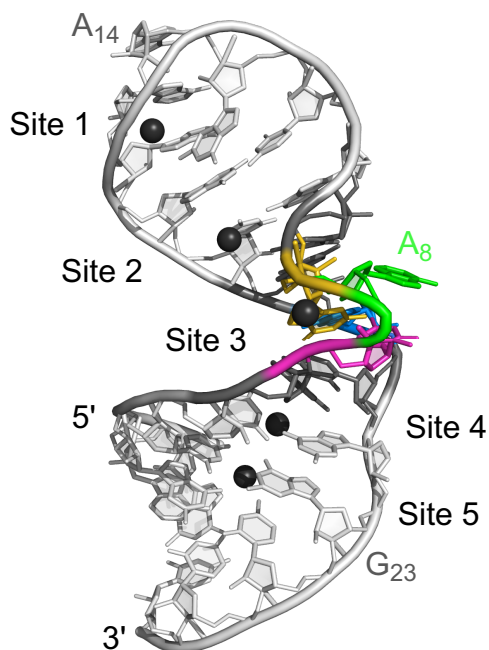


Figure 3.4. Lowest-energy structure of the SLVI RNA determined in the presence of $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ complexes. The five Mg^{2+} ions associated with SLVI are shown (black), but their bound water molecules were omitted for clarity. For SLVI, only the heavy atoms are shown, and the ribbon replacing the phosphorus and the nonbonded oxygen atoms is used to show the backbone. SLVI nucleotides are color-coded: the loop-closing base pairs ($\text{G}_6\text{-C}_{21}$ and $\text{C}_{10}\text{-G}_{19}$) are dark gray, C_7 (C_{755}) is magenta, A_8 (A_{756}) is green, G_9 (G_{757}) is gold, and A_{20} (A_{730}) is blue.

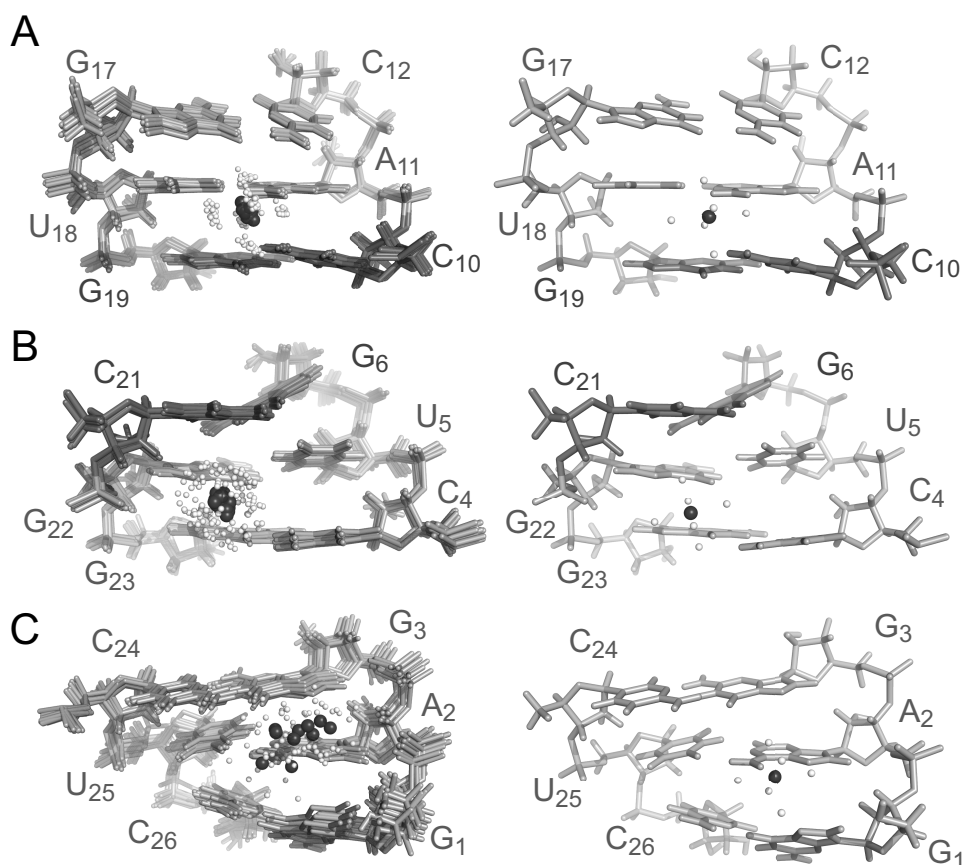


Figure 3.5. $\text{Mg}(\text{H}_2\text{O})_6^{2+}$ -binding sites associated with sites 2, 4 and 5. (A) C_{10} $5'\text{-PO}_4^-$ (site 2), (B) $\text{G}_{22}\text{-U}_5$ wobble (site 4), and (C) $5'$ -end (site 5) of SLVI RNA. In panels A-C a superposition of the 10 lowest-energy structures (left panel) and the lowest-energy structure (right panel) are shown. The Mg^{2+} ions are shown in black, with their bound water molecules in white.

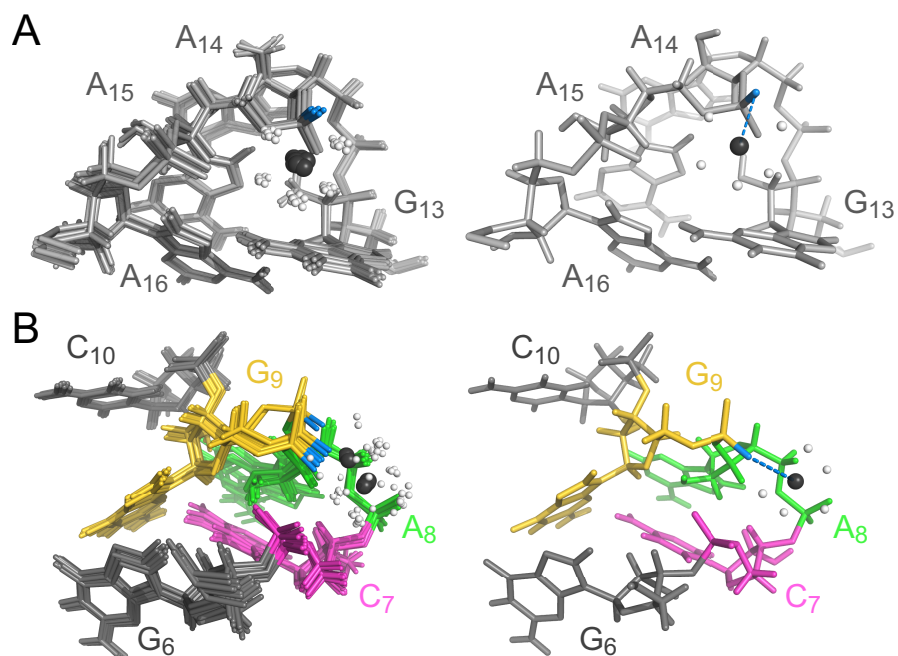


Figure 3.6. $\text{Mg}(\text{H}_2\text{O})_5^{2+}$ -binding sites associated with sites 1 and 3. (A) GAAA hairpin loop (site 1) and (B) S-turn (site 3) of SLVI RNA. In both panels A and B, a superposition of the 10 lowest-energy structures (left panel) and the lowest-energy structure (right panel) are shown. Blue dotted lines represent inner-sphere metal binding involving the A_{15} *pro-S_p* OP in panel A and the G_9 *pro-R_p* OP in panel B. The Mg^{2+} ions are shown in black, with their bound water molecules in white.

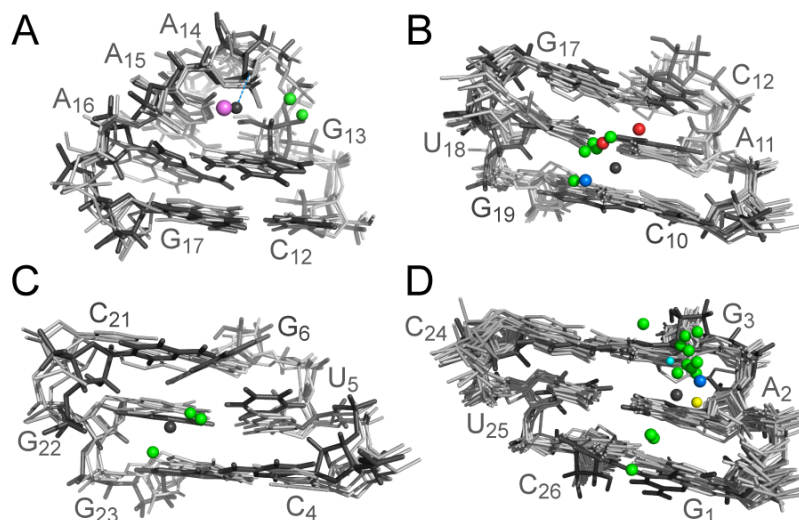


Figure 3.7. Superposition of metal-binding sites in SLVI with PDB structures containing similar metal-binding sites identified using WebFR3D. The metal-binding sites in SLVI are associated with the (A) GNRA hairpin loop (site 1), (B) C₁₀ 5'-PO₄⁻ (site 2), (C) G₂₂-U₅ wobble (site 4), and (D) 5'-end (site 5). The structure and Mg²⁺ ions of SLVI are shown in black, whereas the other structures (PDB entries in Table S3.2 of Supporting Information) are shown in gray with their associated metal ion color coded according to the type of metal (Mg²⁺, green; Na⁺, purple; Co³⁺, red; Ca²⁺, blue; Ni²⁺, yellow; and Ir³⁺; light blue).

SUPPORTING INFORMATION

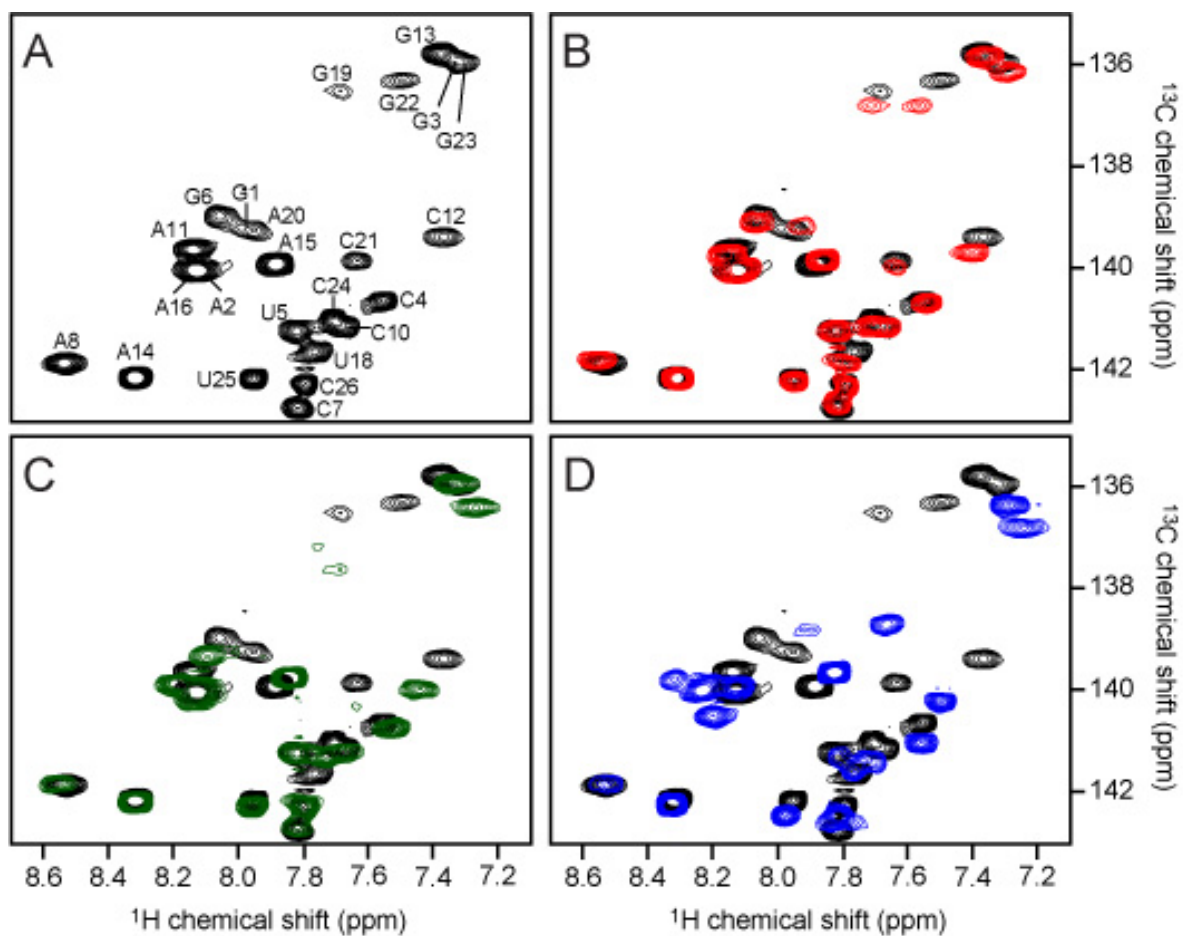


Figure S3.1. Effect of Cd^{2+} ions on SLVI RNA. Comparison of the H6/H8-C6/C8 region of 2D ^1H - ^{13}C CT-HSQC spectra of SLVI at selected points of the Cd^{2+} titration. (A)-(D) Spectra are shown for CdCl_2 : MgCl_2 concentrations of (A)-(D) 0.0:5.0 mM (black), (B) 0.25 mM:4.75 mM (red), (C) 1 mM:4 mM (green), and (D) 4 mM:1 mM (purple).

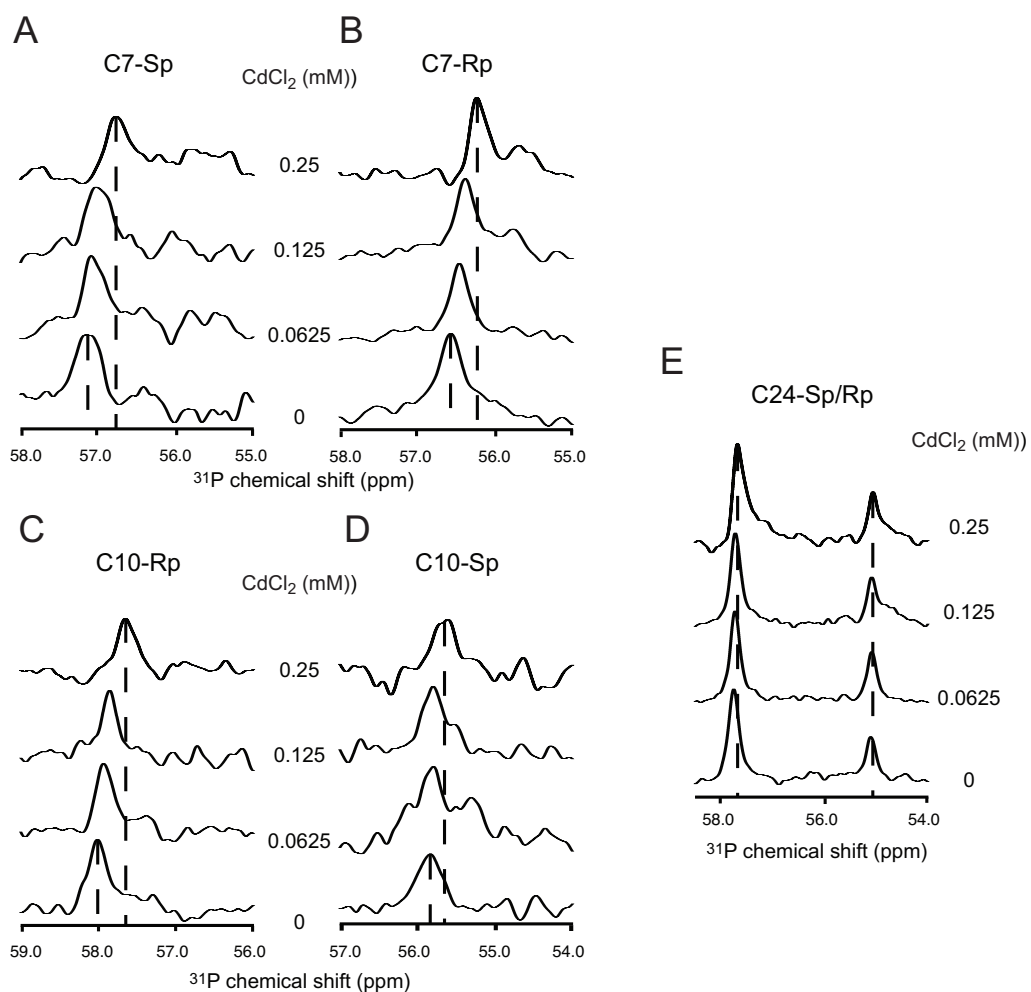


Figure S3.2. Effect of the addition of Cd^{2+} ions on the ^{31}P NMR spectrum of selected SLVI RNAs containing single phosphorothioate modifications. The 1D ^{31}P NMR spectra of (A) SLVI-C7-S_p, (B) SLVI-C7-R_p, (C) SLVI-C10-R_p, (D) SLVI-C10-S_p and (E) SLVI-C24-S_p/R_p were collected in buffer containing CdCl_2 : MgCl_2 concentrations of 0.0:5.0, 0.0625:4.9375, 0.125:4.875 and 0.25:4.75 mM. Only the most downfield region of the spectra containing the signal from the phosphorothioate is shown.

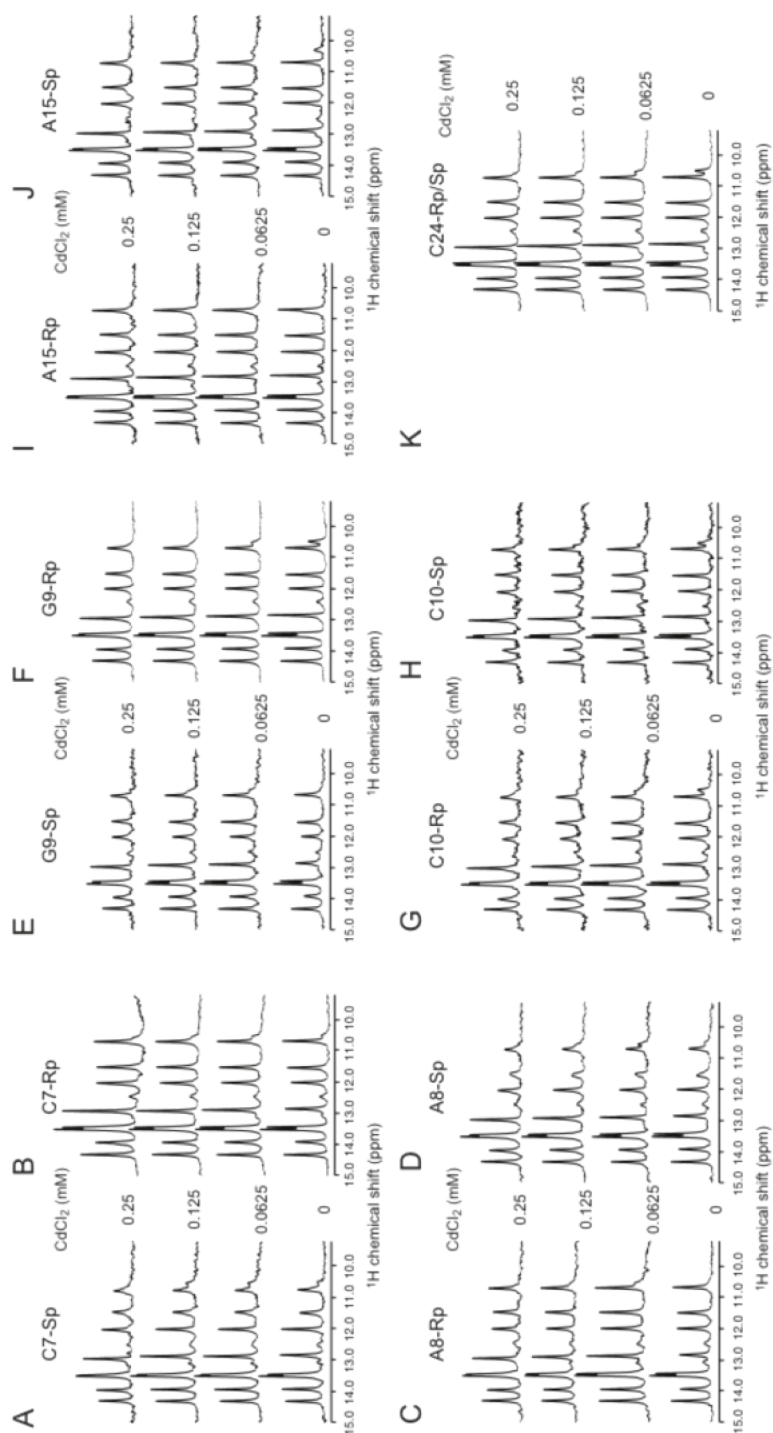


Figure S3.3. Effect of the addition of Cd^{2+} on the imino region of the ^1H NMR spectra of SLVI RNAs containing single phosphorothioate modifications. The 1D ^1H spectra of (A) SLVI-C7-S_p, (B) SLVI-C7-R_p, (C) SLVI-A8-R_p, (D) SLVI-A8-S_p, (E) SLVI-G9-S_p, (F) SLVI-G9-R_p, (G) SLVI-C10-S_p, (H) SLVI-C10-R_p (I) SLVI-A15-R_p, (J) SLVI-A15-S_p, and (K) SLVI-C24-S_p/R_p were collected in buffer containing CdCl_2 : MgCl_2 concentrations of 0.0:5.0, 0.0625:4.9375, 0.125:4.875 and 0.25:4.75 mM.

Table S3.1. Distance restraints derived from Mn²⁺-induced PRE.

Site	Atoms	I_0/I_{Mn}	Restraint
Site 1	G13 N1-H1	51	1.8-7.0 Å
	G13 C8-H8	20	1.8-7.0 Å
	G13 C3'-H3'	8	1.8-7.0 Å
	A16 N7	9	1.8-7.0 Å
	A14 H3'	6	1.8-8.0 Å
	A14 H5' or H5''	5	1.8-8.0 Å
	A15 N7	6	1.8-8.0 Å
	A16 H3'	5	1.8-8.0 Å
	A16 N9	4	1.8-8.0 Å
	A15 N9	7	1.8-8.0 Å
Site 1 or 2	C12 H6	7.5	1.8-10.0 Å
	C12 H5	6	1.8-10.0 Å
	G17 H8	6	1.8-10.0 Å
	G17 H1	5	1.8-10.0 Å
	C12 H41	5	1.8-10.0 Å
	C12 H42	5	1.8-10.0 Å
	G17 H5' or H5''	5	1.8-10.0 Å
Site 2	A11 N7	14	1.8-7.0 Å
	U18 C5-H5	8.5	1.8-7.0 Å
	C10 N4-H42	8	1.8-7.0 Å
	C10 N4-H41	6	1.8-8.0 Å
	G19 H8	7.5	1.8-8.0 Å
	C10 H41	6	1.8-8.0 Å
	A11 N9	5	1.8-8.0 Å
	A11 H8	4	1.8-8.0 Å
	G17 H3'	4	1.8-8.0 Å
U18 H3	4	1.8-8.0 Å	
Site 2 or 3	G9 H8	4	1.8-10.0 Å
Site 3	G6 H2'	6	1.8-8.0 Å
	A8 H8	5	1.8-8.0 Å
	A8 H3'	4	1.8-8.0 Å
	A8 H4'	4	1.8-8.0 Å
	A8 N9	4	1.8-8.0 Å
Site 3 or 4	G6 H8	5	1.8-10.0 Å
Site 4	G23 N7	15	1.8-7.0 Å
	G23 C8-H8	12	1.8-7.0 Å
	G22 C8-H8	9	1.8-7.0 Å
	U5 N3-H3	9	1.8-7.0 Å
	G6 N1-H1	8	1.8-7.0 Å
	A20 N7	5	1.8-8.0 Å
	U5 H5	4	1.8-8.0 Å
Site 4 or 5	G3 N9	6	1.8-10.0 Å
	C24 H42	5	1.8-10.0 Å
	C21 H5	5	1.8-10.0 Å
	C4 H41	4	1.8-10.0 Å
Site 5	G3 N7	15	1.8-7.0 Å
	G1 C8-H8	13	1.8-7.0 Å
	C26 H41	6	1.8-8.0 Å
	A2 N7	5	1.8-8.0 Å
	A2 H8	4	1.8-8.0 Å

Table S3.2. Results from the WebFR3D motif search.

Site	Searched motif ^a	Number of raw hits ^b	Number of hits with a cation ^b	Number of positive hits ^b	PDB entry (positive hits)	Residues (positive hits)	Metal ion ^c (positive hits)
Site 1	NR G – A C – G 5' 3'	111	13	3	3U5F 1FJG 1JTW	1121-1126 897-902 6-11	Mg 2162 Mg 1586 Na 25
Site 2	3' 5' C – G A – U C – G 5' 3'	173	18	8	3U5H 3U5H 2ZNI 3V2F 3V2F 3V2F 3V7E 3BNL	1183-1185/1321-1323 1497-1499/1517-1519 26-28/42-44 589-591/666-668 414-416/2407-2409 2678-2680/2727-2729 48-50/85-87 8-10/13-15	Mg 4027 Mg 3997 Ca 82 Mg 3471 Mg 3288 Mg 3626 NCO 910 NCO 026
Site 3 (search 1)	3' 5' C – G G N A N G – C 5' 3'	1	0	0	-	-	
Site 3 (search 2)	3' 5' N – N G N A N N – N 5' 3'	1	0	0	-	-	
Site 4	3' 5' G – C U – G C – G 5' 3'	92	4	3	1FJG 3V2F 3U5H	1509-1511/1524-1526 1767-1769/1983-1985 1671-1673/1774-1776	Mg 1592 Mg 3524 Mg 4578
Site 5	3' 5' G – C A – U G – C 5' 3'	297	26	16	3U5H 3U5H 3U5H 3U5H 3U5H 1S72 3U5F 3IVK 3V2F 3T4B 3CUN 397D 2XZM 2ZJR 4ENB 1FJG	272- 274/291-293 1379-1381/1424-1426 333-335/28-30 2251-2253/2263-2265 2670-2672/2682-2684 116-118/122-124 1478-1480/1527-1529 72-74/83-85 1277-1279/1291-1293 309-311/325-327 66-68/78-80 26-28/37-39 266-268/275-277 1666-1668/1989-1991 8-10/45-47 407-409/433-435	Mg 4333 Mg 3948 Mg 4491 Mg 4458 Mg4118 Mg 8039 Mg 2087 Mg 137 Mg 3500 Ni 1015 Mg 506 Ca 49 Mg 43 Mg 2901 Iri 104 Mg 1609

^a N are undefined nucleotides and R are purines

^b See Experimental Procedures

^c Refers to metal ion residue in the pdb files, where Mg is a magnesium ion, Na is a sodium ion, Ca is a calcium ion, NCO is a cobalt hexamine ion, Ni is a nickel ion and Iri is a iridium hexamine ion.

4. Détermination de la structure RMN de la jonction III-IV-V et localisation des ions Mg^{2+} par relaxation paramagnétique induite par le Mn^{2+}

Mise en contexte

Les deux études précédentes ont révélé comment la boucle A730 se replie et quel est le rôle structural des ions Mg^{2+} dans l'adoption de la structure tridimensionnelle du site actif du ribozyme VS. Cependant, la formation du site actif requiert la formation de l'interaction *kissing-loop* I/V ainsi que l'association des boucles internes de SLI et de SLVI. Ces interactions tertiaires sont donc dépendantes du repliement du domaine catalytique qui est notamment régi par les jonctions à trois voies III-IV-V et II-III-VI. Ainsi, la détermination de la structure à haute résolution de ces jonctions est essentielle dans la compréhension de la structure complète du ribozyme VS.

Dans cette troisième étude nous nous sommes intéressés à la structure de la jonction III-IV-V. Cette jonction est importante pour l'activité du ribozyme VS puisqu'elle facilite la formation de l'interaction *kissing-loop* I/V en définissant l'orientation de l'hélice SLV. Cette étude, réalisée par spectroscopie RMN, a révélé que la jonction III-IV-V ne nécessite pas la présence d'ions Mg^{2+} pour se replier mais l'ajout d'ions divalents lui permet d'adopter une structure plus compacte. La structure globale de la jonction III-IV-V, en plus d'être en accord avec toutes les données biochimiques, est aussi caractérisée par l'empilement coaxial des tiges III et IV tandis que les tiges III et V décrivent un angle aigu, caractéristique des jonctions de la famille C. Le cœur de la jonction est formé de plusieurs motifs structuraux incluant entre autre quatre triplets de bases et un motif *U-turn*. De plus, nous avons déterminé la structure d'une nouvelle boucle terminale, soit une boucle cCUUGg qui coiffe la tige IV. Cette boucle terminale adopte une structure différente de la boucle gCUUGc. Finalement, six ions Mg^{2+} ont été localisés dans la structure de la jonction III-IV-V par une approche de relaxation paramagnétique induite par le manganèse (*Mn²⁺-induced paramagnetic relaxation enhancement*). Entre autre, la boucle cCUUGg forme un site de liaison de métaux tandis que deux autres ions sont coordonnés dans le cœur de la jonction. Le premier est coordonné par la chaîne ribose-phosphate du motif *U-turn* tandis que le second se situe entre les tiges III et V. Cet ion interagit avec le sillon majeur de la tige III et avec la chaîne ribose-phosphate de l'hélice V. La liaison de ces deux métaux supporte le repliement compact de la jonction en présence d'ions Mg^{2+} .

Dans ce troisième article, j'ai contribué avec le Dr. Pascale Legault à la conception du projet. J'ai aussi réalisé la majorité des synthèses des échantillons avec l'aide de Dominique Chaussé, une employée du laboratoire. De plus, la collecte et l'analyse des

données de RMN, le calcul et l'analyse des structures ainsi que la rédaction de l'article a été réalisé par le Dr. Pascale Legault et moi-même.

Nuclear Magnetic Resonance Structure of the III-IV-V Three-Way
Junction from the Varkud Satellite Ribozyme and Identification of
Magnesium-Binding Sites Using Paramagnetic Relaxation
Enhancement

Eric Bonneau and Pascale Legault*

Département de Biochimie et Médecine Moléculaire, Université de Montréal, C.P. 6128,
Succursale Centre-Ville, Montréal, QC, Canada, H3C 3J7

This research was originally published in Biochemistry:

Biochemistry. 2014 Oct 7;53(39):6264-75. doi: 10.1021/bi500826n. Epub 2014 Sep 19

ABSTRACT

The VS ribozyme is a catalytic RNA found within some natural isolates of *Neurospora* that is being used as a model system to improve our understanding of RNA structure, catalysis, and engineering. The catalytic domain contains five helical domains (SLII-SLVI) that are organized by two three-way junctions. The III-IV-V junction is required for high-affinity binding of the substrate domain (SLI) through formation of a kissing-loop interaction with SLV. Here, we determine the high-resolution nuclear magnetic resonance (NMR) structure of a 47-nucleotide RNA containing the III-IV-V junction (J345). The J345 RNA adopts a Y-shaped fold typical of the family C three-way junctions, with coaxial stacking between stems III and IV and an acute angle between stems III and V. The NMR structure reveals that the core of the III-IV-V junction contains four stacked base triples, a U-turn motif, a cross-strand stacking interaction, an A-minor interaction, and a ribose zipper. In addition, the NMR structure shows that the cCUUGg tetraloop used to stabilize stem IV adopts a novel RNA tetraloop fold, different from the known gCUUGc tetraloop structure. Using Mn²⁺-induced paramagnetic relaxation enhancement, we identify six Mg²⁺-binding sites within J345, including one associated with the cCUUGg tetraloop and two with the junction core. The NMR structure of J345 likely represents the conformation of the III-IV-V junction in the context of the active VS ribozyme and suggests that this junction functions as a dynamic hinge that contributes to substrate recognition and catalysis. Moreover, this study highlights a new role for family C three-way junctions in long-range tertiary interactions.

INTRODUCTION

Multiway junctions are ubiquitous structural elements of nucleic acids found in all domains of life that have numerous roles in the dynamic processes of living cells, such as genetic recombination and mRNA splicing. The most abundant forms of RNA junctions are three-way junctions in which three double-stranded helices are connected at the same branch point, and these are found in a plethora of different RNAs [1-3]. Three-way junctions are essential architectural elements that control tertiary structure folding by supporting the packing and positioning of preformed helical domains, and they often promote the formation of remote tertiary interactions. Three-way junctions also play important functional roles because they are key elements for targeting RNA-binding proteins [4], as well as for forming ligand-binding pockets in riboswitch aptamers [5, 6] and for the folding of catalytically active ribozymes [7, 8].

Three-way junctions generally fold in such a way that two of the three helices are coaxially stacked on one another to form a pseudocontinuous helix [1-3]. Topological analyses of known structures of three-way junctions have led to the classification of these junctions into three major families, named A, B and C [1]. Family C was found to be the most abundant class and the only one containing nonrRNAs. A characteristic feature of family C is that one of the single-stranded regions connecting the three helices is longer than the other two and usually adopts a sharp turn, such as a U-turn or a T-loop. These loop motifs are recurrent in three-way junctions, where they govern the folding of the junction cores and the relative position of the helices [2]. The single-stranded regions connecting the helices may also be involved in base triples and quadruples, as well as A-minor motifs [9] and ribose zippers [10]. Metal cations, typically Mg^{2+} ions, often assist the folding of three-way junctions by stabilizing interactions between distant helices, inducing conformational changes and neutralizing the electrostatic repulsion of negatively charged phosphates that may come into close contact at the junction core [11-14]. Despite this wealth of structural knowledge available on three-way junctions, it is still difficult to predict their structure and metal-binding sites on the basis of sequence alone.

The *Neurospora* Varkud Satellite (VS) RNA is a naturally occurring ribozyme and a member of the family of small nucleolytic ribozymes [15-20]. It catalyzes both cleavage

and ligation reactions of a specific phosphodiester bond [15, 21], and metal cations are required for both folding and catalysis [22-29]. The minimal functional VS ribozyme contains six stem-loops (or stems) separated into two functional domains: stem-loop I (SLI) forms the substrate domain, whereas stem-loops II-VI (SLII-SLVI, respectively) define the catalytic domain (Figure 4.1A; [23]) Substrate recognition involves formation of a high-affinity kissing-loop interaction between SLI and SLV, and this interaction induces a helix shift in SLI that conformationally activates the substrate for catalysis (Figure 4.1A; [30-34]) Like the hairpin ribozyme, the active site of the VS ribozyme is formed by the interaction of two internal loops, the substrate internal loop of SLI and the A730 internal loop of SLVI [18, 20, 35, 36]. A general acid-base mechanism has been proposed for the VS ribozyme cleavage reaction, with G₆₃₈ of SLI as the general base and A₇₅₆ of the A730 loop of SLVI as the general acid [35, 37-43].

In the VS ribozyme, the structure of the catalytic domain is governed by two three-way junctions: the II-III-VI and III-IV-V junctions (Figure 4.1A; [23]) These junctions play a critical role in defining the relative orientation of the connected helices [44, 45]. Although the III-IV-V junction is not absolutely required for activity, it is an important architectural domain that mediates recognition of the SLI substrate domain and thereby greatly increases the catalytic activity of the VS ribozyme [28]. Specifically, base substitutions of most of the core residues within the III-IV-V junction interfere with both the cleavage and ligation activities [38, 45, 46]. Furthermore, three residues in the longest single-stranded region are proposed to form a U-turn motif within the junction core [46]. A family C topology has been predicted for the III-IV-V junction by several laboratories [1, 3], but there is no consensus from these predictions in terms of coaxial stacking, which may involve either stems IV and V [47] or stems III and IV [1, 3]. The latter prediction is in agreement with low-resolution models of the VS ribozyme based on fluorescence resonance energy transfer (FRET; [45]) and small-angle X-ray scattering (SAXS; [48]).

In this study, we determine the nuclear magnetic resonance (NMR) solution structure of the J345 RNA (Figure 4.1B), a 47-nucleotide RNA that encompasses the VS ribozyme III-IV-V junction. The NMR structure reveals that the J345 RNA adopts a Y-shaped structure that belongs to family C of three-way junctions. The structure of the three-way junction within J345 is stabilized by a complex network of interactions, which includes coaxial stacking of stems III and IV. In addition, we determine that the cCUUGg tetraloop used to

stabilize stem IV adopts a structure that is different from the known gCUUGc tetraloop structures. Using Mn^{2+} -induced paramagnetic relaxation enhancement (PRE), [49] we characterize six Mg^{2+} -binding sites within the J345 RNA. Our study provides insights into the global folding of the VS ribozyme, particularly the structural role of the III-IV-V junction in positioning the SLV loop for SLI substrate recognition.

EXPERIMENTAL PROCEDURES

Plasmids and DNA Templates. Double-stranded PCR fragments encoding J345-VS and S3-VS RNAs and flanked by a T7 promoter were inserted into the *HindIII-EcoRI* sites of the pTZ19R-derived pTR-4 vector [50] to generate the pJ345 and pS3 plasmids, respectively. The J345-VS and S3-VS RNAs correspond to the J345 (Figure 4.1B) and S3 RNAs (Figure 4.2), respectively, with a Varkud Satellite (VS) ribozyme substrate at their 3'-ends [50]. Prior to transcription, the plasmids were fully linearized using *EcoRI* (New England Biolabs). For transcription of S12, S3₃₀₋₃₈ and S3₃₀₋₄₀ RNAs (Figure 4.2), partially single-stranded synthetic DNA templates were used (Integrated DNA Technologies; [51]).

RNA Synthesis and Purification. Unlabeled, ^{15}N -labeled, and $^{13}C/^{15}N$ -labeled RNAs (J345-VS, S12, S3-VS, S3₃₀₋₃₈ and S3₃₀₋₄₀) were synthesized in vitro from their DNA template using a T7 RNA polymerase with an N-terminal His₆ tag prepared in house [52] and the appropriate NTPs, either unlabeled NTPs (Sigma-Aldrich) or isotopically-labeled NTPs ($^{13}C/^{15}N$ - or ^{15}N -labeled) prepared in house [53]. Following the transcription reactions of J345-VS and S3-VS, 25 mM $MgCl_2$ and purified trans-cleaving TR-4 VS ribozyme (1:10 VS ribozyme:full-length transcribed RNA molar ratio; [50]) were added to the transcription mixtures and incubated for 8 h at 37°C to produce the J345 and S3 RNAs with homogeneous 3'-ends. J345, S12, and S3 RNAs were treated with calf intestinal alkaline phosphatase (CIP) (Roche Diagnostics) to remove their 5'-phosphates. The unlabeled S3₃₉₋₄₈ and S3₄₁₋₄₈ RNAs were chemically synthesized (Thermo Fisher Scientific). The segmentally labeled [^{15}N]{30-38}-S3 and ^{15}N {30-40}-S3 RNAs (Figure 4.2) were produced by RNA ligation with the T4 DNA ligase, two S3 fragments and a single-stranded DNA splint of 19 nucleotides that is fully complementary to S3. Detailed procedures for segmental isotope labeling and preparation of S12/S3 complexes are provided in the Supporting Information.

All RNAs were purified by denaturing gel electrophoresis followed by DEAE-Sepharose chromatography.

NMR Sample Preparation. The purified J345 RNAs and the different S12/S3 complexes were concentrated and exchanged in NMR buffer A [10 mM sodium cacodylate (pH 6.5), 50 mM KCl, and 0.05 mM NaN_3] in a 90% H_2O /10% D_2O mixture with Amicon Ultra-4 centrifugation devices (Millipore). The RNAs were then heated at 37°C for 2 min and cooled in an ice/water mixture for 5 min before the change to the final NMR buffer [NMR buffer A with 5 mM MgCl_2 (99.995%) (Sigma-Aldrich)]. For NMR studies in D_2O , the samples were obtained by four cycles of lyophilization and resuspension in 99.996% D_2O .

NMR Spectroscopy. All NMR experiments were conducted on a Varian ^{Unity}INOVA 600 MHz spectrometer equipped with a pulse-field gradient unit and an actively-shielded z-gradient $^1\text{H}/^{13}\text{C}/^{15}\text{N}$ triple-resonance probe. The resonance assignments and distance restraints were derived from 2D ^1H - ^{15}N HSQC, 2D HNN-COSY, 2D ^1H - ^{15}N CPMG-NOESY, 2D G-specific H(NC)-TOCSY-(C)H, 2D ^1H - ^{13}C CT-HSQC, 2D ^1H - ^{13}C HMQC, 2D ^1H - ^{15}N MQ-(HC)N(C)H, 3D CT-HCCH-COSY, 3D HCCH-TOCSY, 3D ^{15}N -edited NOESY-HSQC, 3D ^{13}C -edited HMQC-NOESY, and 3D $^{15}\text{N}/^{13}\text{C}$ -edited NOESY-HSQC spectra (Tables S4.1 and S4.2 of the Supporting Information; [34, 54, 55]). NMR data were processed using the NMRPipe/NMRDraw package [56] and analyzed with the CCPNMR suite [57].

Structural Restraints. The NOE-derived distance restraints were separated into four classes [strong (1.8-3.3 Å), medium (1.8-4.5 Å), weak (1.8-5.5 Å) and very weak (2.8-8.0 Å)] on the basis of NOE cross-peak intensities. On the basis of NMR evidence from NOESY and HNN-COSY spectra, canonical distance restraints were employed to define the hydrogen-bond pattern and planarity of the WC base pairs in helical regions (SLIII: 1-5/42-46, SLIV: 8-13/16-21, SLV: 24-27/32-35) as well as the typical G_{28} - A_{31} base pair of the GAAA tetraloop [58, 59]. Dihedral angle restraints for the sugar puckers (δ) and other backbone dihedral angles (α , γ , χ , and ζ) were defined from comparative NOE analyses.

On the basis of NMR evidence, backbone torsion angles of residues in helical regions were restrained to A-form values ($\pm 15^\circ$).

Metal Ion Binding Studies. Manganese (Mn^{2+}) titrations were performed with two J345 samples (1.25 mM ^{15}N -labeled J345 and 2.0 mM $^{13}\text{C}/^{15}\text{N}$ -labeled J345) in NMR buffer A with 5 mM MgCl_2 , as previously described [49]. Mn^{2+} -induced paramagnetic relaxation enhancement (PRE) was monitored by 2D ^1H - ^{15}N HSQC [60] spectra at 15°C and 2D ^1H - ^{13}C CT-HSQC [61, 62] spectra at 25°C [49]. The RNA-metal distance restraints were derived from Mn^{2+} -induced PRE using the ratio of signal intensity (I_0/I_{Mn}) determined from spectra collected at $0 \mu\text{M}$ MnCl_2 (I_0) and $10 \mu\text{M}$ MnCl_2 (I_{Mn} ; [49]).

Structure Calculation. Three-dimensional structures of J345 were calculated with restrained molecular dynamics and simulated annealing in X-PLOR-NIH version 2.1.9 [63], from structures with randomized backbone angles, as previously described [55]. A force field was used that included bond, angle, improper, and repulsive van der Waals energy terms as well as NOE and torsion-angle pseudoenergy terms. Electrostatic contributions were not included in the force field. Three-dimensional structures of J345 bound to hydrated magnesium complexes [$\text{Mg}(\text{H}_2\text{O})_n^{2+}$], termed J345^{Mg}, were calculated as described for J345 using $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ coordinates and parameters, as previously described [49]. A pentahydrated Mg^{2+} ion [$\text{Mg}(\text{H}_2\text{O})_5^{2+}$] was used at site 1 (GAAA tetraloop; [49]), whereas sites 2-6 were modeled with hexahydrated Mg^{2+} ions [$\text{Mg}(\text{H}_2\text{O})_6^{2+}$]. For both J345 and J345^{Mg}, the 20 lowest-energy structures that satisfied all the experimental restraints (all distance violations $< 0.2 \text{ \AA}$ and all torsion-angle violations $< 5^\circ$) were selected for analysis and used to calculate an average structure that was minimized against all restraints. Structures were visualized with PyMOL Molecular Graphics System version 1.3 (Schrödinger, LLC) and analyzed with PyMOL and Curves+ [64]. Reported values of rmsds, interatomic distances, and interhelical angles are given as average values with standard deviations for the 20 lowest-energy structures, unless otherwise mentioned. Interhelical angles were calculated using residues 1-5 and 42-46 of stem III, residues 8-12 and 17-21 of stem IV, and residues 24-27 and 32-35 of stem V.

RESULTS

The J345 RNA Adopts a Compact Structure in the Presence of Mg²⁺ Ions. The J345 RNA is a 47-nucleotide three-way junction that incorporates all the single-stranded residues of the junction core as well as the proposed proximal base pairs of stems III-V found in the wild-type VS ribozyme (Figure 4.1A,B; [23]). In addition, the J345 RNA contains two non-natural terminal loops that were included to facilitate NMR assignment: a CUUG tetraloop [65] in stem-loop IV and a GAAA tetraloop [58, 59] in stem-loop V (Figure 4.1B). Both tetraloops are closed by a C-G base pair and are termed herein cCUUGg and cGAAAg tetraloops, respectively.

NMR studies of J345 were conducted in the presence of 5 mM MgCl₂, because the VS ribozyme requires metal cations, preferentially Mg²⁺ ions, for its folding and catalytic activities [22-29]. Under these conditions, the 2D ¹H-¹⁵N HSQC spectrum (Figure 4.1C) is well-dispersed with imino signals representing 24 of the 27 imino groups in J345, in agreement with formation of a unique stable structure for J345. All 24 imino signals were also observed in the absence of Mg²⁺ ions (Figure S4.1 of the Supporting Information), indicating that addition of Mg²⁺ ions induces only minor changes in the structure of J345. These imino signals could be assigned from NOESY spectra, except for the U imino signals from the junction core.

To allow unambiguous assignment of these U imino signals, a segmental ¹⁵N isotope labeling approach was developed using a two-stranded complex similar to J345, designated the S12/S3 complex. The 2D ¹H-¹⁵N HSQC spectrum of [¹⁵N]-S12/[¹⁵N]-S3 (Figure S4.2 of the Supporting Information) together with 2D HNN-COSY and 3D ¹⁵N-edited HSQC-NOESY spectra (not shown) indicates that the S12/S3 complex adopts a structure that is very similar to that of J345. This 2D ¹H-¹⁵N HSQC spectrum was then compared with that of differentially-labeled S12/S3 complexes (Figure 4.2). The U₆ and U₂₂ imino signals disappear with the unlabeled S12/[¹⁵N]-S3 complex (Figure 4.2B), the U₄₁ imino signal disappears with the [¹⁵N]-S12/[¹⁵N]{30-40}-S3 complex (Figure 4.2C); whereas the U₄₀ and U₄₁ imino signals disappear with the [¹⁵N]-S12/[¹⁵N]{30-38}-S3 complex (Figure 4.2D). The unambiguous assignment of these U imino signals allowed further interpretation of NOESY spectra (Figure S4.3 of the Supporting Information). Intense NOEs between the U₆ and U₄₁ imino protons are consistent with formation of a *cis* WC/WC U₆-U₄₁ base pair within the core. In addition, NOE connectivities involving imino protons of four uridines

from the core (U_6 , U_{22} , U_{40} , and U_{41}) as well as G_{42} and G_{21} from stems III and IV provide evidence of a highly organized core in which stems III and IV are coaxially stacked, as previously proposed [1, 3, 45, 48].

The Overall NMR Structure of the J345 RNA. The three-dimensional structure of J345 was determined using NOE-derived distance restraints and dihedral-angle restraints. Its overall structure is well-defined by the NMR data with a heavy-atoms rmsd of 3.72 ± 1.62 Å for the 20 lowest-energy structures (Table 4.1 and Figure 4.3). The J345 RNA adopts a Y-shaped fold with a well-defined core domain (rmsd of 0.97 ± 0.13 Å) that orients stem III (rmsd of 0.35 ± 0.09 Å) and stem V (rmsd of 0.82 ± 0.31 Å) side by side and away from stem IV (rmsd of 0.36 ± 0.09 Å). Stems III and V define the most acute interhelical angle ($\phi_{III-V} = 32.8^\circ \pm 22.6^\circ$), whereas the other two stem pairs define obtuse interhelical angles ($\phi_{IV-V} = 149.3^\circ \pm 14.2^\circ$, and $\phi_{III-IV} = 167.1^\circ \pm 10.1^\circ$), with stems III and IV being nearly coaxial.

Compact cCUUGg and cGAAAg Tetraloops Capping Stems IV and V. The $cG_{28}A_{29}A_{30}A_{31}g$ tetraloop in stem-loop V adopts a typical GNRA fold [58, 59], with a *cis* WC/WC C-G closing base pair ($C_{27}-G_{32}$), a *trans* sugar edge/Hoogsteen G-A base pair between the first and last residues of the loop ($G_{28}-A_{31}$), a sharp backbone turn between the first and second residues (G_{28} and A_{29}), 3'-stacking among the last three bases of the loop, and two typical hydrogen bonds (not shown). In contrast, the cCUUGg tetraloop adopts a structure different from that previously reported for the related gCUUGc tetraloop [65]. In J345, the $cC_{13}U_{14}U_{15}G_{16}g$ tetraloop is stabilized by two *cis* WC/WC C-G base pairs: the $C_{12}-G_{17}$ closing base pair and the $C_{13}-G_{16}$ base pair between the first and last residues of the loop (Figure 4.4). An analysis of rotamer geometries yields non-A-form values for residues U_{14} , U_{15} , and G_{16} [66], in agreement with the backbone turn of this cCUUGg tetraloop not being as sharp as that of the GNRA tetraloop (Figure 4.4). In addition, the cCUUGg tetraloop displays 5'-stacking among the first three bases of the loop [C_{13} , U_{14} , and U_{15} (Figure 4.4)]. All the residues in the cCUUGg tetraloop have glycosidic angles in the anti conformation and sugar puckers in the C3'-endo conformation, with the exception of U_{14} and U_{15} that adopt a C2'-endo sugar pucker. There are two hydrogen bonds that potentially stabilize the cCUUGg loop as evidenced by short interatomic distances involving U_{15} O2' (Figure 4.4B).

In summary, the cCUUGg tetraloop bears a structural resemblance to the cGAAAg tetraloop with respect to the formation of base pairing, hydrogen bonding, and stacking interactions, but there are significant differences in terms of the residues involved in these interactions and the backbone geometry.

Formation of a U-Turn at the Junction Core. As proposed previously [46], residues U₃₇, G₃₈, and A₃₉ adopt a canonical U-turn structure within the well-defined core domain of J345 (Figure 4.5; [67]). In the J345 structure, the U-turn stacks on stem V and produces a sharp turn in the backbone after U₃₇ that positions stem V side by side with respect to stem III. In addition, two hydrogen bonds, typical of the U-turn fold, are observed between U₃₇ 2'-OH and A₃₉ N7 and between U₃₇ H3 and A₃₉ 3'-phosphate (Figure 4.5B).

The G₃₆ residue preceding the U-turn does not form a canonical *cis* WC/WC base pair with C₂₃, as previously proposed [23]. Rather, G₃₆ forms a cross-strand hydrogen bond involving its amino group and the phosphate of C₂₃ [G₃₆ N2-C₂₃ 5'-phosphate oxygen distance of 3.40 ± 0.50 Å; (Figure 4.5)], while the C₂₃ base is flipped out to fulfill other interactions within the core. These cross-strand interactions involving G₃₆ likely stabilize the U-turn by allowing continuous stacking among U₃₇, G₃₆, and the G₂₄-U₃₅ base pair.

A Complex Network of Interactions within the Core of the III-IV-V Junction. The core domain of J345 is stabilized by a stack of four base triples that is well-defined by the NMR data (Figures 4.5 and 4.6). Within the U₆-U₄₁-C₂₃ base triple (rmsd of 0.81 ± 0.19 Å), U₆ and U₄₁ form a *cis* WC/WC U-U base pair, and the major groove of this U₆-U₄₁ base pair interacts with the flipped-out base of C₂₃ through a single hydrogen bond. The U₆-U₄₁-C₂₃ base triple stacks on the C₇-U₄₀-U₂₂ base triple (rmsd of 0.68 ± 0.15 Å), which is stabilized by three hydrogen bonds involving the WC edges of these residues. The C₇-U₄₀-U₂₂ base triple stacks on the C₈-G₂₁-A₃₉ base triple (rmsd of 0.46 ± 0.13 Å), where C₈ and G₂₁ form a canonical *cis* WC/WC base pair at the end of stem IV, while A₃₉ of the U-turn participates in a type I A-minor interaction with the C₈-G₂₁ base pair [9]. Finally, the C₈-G₂₁-A₃₉ base triple stacks on the C₉-G₂₀-G₃₈ base triple (rmsd of 0.66 ± 0.25 Å), where C₉ and G₂₀ form a canonical *cis* WC/WC base pair in stem IV, while G₃₈ interacts with the minor groove of the C₉-G₂₀ base pair. Although the base of G₃₈ is not well-defined in the ensemble of NMR

structures (Figure 4.5), G₃₈ could potentially form several hydrogen bonds with the riboses of C₉ (Figure 4.6) and G₁₀ [G₃₈ NH₂ to G₁₀ O4' and O2' (not shown)].

The stack of four base triples within the J345 core is also stabilized by a ribose zipper [10], which involves close backbone contacts of G₃₈ and A₃₉ from the U-turn with C₈ and C₉ of stem IV. As is typical of ribose zippers [10], several hydrogen bonds are observed that involve consecutive 2'-OH groups (Figure S4.4 of the Supporting Information). Such a network of interactions allows close packing of the U-turn within the junction core.

Identification of Mg(H₂O)_n²⁺-Binding Sites. The structure of J345 in complex with Mg(H₂O)_n²⁺ (J345^{Mg}) was determined by adding metal-RNA restraints derived from Mn²⁺-induced PRE to the existing set of experimental restraints (Table 4.1 and Table S4.3 of the Supporting Information). The overall structure of the J345^{Mg} RNA is very well-defined by the NMR data with a heavy atom rmsd of 2.63 ± 0.80 Å (Table 4.1). In comparison with J345, the global structure of J345^{Mg} is better defined, even though local structural elements have similar resolutions in both structures.

The J345^{Mg} structure displays six Mg²⁺-binding sites (Figure 4.7A), and potential metal ligands at these sites were defined on the basis of distance measurements (Table S4.4 of the Supporting Information). These sites defined by the NMR data should be viewed as preferential sites for Mg²⁺-ion binding, and no inference can be made in terms of the occupancy of the metal at a given site. The Mg(H₂O)₅²⁺ at site 1 is associated with the cG₂₈A₂₉A₃₀A₃₁g tetraloop, whereas the Mg(H₂O)₆²⁺ species at sites 2 and 4 are associated with the major grooves of stems V and IV, respectively (Figure S4.5A-C of the Supporting Information), in agreement with previous observations [49]. The Mg(H₂O)₆²⁺ at site 3 is associated with the cCUUGg tetraloop (Figure 4.7B). The O6 and N7 atoms of G₁₆ and G₁₇ act as potential outer-sphere ligands, whereas the G₁₆ 5'-PO₄⁻ contributes to electrostatic stabilization of the Mg²⁺ ion. The Mg(H₂O)₆²⁺ at site 5 is associated with the minor groove of the U-turn within the core domain (Figure 4.7C), where the ribose phosphate of the U-turn (residues 37-40) provides several outer-sphere ligands as well as electrostatic stabilization. Remarkably, a cation-π interaction is also observed between this Mg(H₂O)₆²⁺ and the U₃₇ base that defines the U-turn. The Mg²⁺ ion at this position is likely important for stabilization of the U-turn fold. The Mg(H₂O)₆²⁺ at site 6 is located near the U₆-U₄₁-C₂₃ base triple of the core domain within a pocket formed by the major groove of stem III and the

phosphate backbone of stem V (Figure 4.7D). Residues C₂₃, G₂₄, G₄₂, and G₄₃ provide several potential ligands, whereas the 5'-phosphates of G₂₄ and G₄₂ contribute to electrostatic stabilization of the Mg²⁺ ion. Site 6 is unique in that Mn²⁺-induced PRE affects residues of both stems III and V (Table S4.3 of the Supporting Information). Thus, RNA-metal restraints at this site help improve the overall rmsd of J345 by better defining the relative orientation of these two stems.

DISCUSSION

As part of our quest to understand the structural basis for substrate recognition and cleavage by the VS ribozyme, we determined the high-resolution solution structure of J345, which contains a three-way junction that plays a key role in defining the global architecture of the VS ribozyme. In addition, we localized Mg²⁺-binding sites within J345 using Mn²⁺-induced PRE. The structure of J345 reveals an intricate network of interactions at the junction core as well as a novel fold for the cCUUGg tetraloop that was introduced to stabilize stem-loop IV, and both are discussed below in light of previous studies.

A Novel cCUUGg Tetraloop Structure. To date, thermodynamic and structural studies of CUUG tetraloops have mostly focused on the gCUUGc sequence because of its greater prevalence in rRNA [68, 69]. The gCUUGc and cCUUGg tetraloops are considered unusually stable tetraloops [70], although they are less stable than the more common cGNRAg and cUNCGg tetraloops [71, 72]. At this time, symbolic searches within the nonredundant list of RNA-containing structures from the Protein Data Bank (PDB) using WebFR3D [73] revealed two essentially identical structures of gCUUGc tetraloops, one from the NMR structure of a short RNA hairpin (PDB entry 1RNG; [65]) and one from the X-ray structure of the *Escherichia coli* ribosome (PDB entry 2AW7; [74]); however, there is no available structure for the cCUUGg loop.

The NMR structure of J345 reveals that the cCUUGg tetraloop structure is significantly different from those reported for the gCUUGc tetraloop, in agreement with previous circular dichroism studies [70]. One important characteristic of the gCUUGc structures is that the first uridine is extruded into the minor groove, forming hydrogen bonds with the closing and loop base pairs [65, 74]. In contrast, the cCUUGg tetraloop adopts a compact

fold with consecutive stacking of the first three residues. This main structural difference may arise from the inversion of the closing base pair (G-C to C-G), which does not allow hydrogen bonding of an extruded U residue in the minor groove in the cCUUGg tetraloop [65, 74]. This structural difference could also be due to the presence of Mg^{2+} ions in our study, because the NMR structure of the gCUUGc tetraloop was determined in the absence of Mg^{2+} ions [65]. Using Mn^{2+} -induced PRE, we identified a Mg^{2+} -binding site associated with the major groove of the cCUUGg tetraloop that likely helps stabilize the observed fold.

An important feature of the cCUUGg tetraloop is that the WC edges of the loop residues are all oriented toward the major groove. In contrast, both the GNRA and gCUUGc tetraloops expose loop bases in the minor groove. In the GNRA tetraloop, these bases are known to play important roles in forming tertiary interactions [75, 76]. Although the gCUUGc tetraloop is free in the context of the folded *E. coli* ribosome [74], protrusion of its first uridine into the minor groove could allow for RNA-RNA or RNA-protein interactions in other contexts [65]. In contrast, the exposed bases of the cCUUGg tetraloop appear to be more retracted in the major groove and less accessible for tertiary contacts, and this may explain the low frequency of this tetraloop in rRNA.

Compatibility of the J345 Structure with Biochemical Data. The NMR structure of J345 reveals a complex network of interactions for the VS ribozyme III-IV-V junction that is in agreement with previous biochemical data. The NMR structure clearly demonstrates that residues U_{710} , G_{711} , and A_{712} adopt a canonical U-turn structure (Figure 4.5), in agreement with its functional substitution with a hairpin, as well as mutational and chemical probing data [23, 24, 26, 28, 29, 38, 45, 46]. In particular, the effect of base substitutions on both cleavage and ligation activities is in agreement with the UNR consensus sequence (N is any base, and R is A or G) of the U-turn [38, 45, 46]. Furthermore, the specific requirement for an adenine at the R position is in agreement with the observed A-minor interaction involving A_{712} (Figure 4.6).

Kinetic studies of VS ribozyme variants carrying single-nucleotide substitutions have previously highlighted the importance of several residues in and around the junction [38, 45, 46]. For sake of simplicity, we restrict our discussion to the effect of base substitutions on cleavage activity of the VS ribozyme [45, 46]; however, similar effects on its ligation

activity were also observed [38]. In particular, base substitutions of all junction residues initially predicted to be single-stranded (Figure 4.1A) reduced cleavage activity, except for G₇₁₁ [45, 46]. This is in agreement with the NMR structure of J345 in which all bases within the junction core are directly involved in stable hydrogen-bonding and stacking interactions (Figures 4.5 and 4.6), except for the less-well-defined G₇₁₁ (G₃₈) that likely forms less stable interactions. In terms of the junction proximal base pairs, only a 4-fold reduction in cleavage activity was observed for inversion of the closing base pair of stem III (C₆₆₃-G₇₁₅), whereas inversion of the proposed closing base pairs of stem IV (C₆₆₆-G₆₈₅) and stem V (C₆₈₇-G₇₀₉) reduced the VS ribozyme cleavage activity by 30- and 125-fold, respectively [45]. These results are consistent with the C₆₆₆-G₆₈₅ base pair forming a base triple with A₇₁₂ (Figure 4.6) and with the previously proposed C₆₈₇-G₇₀₉ base pair (Figure 4.1A,B) not being present in the J345 structure determined here [23]. Rather, C₆₈₇ forms a base triple (U₆₆₄-U₇₁₄-C₆₈₇), and G₇₀₉ forms cross-strand stacking and hydrogen bonding interactions (Figures 4.5 and 4.6).

Chemical probing experiments conducted in the presence of Mg²⁺ ions are also in agreement with the NMR structure of J345. The WC edges of all the junction residues initially predicted to be single-stranded are protected from chemical modifications under native conditions, with the exception of G₇₁₁ (G₃₈; [23, 28, 29]), in agreement with the J345 structure. In addition, chemical modifications of bases that play an important role in the structure of the III-IV-V junction have been shown to significantly interfere with self-cleavage; this includes carboxyethylation at the N7 position of G₇₀₉ (G₃₆) and A₇₁₂ (A₃₉) by diethyl pyrocarbonate (DEPC) and removal of U₆₈₆ (U₂₂) and U₇₁₀ (U₃₇) by hydrazine [24]. The involvement of A₇₁₂ in the U-turn and A-minor interaction (Figures 4.5 and 4.6) also helps explain results of adenine analogue interference experiments, in which chemical modifications at the N7 and C2 positions significantly inhibit cleavage activity [37, 77]. Considering the backbone, the 5'-phosphates of A₇₁₂ (A₃₉) and U₇₁₃ (U₄₀) are protected from ethylnitrosourea modifications [26], which is supported by Mg²⁺ binding (Figure 4.7) and hydrogen bonding within the U-turn (Figure 4.5B). In addition, C₆₆₆ (C₈), A₆₆₇ (C₉), U₆₈₆ (U₂₂), C₆₈₇ (C₂₃), and A₇₁₂ (A₃₉) are protected from hydroxyl radical footprinting [27], consistent with the poor accessibility of their C4' atoms within the ribose zipper [for C₆₆₆, A₆₆₇ and A₇₁₂ (Figure S4.4 of the Supporting Information)] and base triples [for U₆₈₆, C₆₈₇ and A₇₁₂ (Figure 4.6)]. Moreover, 2'-deoxynucleotide modifications of several residues in

the III-IV-V junction (C_{666} , U_{686} , U_{710} , A_{712} , and U_{713}) were shown to significantly inhibit the cleavage and/or ligation activities of the VS ribozyme, in agreement with the importance of these 2'-hydroxyls in the ribose zipper (Figure S4.4 of the Supporting Information), the U-turn (Figure 4.5), and several base triples (Figure 4.6; [37, 78]). Hence, the NMR structure of J345 is consistent with previous biochemical data and most likely represents the conformation of the III-IV-V junction in the context of the active VS ribozyme.

The Topology of the III-IV-V Junction is Typical of the Family C Three-Way Junctions. The NMR structure of J345 confirms that the III-IV-V junction of the VS ribozyme adopts a global Y-shaped fold typical of family C three-way junctions, as previously proposed [1, 3]. Previous predictions also indicated two potential coaxial stacking interactions, either between stems IV and V [47] or between stems III and IV [1, 3]. The NMR structure establishes that stems III and IV are coaxially stacked and that there is an acute angle between stems III and V. This coaxial stacking of stems III and IV is also in agreement with low-resolution structural models based on FRET and SAXS data [45, 48]. Although these models could not precisely account for the structural complexity of the III-IV-V junction, they both predicted an acute interhelical angle between stems III and V (ϕ_{III-V} ; [45, 48]). Interestingly, the ϕ_{III-V} value of $\sim 40^\circ$ estimated from the SAXS model [48] falls within the range observed in the ensemble of NMR structures of J345 ($\phi_{III-V} = 32.8^\circ \pm 22.6$).

Diverse Roles for Three-Way Junctions in Long-Range Tertiary Interactions. The structure of the III-IV-V junction relies on numerous local structural elements that have been previously found in several family C three-way junctions, namely a U-turn motif, noncanonical base pairs and base triples, the formation of A-minor and cross-strand interactions, and a ribose zipper [1, 5-8, 79-90]. However, the III-IV-V junction does not include a long-range interaction between the helices that are positioned side-by-side (stems III and V; [2]), as commonly found in family C three-way junctions from various RNAs [1, 5, 7, 8, 79-90]. Instead, the III-IV-V junction of the VS ribozyme allows for the interaction between SLV and an exogenous stem-loop (SLI) that does not contribute to forming the junction [30-34], and, to the best of our knowledge, this type of tertiary interaction for family C three-way junctions has not been previously reported. Thus, this study expands on

our understanding of the diverse roles that three-way junctions play in the organization of complex RNA structures, which could have useful applications in nanotechnologies.

Effect of Mg²⁺ Ions on the Folding of the III-IV-V Junction. Chemical probing data of the VS ribozyme have shown increased levels of protection of several residues of the junction upon addition of Mg²⁺ ions [23, 24, 26-29]. In addition, native gel and FRET studies have provided evidence of a Mg²⁺-induced structural transition that alters the global structure of the III-IV-V junction [45]. From these studies, it was proposed that the junction adopts an extended structure in the absence of divalent metal ions that folds upon addition of Mg²⁺ ions [45]. However, the NMR data of J345 presented here indicate that the III-IV-V junction is prefolded in the absence of divalent cations and that the local structural features of the junction are not considerably altered upon addition of Mg²⁺ ions. Moreover, the presence of two divalent metal-binding sites in the core of the III-IV-V junction indicates that these divalent cations help stabilize both the U-turn motif and the close packing of stems III and V. This is in agreement with results from FRET studies in which an increasing Mg²⁺ ion concentration systematically reduced the end-to-end distance between stems III and V [45]. Taken together, these data support a model in which Mg²⁺ ion binding helps stabilize a preformed structure at the III-IV-V junction to allow close packing of stems III and V.

The III-IV-V Junction Forms a Dynamic Hinge for SLI Substrate Recognition and Catalysis. The main role of the III-IV-V junction in the VS ribozyme is to properly orient SLV for SLI substrate binding and catalysis. As currently viewed, SLV forms a high-affinity kissing loop interaction with SLI [33], and this facilitates the association between the SLI and SLVI internal loops to form the active site. With this study, we establish that SLV is not locked by tertiary interactions with SLIII as would be expected for a typical family C three-way junction. Rather, SLV appears to have a certain degree of flexibility with respect to SLIII, and thus, the junction most likely functions as a dynamic hinge to assist SLV in exploring the conformational space needed to facilitate SLI binding and catalysis. In particular, a certain degree of flexibility may be required to allow formation of the active site through the intimate association of the SLI and SLVI internal loops.

Acknowledgments

We thank Thomas R. Cech for providing the expression vector for the T4 DNA ligase, Mathieu Lussier-Price for help with the purification of the T4 DNA ligase, Dominique Chaussé for help with sample preparation, and Ryan Richter for providing computer scripts for structural analysis and for computer support. We also thank Nicolas Girard, Patricia Bouchard and James G. Omichinski for discussions and critical reading of the manuscript.

Supporting Information

Experimental procedures providing more details about the methods used in this study; figures that show NMR evidence that J345 adopts a single stable structure in the absence of Mg^{2+} ions that is similar to that formed in the presence of Mg^{2+} ions (Figure S4.1), NMR evidence that the S12/S3 complex adopts a structure similar to that of J345 (Figure S4.2), NMR evidence that J345 forms a highly organized core with helical stacking between stems III and IV (Figure S4.3), representation of the ribose zipper within the core domain of the J345 RNA (Figure S4.4), and representations of additional $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ -binding sites within J345^{Mg} (Figure S4.5); and tables that show the resonance assignment of J345 (Tables S4.1 and S4.2), distance restraints derived from Mn^{2+} -induced PRE used for the J345^{Mg} structure calculation (Table S4.3), and structural characteristics of $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ -binding sites in J345^{Mg} (Table S4.4). This material is available free of charge via the Internet at <http://pubs.acs.org>.

Accession Codes

NMR chemical shifts, structural restraints, and structural coordinates have been deposited for J345 and J345^{Mg} as BMRB entries 25163 and 25164 and PDB entries 2mtj and 2mtk, respectively.

Author Information

Funding Sources

This work was supported by the Canadian Institutes of Health Research (CIHR) (Grant MOP-86502 to P.L.). E.B. was supported by a Frederick Banting and Charles Best Canada Graduate Scholarship Ph.D. scholarship from CIHR and graduate scholarships from the Université de Montréal. P.L. holds a Canada Research Chair in Structural Biology and Engineering of RNA. Funding for open access charge provided by CIHR.

Notes

The authors declare no competing financial interest.

Abbreviations

2D, two-dimensional; 3D, three-dimensional; FRET, fluorescence resonance energy transfer; HNN-COSY, HNN-correlated spectroscopy; HSQC, heteronuclear single quantum coherence; Mg^{2+} , magnesium ion; Mn^{2+} , manganese ion; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; nt, nucleotides; PDB, protein data bank; PRE, paramagnetic relaxation enhancement; SLI, stem-loop I; SLII, stem-loop II, SLV, stem-loop V; SLVI, stem-loop VI; rmsd, root-mean-square deviation; SAXS, small angle X-ray scattering; VS, Varkud satellite; WC, Watson-Crick.

REFERENCES

1. Lescoute, A. and E. Westhof, *Topology of three-way junctions in folded RNAs*. RNA, 2006. **12**(1): p. 83-93.
2. de la Pena, M., D. Dufour, and J. Gallego, *Three-way RNA junctions with remote tertiary contacts: a recurrent and highly versatile fold*. RNA, 2009. **15**(11): p. 1949-64.
3. Laing, C., et al., *Predicting coaxial helical stacking in RNA junctions*. Nucleic Acids Res., 2012. **40**(2): p. 487-98.
4. Woodson, S.A. and N.B. Leontis, *Structure and dynamics of ribosomal RNA*. Curr. Opin. Struct. Biol., 1998. **8**(3): p. 294-300.

5. Batey, R.T., S.D. Gilbert, and R.K. Montange, *Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine*. *Nature*, 2004. **432**(7015): p. 411-5.
6. Serganov, A., et al., *Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs*. *Chem. Biol.*, 2004. **11**(12): p. 1729-41.
7. Cate, J.H., et al., *Crystal structure of a group I ribozyme domain: principles of RNA packing*. *Science*, 1996. **273**: p. 1678-1685.
8. Krasilnikov, A.S., et al., *Crystal structure of the specificity domain of ribonuclease P*. *Nature*, 2003. **421**(6924): p. 760-4.
9. Nissen, P., et al., *RNA tertiary interaction in the large ribosomal subunit: the A-minor motif*. *Proc. Natl. Acad. Sci. U.S.A.*, 2001. **98**(9): p. 4899-4903.
10. Tamura, M. and S.R. Holbrook, *Sequence and structural conservation in RNA ribose zippers*. *J. Mol. Biol.*, 2002. **320**: p. 455-474.
11. Draper, D.E., *A guide to ions and RNA structure*. *RNA*, 2004. **10**: p. 335-343.
12. Wakeman, C.A., A. Ramesh, and W.C. Winkler, *Multiple metal-binding cores are required for metalloregulation by M-box riboswitch RNAs*. *J. Mol. Biol.*, 2009. **392**(3): p. 723-35.
13. Chen, G., Z.J. Tan, and S.J. Chen, *Salt-dependent folding energy landscape of RNA three-way junction*. *Biophys J*, 2010. **98**(1): p. 111-20.
14. Zhang, H., et al., *Crystal structure of 3WJ core revealing divalent ion-promoted thermostability and assembly of the Phi29 hexameric motor pRNA*. *RNA*, 2013. **19**(9): p. 1226-37.
15. Saville, B.J. and R.A. Collins, *A site-specific self-cleavage reaction performed by a novel RNA in Neurospora mitochondria*. *Cell*, 1990. **61**: p. 685-696.
16. Collins, R.A., *The Neurospora Varkud satellite ribozyme*. *Bioch. Soc. Trans. Rev.*, 2002. **30**: p. 1122-1126.
17. Lilley, D.M., *The Varkud satellite ribozyme*. *RNA*, 2004. **10**: p. 151-158.
18. Lilley, D.M.J., *The hairpin and Varkud satellite ribozymes*, in *Ribozymes and RNA catalysis*, D.M.J. Lilley and F. Eckstein, Editors. 2008, Royal Society of Chemistry: Cambridge. p. 66-91.

19. Cochrane, J.C. and S.A. Strobel, *Catalytic strategies of self-cleaving ribozymes*. *Acc. Chem. Res.*, 2008. **41**(8): p. 1027-35.
20. Wilson, T.J. and D.M. Lilley, *Do the hairpin and VS ribozymes share a common catalytic mechanism based on general acid-base catalysis? A critical assessment of available experimental data*. *RNA*, 2011. **17**(2): p. 213-21.
21. Saville, B.L. and R.A. Collins, *RNA-mediated ligation of self-cleavage products of a Neurospora mitochondrial plasmid transcript*. *Proc. Natl. Acad. Sci. U. S. A.*, 1991. **88**: p. 8826-8830.
22. Collins, R.A. and J.E. Olive, *Reaction conditions and kinetics of self-cleavage of a ribozyme derived from Neurospora VS RNA*. *Biochemistry*, 1993. **32**: p. 2795-2799.
23. Beattie, T.L., J.E. Olive, and R.A. Collins, *A secondary-structure model for the self-cleaving region of Neurospora VS RNA*. *Proc. Natl. Acad. Sci. U. S. A.*, 1995. **92**: p. 4686-4690.
24. Beattie, T.L. and R.A. Collins, *Identification of functional domains in the self-cleaving Neurospora VS ribozyme using damage selection*. *J. Mol. Biol.*, 1997. **267**: p. 830-840.
25. Murray, J.B., et al., *The hammerhead, hairpin and VS ribozymes are catalytically proficient in monovalent cations alone*. *Chem. Biol.*, 1998. **5**: p. 587-595.
26. Sood, V.D., T.L. Beattie, and R.A. Collins, *Identification of phosphate groups involved in metal binding and tertiary interactions in the core of the Neurospora VS ribozyme*. *J. Mol. Biol.*, 1998. **282**: p. 741-750.
27. Hiley, S.L. and R.A. Collins, *Rapid formation of a solvent-inaccessible core in the Neurospora Varkud satellite ribozyme*. *EMBO J.*, 2001. **20**(19): p. 5461-5469.
28. Sood, V.D. and R.A. Collins, *Identification of the catalytic subdomain of the VS ribozyme and evidence for remarkable sequence tolerance in the active site loop*. *J. Mol. Biol.*, 2002. **320**: p. 443-454.
29. Maguire, J.L. and R.A. Collins, *Effects of cobalt hexammine on folding and self-cleavage of the Neurospora VS ribozyme*. *J. Mol. Biol.*, 2001. **309**: p. 45-56.
30. Rastogi, T., et al., *A long-range pseudoknot is required for activity of the Neurospora VS ribozyme*. *EMBO J.*, 1996. **15**(11): p. 2820-2825.
31. Andersen, A. and R.A. Collins, *Rearrangement of a stable RNA secondary structure during VS ribozyme catalysis*. *Mol. Cell*, 2000. **5**: p. 469-478.

32. Andersen, A.A. and R.A. Collins, *Intramolecular secondary structure rearrangement by the kissing interaction of the Neurospora VS ribozyme*. Proc. Natl. Acad. Sci U. S. A., 2001. **98**(14): p. 7730-7735.
33. Fernandez, I.S., et al., *Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome*. Cell, 2014. **157**(4): p. 823-31.
34. Bouchard, P. and P. Legault, *Structural insights into substrate recognition by the Neurospora Varkud satellite ribozyme: importance of U-turns at the kissing-loop junction*. Biochemistry, 2014. **53**(1): p. 258-69.
35. Wilson, T.J., A.C. McLeod, and D.M. Lilley, *A guanine nucleobase important for catalysis by the VS ribozyme*. EMBO J., 2007. **26**(10): p. 2489-500.
36. Desjardins, G., et al., *NMR structure of the A730 loop of the Neurospora VS ribozyme: insights into the formation of the active site*. Nucleic Acids Res., 2011. **39**(10): p. 4427-37.
37. Jones, F.D. and S.A. Strobel, *Ionization of a critical adenosine residue in the Neurospora Varkud satellite ribozyme active site*. Biochemistry, 2003. **42**: p. 4265-4276.
38. McLeod, A.C. and D.M. Lilley, *Efficient, pH-dependent RNA ligation by the VS ribozyme in trans*. Biochemistry, 2004. **43**(4): p. 1118-25.
39. Zhao, Z.Y., et al., *Nucleobase participation in ribozyme catalysis*. J. Am. Chem. Soc., 2005. **127**(14): p. 5026-7.
40. Smith, M.D. and R.A. Collins, *Evidence for proton transfer in the rate-limiting step of a fast-cleaving Varkud satellite ribozyme*. Proc. Natl. Acad. Sci. U. S. A., 2007. **104**(14): p. 5818-23.
41. Jaikaran, D., et al., *An important role of G638 in the cis-cleavage reaction of the Neurospora VS ribozyme revealed by a novel nucleotide analog incorporation method*. RNA, 2008. **14**(5): p. 938-49.
42. Smith, M.D., et al., *The ionic environment determines ribozyme cleavage rate by modulation of nucleobase pKa*. RNA, 2008. **14**(9): p. 1942-9.
43. Wilson, T.J., et al., *Nucleobase-mediated general acid-base catalysis in the Varkud satellite ribozyme*. Proc. Natl. Acad. Sci. U. S. A., 2010. **107**(26): p. 11751-6.

44. Lafontaine, D.A., D.G. Norman, and D.M. Lilley, *Structure, folding and activity of the VS ribozyme: importance of the 2-3-6 helical junction*. EMBO J., 2001. **20**(6): p. 1415-1424.
45. Lafontaine, D.A., D.G. Norman, and D.M. Lilley, *The global structure of the VS ribozyme*. EMBO J., 2002. **21**(10): p. 2461-2471.
46. Sood, V.D. and R.A. Collins, *Functional equivalence of the uridine turn and the hairpin as building blocks of tertiary structure in the Neurospora VS ribozyme*. J. Mol. Biol., 2001. **313**: p. 1013-1019.
47. Tyagi, R. and D.H. Mathews, *Predicting helical coaxial stacking in RNA multibranch loops*. RNA, 2007. **13**(7): p. 939-51.
48. Lipfert, J., et al., *The complete VS ribozyme in solution studied by small-angle X-ray scattering*. Structure, 2008. **16**(9): p. 1357-67.
49. Bonneau, E. and P. Legault, *NMR localization of divalent cations at the active site of the Neurospora VS ribozyme provides insights into RNA-metal-ion interactions*. Biochemistry, 2014. **53**(3): p. 579-90.
50. Rastogi, T. and R.A. Collins, *Smaller, faster ribozymes reveal the catalytic core of Neurospora VS RNA*. J. Mol. Biol., 1998. **277**: p. 215-224.
51. Milligan, J.F., et al., *Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates*. Nucleic Acids Res., 1987. **15**: p. 8783-8798.
52. Salvail-Lacoste, A., et al., *Affinity purification of T7 RNA transcripts with homogeneous ends using ARiBo and CRISPR tags*. RNA, 2013. **19**(7): p. 1003-14.
53. Nikonowicz, E.P., et al., *Preparation of ¹³C and ¹⁵N labelled RNAs for heteronuclear multidimensional NMR studies*. Nucleic Acids Res., 1992. **20**: p. 4507-4513.
54. Latham, M.P., et al., *NMR methods for studying the structure and dynamics of RNA*. Chembiochem, 2005. **6**(9): p. 1492-505.
55. Campbell, D.O. and P. Legault, *NMR structure of the Varkud satellite ribozyme stem-loop V RNA and magnesium-ion binding from chemical-shift mapping*. Biochemistry, 2005. **44**: p. 4157-4170.
56. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes*. J. Biomol. NMR, 1995. **6**: p. 277-293.

57. Greber, B.J., et al., *Architecture of the large subunit of the mammalian mitochondrial ribosome*. Nature, 2014. **505**(7484): p. 515-9.
58. Heus, H.A. and A. Pardi, *Structural features that give rise to the unusual stability of RNA hairpins containing GNRA tetraloops*. Science, 1991. **253**(5016): p. 191-194.
59. Jucker, F.M., et al., *A network of heterogeneous hydrogen bonds in GNRA tetraloops*. J. Mol. Biol., 1996. **264**(5): p. 968-80.
60. Kay, L.E., P. Keifer, and T. Saarinen, *Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity*. J. Am. Chem. Soc., 1992. **114**: p. 10663-10665.
61. Vuister, G.W. and A. Bax, *Resolution enhancement and spectral editing of uniformly ¹³C-enriched proteins by homonuclear broadband ¹³C decoupling*. J. Magn. Reson., 1992. **98**: p. 428-435.
62. Santoro, J. and G.C. King, *A constant-time 2D overbroaden experiment for inverse correlation of isotopically enriched species*. J. Magn. Reson., 1992. **97**: p. 202-207.
63. Schwieters, C.D., et al., *The Xplor-NIH NMR molecular structure determination package*. J. Magn. Reson., 2003. **160**: p. 66-74.
64. Lavery, R., et al., *Conformational analysis of nucleic acids revisited: Curves+*. Nucleic Acids Res., 2009. **37**(17): p. 5917-29.
65. Jucker, F.M. and A. Pardi, *Solution structure of the CUUG hairpin loop: a novel RNA tetraloop motif*. Biochemistry, 1995. **34**(44): p. 14416-27.
66. Klinge, S., et al., *Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6*. Science, 2011. **334**(6058): p. 941-8.
67. Campbell, D.O., et al., *NMR structure of Varkud satellite ribozyme stem-loop V in the presence of magnesium ions and localization of metal-binding sites*. Biochemistry, 2006. **45**(35): p. 10591-605.
68. Woese, C.R., S. Winker, and R.R. Gutell, *Architecture of ribosomal RNA: constraints on the sequence of "tetra-loops"*. Proc. Natl. Acad. Sci. U. S. A., 1990. **87**(21): p. 8467-71.
69. Wolters, J., *The nature of preferred hairpin structures in 16S-like rRNA variable regions*. Nucleic Acids Res., 1992. **20**(8): p. 1843-50.

70. Proctor, D.J., et al., *Isolation and characterization of a family of stable RNA tetraloops with the motif YNMG that participate in tertiary interactions*. *Biochemistry*, 2002. **41**(40): p. 12062-75.
71. Antao, V.P., S.Y. Lai, and I. Tinoco, Jr., *A thermodynamic study of unusually stable RNA and DNA hairpins*. *Nucleic Acids Res.*, 1991. **19**(21): p. 5901-5.
72. Baumruk, V., et al., *Comparison between CUUG and UUCG tetraloops: thermodynamic stability and structural features analyzed by UV absorption and vibrational spectroscopy*. *Nucleic Acids Res.*, 2001. **29**(19): p. 4089-96.
73. Petrov, A.I., C.L. Zirbel, and N.B. Leontis, *WebFR3D--a server for finding, aligning and analyzing recurrent RNA 3D motifs*. *Nucleic Acids Res.*, 2011. **39**(Web Server issue): p. W50-5.
74. Schuwirth, B.S., et al., *Structures of the bacterial ribosome at 3.5 Å resolution*. *Science*, 2005. **310**(5749): p. 827-34.
75. Pley, H.W., K.M. Flaherty, and D.B. McKay, *Model for an RNA tertiary interaction from the structure of an intermolecular complex between a GAAA tetraloop and an RNA helix*. *Nature*, 1994. **372**(6501): p. 111-3.
76. Costa, M. and F. Michel, *Frequent use of the same tertiary motif by self-folding RNAs*. *EMBO J.*, 1995. **14**(6): p. 1276-85.
77. Suydam, I.T. and S.A. Strobel, *Fluorine substituted adenosines as probes of nucleobase protonation in functional RNAs*. *J. Am. Chem. Soc.*, 2008. **130**(41): p. 13639-48.
78. Sood, V.D., S. Yekta, and R.A. Collins, *The contribution of 2'-hydroxyls to the cleavage activity of the Neurospora VS ribozyme*. *Nucleic Acids Res.*, 2002. **30**(5): p. 1132-1138.
79. Wimberly, B.T., et al., *Structure of the 30S ribosomal subunit*. *Nature*, 2000. **407**(6802): p. 327-39.
80. Klein, D.J., P.B. Moore, and T.A. Steitz, *The roles of ribosomal proteins in the structure assembly, and evolution of the large ribosomal subunit*. *J. Mol. Biol.*, 2004. **340**(1): p. 141-77.
81. Conn, G.L., et al., *A compact RNA tertiary structure contains a buried backbone-K⁺ complex*. *J. Mol. Biol.*, 2002. **318**(4): p. 963-73.

82. Weichenrieder, O., et al., *Structure and assembly of the Alu domain of the mammalian signal recognition particle*. Nature, 2000. **408**(6809): p. 167-73.
83. Kuglstatter, A., C. Oubridge, and K. Nagai, *Induced structural changes of 7SL RNA during the assembly of human signal recognition particle*. Nat. Struct. Biol., 2002. **9**(10): p. 740-4.
84. Scott, W.G., J.T. Finch, and A. Klug, *The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage*. Cell, 1995. **81**: p. 991-1002.
85. Golden, B.L., H. Kim, and E. Chase, *Crystal structure of a phage Twort group I ribozyme-product complex*. Nature Struct. & Mol. Biol., 2005. **12**(1): p. 82-88.
86. Wu, H. and J. Feigon, *H/ACA small nucleolar RNA pseudouridylation pockets bind substrate RNA to form three-way junctions that position the target U for modification*. Proc. Natl. Acad. Sci. U.S.A., 2007. **104**(16): p. 6655-60.
87. Pikovskaya, O., et al., *Structural principles of nucleoside selectivity in a 2'-deoxyguanosine riboswitch*. Nat. Chem. Biol., 2011. **7**(10): p. 748-55.
88. Huang, L., et al., *Long-range pseudoknot interactions dictate the regulatory response in the tetrahydrofolate riboswitch*. Proc. Natl. Acad. Sci. U.S.A., 2011. **108**(36): p. 14801-6.
89. Donghi, D., et al., *The structural stabilization of the kappa three-way junction by Mg(II) represents the first step in the folding of a group II intron*. Nucleic Acids Res., 2013. **41**(4): p. 2489-504.
90. Chapman, E.G., et al., *The structural basis of pathogenic subgenomic flavivirus RNA (sfRNA) production*. Science, 2014. **344**(6181): p. 307-10.
91. Strobel, S.A. and T.R. Cech, *Tertiary interactions with the internal guide sequence mediate docking of the P1 helix into the catalytic core of the Tetrahymena ribozyme*. Biochemistry, 1993. **32**(49): p. 13593-604.
92. Piotto, M., V. Saudek, and V. Skleňár, *Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions*. J. Biomol. NMR, 1992. **2**: p. 661-665.
93. Dingley, A.J. and S. Grzesiek, *Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide $^2J_{NN}$ couplings*. J. Am. Chem. Soc., 1998. **120**: p. 8293-8297.

94. Mueller, L., P. Legault, and A. Pardi, *Improved RNA structure determination by detection of NOE contacts to exchange-broadened amino groups*. J. Am. Chem. Soc., 1995. **117**: p. 11043-11048.
95. Simorre, J.P., et al., *Correlation of the guanosine exchangeable and non-exchangeable base protons in $^{13}\text{C}/^{15}\text{N}$ -labeled RNA with an HNC-TOCSY-CH experiment*. J. Biomol. NMR, 1996. **7**: p. 153-156.
96. Zhang, O., et al., *Backbone ^1H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques*. J. Biomol. NMR, 1994. **4**(6): p. 845-58.
97. Pascal, S.M., et al., *Simultaneous acquisition of ^{15}N and ^{13}C -edited NOE spectra of proteins dissolved in H_2O* . J. Magn. Res. Series B, 1994. **103**: p. 197-201.
98. Ikura, M., et al., *Three-dimensional NOESY-HMQC spectroscopy of a ^{13}C -labeled protein*. J. Magn. Reson., 1990. **86**: p. 204-209.
99. Skleñár, V., et al., *Optimization of triple-resonance HCN experiments for application to larger RNA oligonucleotides*. J. Magn. Reson., 1998. **130**: p. 119-124.
100. Pardi, A. and E.P. Nikonowicz, *Simple procedure for resonance assignment of the sugar protons in ^{13}C -labeled RNAs*. J. Am. Chem. Soc., 1992. **114**: p. 9202-9203.
101. Wishart, D.S., et al., *^1H , ^{13}C , ^{15}N chemical shift referencing in biomolecular NMR*. J. Biomol. NMR, 1995. **6**: p. 135-140.
102. Maderia, M., T.E. Horton, and V.J. DeRose, *Metal interactions with a GAAA RNA tetraloop characterized by ^{31}P NMR and phosphorothioate substitution*. Biochemistry, 2000. **39**: p. 8193-8200.
103. McFail-Isom, L., X. Shui, and L.D. Williams, *Divalent cations stabilize unstacked conformations of DNA and RNA by interacting with base pi systems*. Biochemistry, 1998. **37**(49): p. 17105-11.

TABLE**Table 4.1. Structural statistics for J345 and J345^{Mg}.**

	J345	J345 ^{Mg}
no. of experimental restraints		
NOE-derived distance restraints	1780	1780
inter-nucleotide	897	897
intra-nucleotide	855	855
ambiguous	28	28
hydrogen bond restraints	84	84
base pair planarity restraints	30	30
Mg ²⁺ -RNA distance restraints	0	92
dihedral angle restraints	161	161
total number of restraints	2055	2147
rmsd from experimental restraints		
NOE (Å) (none > 0.2)	0.0088 ± 0.0004	0.0111 ± 0.0002
dihedral (°) (none > 5°)	0.10 ± 0.01	0.11 ± 0.01
rmsd from idealized geometry		
bonds (Å)	0.00368 ± 0.00003	0.00373 ± 0.00003
angles (°)	0.9969 ± 0.0007	0.9866 ± 0.0007
impropers (°)	0.392 ± 0.001	0.396 ± 0.002
heavy-atom rmsd to the minimized average structure (Å)		
overall (residues 2-45)	3.72 ± 1.62	2.63 ± 0.80
stem III (residues 2-5 and 42-45)	0.35 ± 0.09	0.31 ± 0.08
stem-loop IV (residues 10-19)	0.36 ± 0.09	0.38 ± 0.09
stem-loop V (residues 24-35)	0.82 ± 0.31	0.55 ± 0.18
core domain (residues 6-9, 20-23 and 36-41)	0.97 ± 0.13	1.15 ± 0.16
U ₆ -U ₄₁ -C ₂₃ base triple	0.81 ± 0.19	0.78 ± 0.24
C ₇ -U ₄₀ -U ₂₂ base triple	0.68 ± 0.15	0.86 ± 0.22
C ₈ -G ₂₁ -A ₃₉ base triple	0.46 ± 0.13	0.69 ± 0.18
C ₉ -G ₂₀ -G ₃₈ base triple	0.66 ± 0.25	0.78 ± 0.23

FIGURES

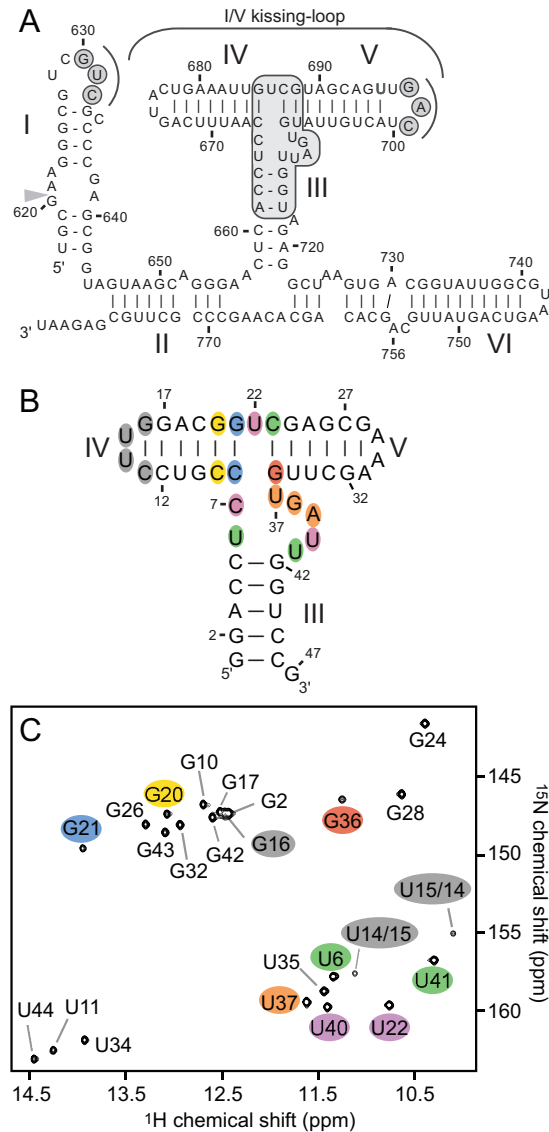


Figure 4.1. J345 RNA adopts a single stable structure in the presence of Mg^{2+} ions. (A) Sequence and proposed secondary structure of a *cis*-cleaving VS ribozyme (including residues 617-783).²³ The cleavage site is shown with a gray arrowhead. Gray shading is used to highlight the nucleotides involved in WC base pairs at the I-V kissing loop interaction as well as residues from the III-IV-V junction that are present in the J345 RNA. (B) Sequence and proposed secondary structure²³ of the 47-nucleotide J345 RNA used in this study. (C) 2D ^1H - ^{15}N HSQC spectrum of the ^{15}N -labeled J345 RNA in the presence of 5 mM MgCl_2 . In panels B and C, the shading of residues is color-coded according to structural elements present in the NMR structure (see Figure 4.3).

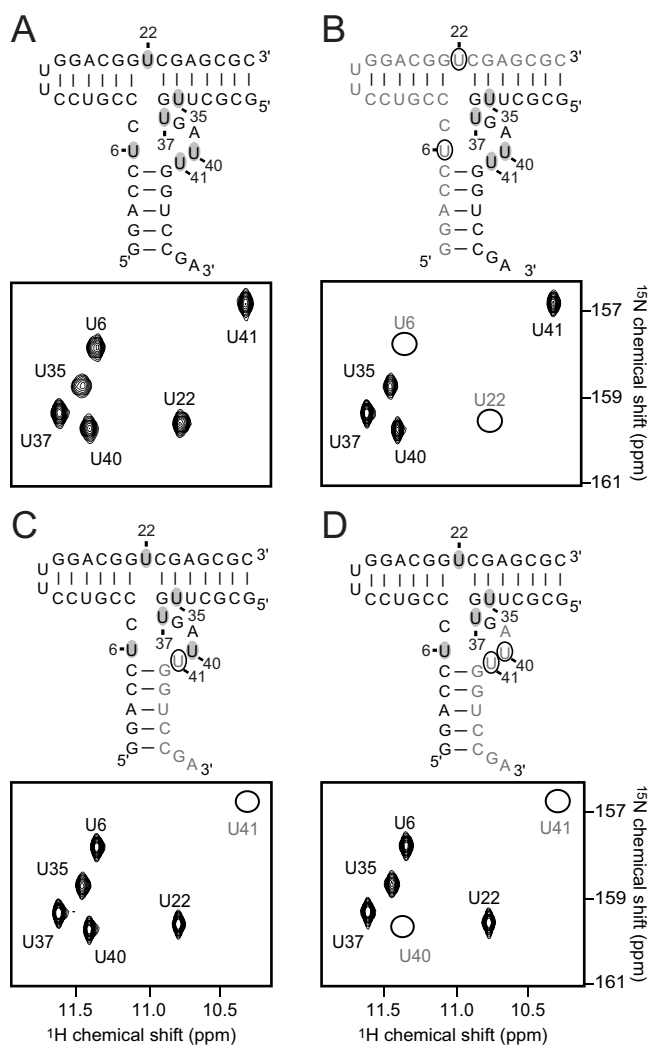


Figure 4.2. Unambiguous assignment of the uridine imino protons from the core of J345 using ^{15}N segmental labeling. Several complexes similar to the J345 RNA were formed between the S12 and S3 RNAs using different segmental ^{15}N labeling. The sequences and secondary structures of these complexes are shown along with a region of their 2D ^1H - ^{15}N HSQC spectrum: **(A)** [^{15}N]-S12/[^{15}N]-S3, **(B)** unlabeled S12/[^{15}N]-S3, **(C)** [^{15}N]-S12/[^{15}N]{30-40}-S3 and **(D)** [^{15}N]-S12/[^{15}N]{30-38}-S3. All RNA complexes are in the presence of 5 mM MgCl_2 . To illustrate the different ^{15}N -labeling schemes, ^{15}N -labeled residues are colored black and unlabeled residues gray. In addition, ^{15}N -labeled and unlabeled uridines from the core are highlighted with gray filled ovals and empty ovals, respectively. Peaks that are absent from the HSQC spectrum are also denoted with empty ovals. The signals of U_{14} and U_{15} are also observed in this region of the spectrum in panels **A**, **C**, and **D** as in Figure 4.1C, but are too weak to appear at the intensity level displayed here.

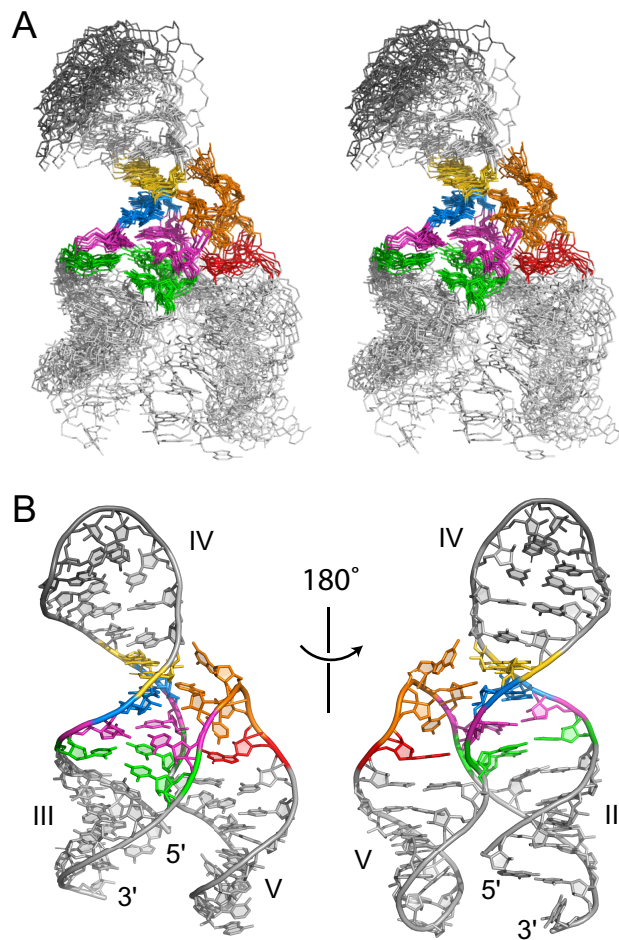


Figure 4.3. NMR solution structure of the J345 RNA. (A) Stereoview of the 10 lowest-energy structures. Only heavy atoms of core residues (residues 6-9, 20-23, and 36-41) were used for the superposition, but all residues are shown. (B) Stick representations of the minimized average structure of the J345 RNA. In panels A and B, residues are color-coded according to structural elements: the U₆-U₄₁-C₂₃ base triple (green), the C₇-U₄₀-U₂₂ base triple (magenta), the C₈-G₂₁ base pair (blue), the C₉-G₂₀ base pair (gold), the U-turn (residues U₃₇, G₃₈, and A₃₉) (orange), G₃₆ (red), and the CUUG tetraloop (residues C₁₃-G₁₆) (dark gray).

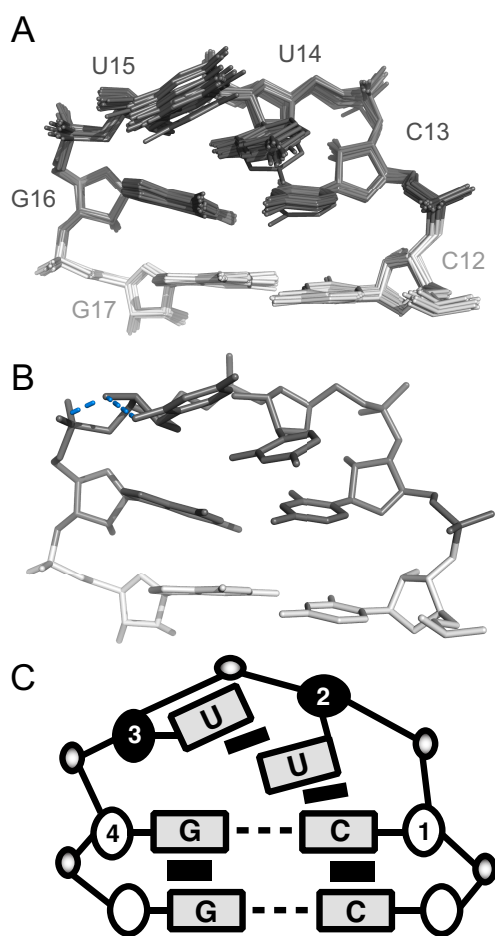


Figure 4.4. The cCUUGg tetraloop of the J345 RNA. (A) Superposition of the 20 lowest-energy structures. Only heavy atoms of residues 12-17 were used for the superposition and are shown. (B) Stick representation of the lowest-energy structure of the cCUUGg tetraloop. Cyan dashed lines represent hydrogen bonds that are defined on the basis of short distances between heavy atoms in the ensemble of structures: U₁₅ O2' and the 5'-phosphate oxygen of G₁₆ ($2.51 \pm 0.27 \text{ \AA}$) and U₁₅ O2' and U₁₅ O2 ($2.82 \pm 0.05 \text{ \AA}$). In panels A and B, the CUUG residues are colored dark gray, whereas those forming the closing base pair (C₁₂ and G₁₇) are colored light gray. (C) Schematic representation of the cCUUGg loop. The light gray boxes represent the bases, the large ovals the riboses in a C2'-endo (filled) or C3'-endo (empty) conformation, and the small ovals the phosphates. Dashed lines denote WC base pairs, and black rectangles indicate base stacking interactions.

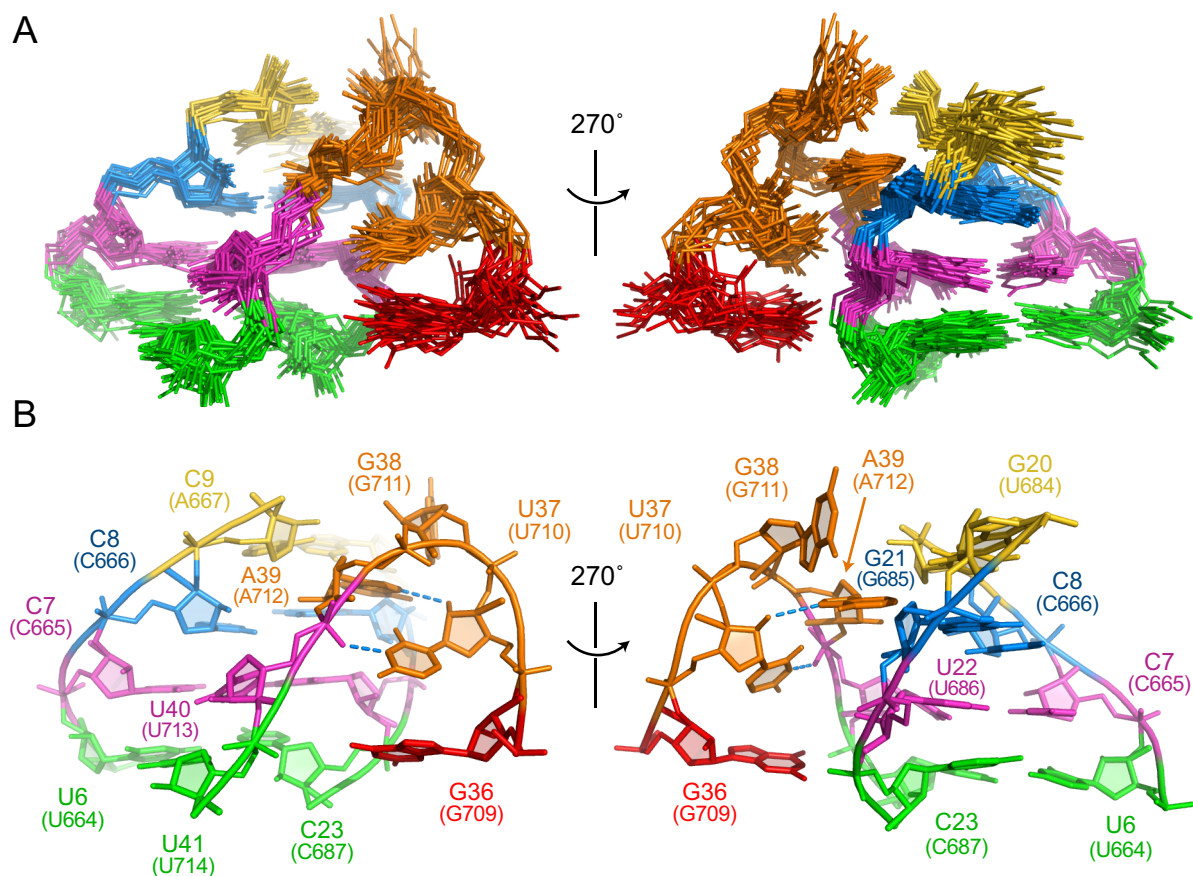


Figure 4.5. NMR structure of the core domain of the J345 RNA. (A) Superposition of the 20 lowest-energy structures. Only heavy atoms of core residues (6-9, 20-23, 36-41) were used for the superposition and are shown. (B) Stick representations of the lowest-energy structure of the core domain. The cyan dashed lines represent characteristic hydrogen bonds of the U-turn motif that are defined on the basis of short distances between heavy atoms in the ensemble of structures: U₃₇ O2' to A₃₉ N7 (3.12 ± 0.44 Å) and U₃₇ N3 to A₃₉ 3'-phosphate oxygen (19 of 20 structures, 3.08 ± 0.56 Å).

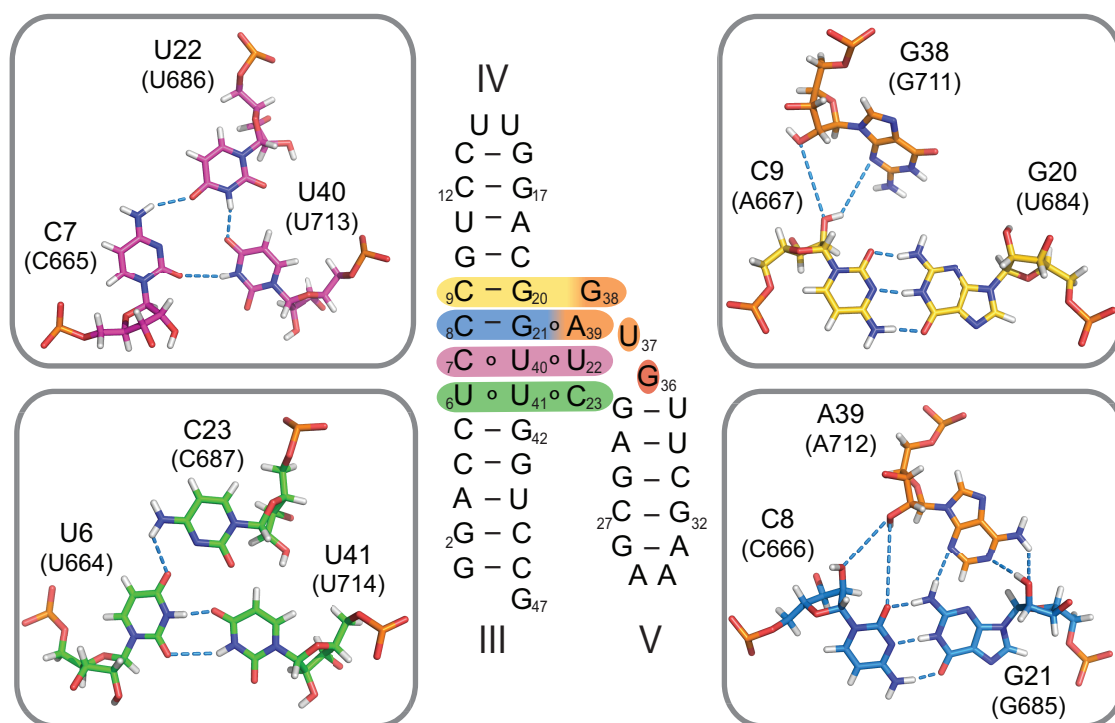


Figure 4.6. Base triples within the core domain of the J345 RNA. The four base triples are located on the secondary structure of J345 derived from the NMR structure (center), and their hydrogen bonding pattern is shown on the lowest-energy structure (sides). Cyan dashed lines represent hydrogen bonds that are defined on the basis of short distances (≤ 4.0 Å) between heavy atoms observed in the ensemble of structures.

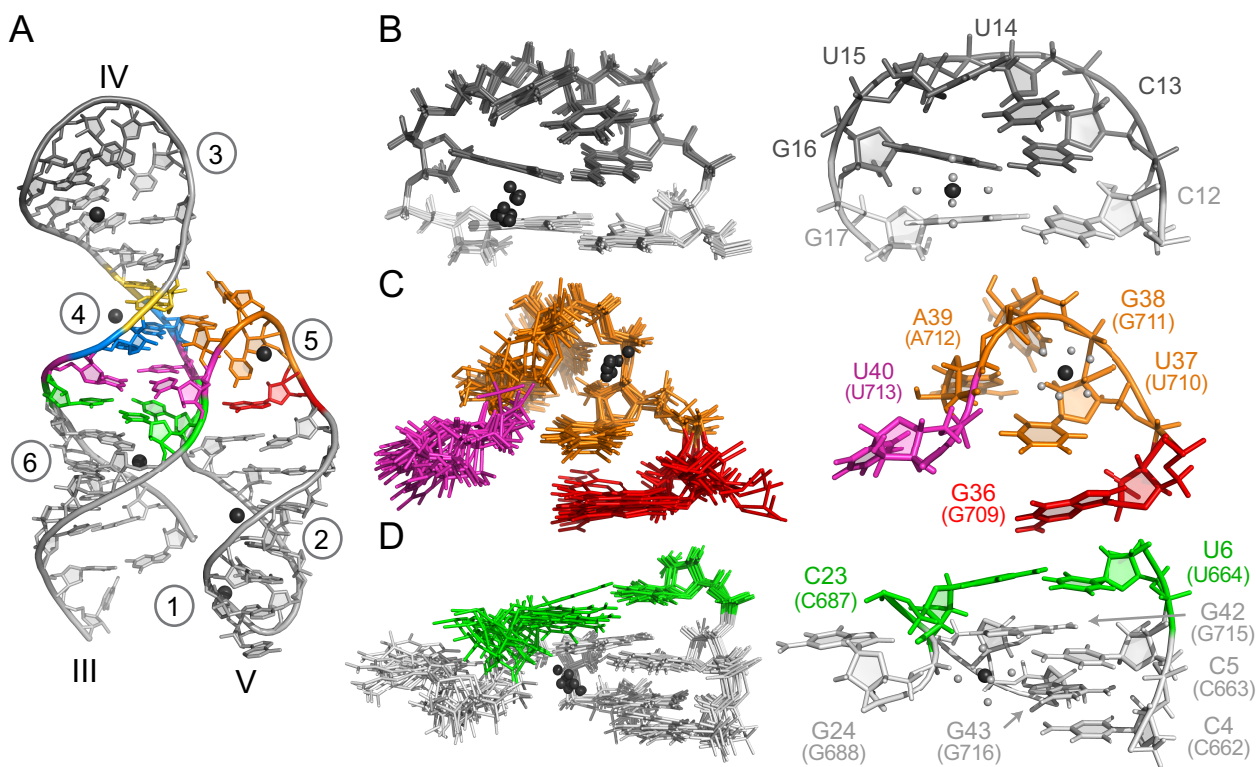


Figure 4.7. $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ -binding sites within the J345^{Mg} RNA. (A) Stick representation of the lowest-energy structure of J345^{Mg} with the six metal-binding sites, numbered 1-6. (B-D) $\text{Mg}(\text{H}_2\text{O})_6^{2+}$ -binding sites associated with (B) the cCUUGg tetraloop (site 3; in 20 of 20 structures), (C) the major groove face of the U-turn (site 5; in 16 of 20 structures) and (D) the core domain (site 6; in 20 of 20 structures). For each $\text{Mg}(\text{H}_2\text{O})_6^{2+}$ -binding site, the heavy atom superposition of the 10 lowest-energy structures is shown (left panels; only the residues shown were used for the superposition) along with a stick representation of the lowest-energy structure (right panels). The Mg^{2+} ions are colored black, with their bound water molecules colored dark gray.

SUPPORTING INFORMATION

Supporting Experimental Procedures

Segmental isotope labeling of S3 RNA. The segmentally-labeled $^{15}\text{N}\{30-38\}$ -S3 and $^{15}\text{N}\{30-40\}$ -S3 RNAs were produced by RNA ligation with the T4 DNA ligase, two S3 fragments and a single-stranded DNA splint of 19 nucleotides that is fully complementary to S3. The T4 DNA ligase was expressed with a C-terminal His₆ tag from the 3H11 vector in BL21(DE3) (generous gift from Dr. Thomas R. Cech) and purified by nickel-affinity chromatography and Q Sepharose anion-exchange FPLC [1]. The two S3 fragments used for production of $^{15}\text{N}\{30-38\}$ -S3 are ^{15}N -labeled S3₃₀₋₃₈ and unlabeled S3₃₉₋₄₈, whereas those used for production of $^{15}\text{N}\{30-40\}$ -S3 are ^{15}N -labeled S3₃₀₋₄₀ and unlabeled S3₄₁₋₄₈. To allow the desired ligation reaction and minimize ligation by-products, the acceptor RNAs (^{15}N -labeled S3₃₀₋₃₈ and S3₃₀₋₄₀) were transcribed with a 5'-triphosphate that was kept as is, whereas the donor RNAs (unlabeled S3₃₉₋₄₈ and S3₄₁₋₄₈) were chemically synthesized with a 5'-OH that was converted to a 5'-monophosphate using the T4 polynucleotide kinase (PNK). The RNA was incubated for 30 min at 37°C with 3 U of T4 PNK (New England Biolabs, MA, USA) per nmole of RNA in ligation buffer (70 mM Tris pH 7.6, 10 mM MgCl₂, 5 mM fresh filtered DTT and 6 mM ATP). The phosphorylation reaction was stopped by heat inactivation of the enzyme at 65°C for 5 min. The DNA splint was chemically synthesized (Integrated DNA Technologies, IA) and purified by denaturing gel electrophoresis.

The pre-ligation complexes were assembled in molar ratios of either 2:1:2 ^{15}N -S3₃₀₋₄₀:S3₄₁₋₄₈:splint or 2:1:2.5 ^{15}N -S3₃₀₋₃₈:S3₃₉₋₄₈:splint) in pre-ligation buffer (5 mM Tris-HCl pH 7.5, 5% (v/v) PEG 4000, 10 mM fresh filtered DTT, 10 mM MgCl₂) and annealed by heating at 70°C for 2 min and slowly cooling down at room temperature (RT) for 10 min. The formation of the ternary complexes was verified by native gel electrophoresis. The ligation reaction was performed at 37°C for 90 min with annealed pre-ligation complexes at a concentration of donor RNA of either 445 nmoles of S3₄₁₋₄₈ or 335 nmoles of S3₃₉₋₄₈ in pre-ligation buffer supplemented with 3 mM ATP, 50 U/mL RNAsin (Promega, WI, USA) and 100 U T4 DNA ligase per nmole of donor RNA. Both reactions yielded a 76% ligation rate with respect to the donor RNA. The resulting $^{15}\text{N}\{30-38\}$ -S3 and $^{15}\text{N}\{30-40\}$ -S3 RNAs were treated with CIP and purified as for the ^{15}N -labeled S3 RNA.

Preparation of S12/S3 complexes. The S12/S3 complexes were formed by annealing 0.4 mM S12 and 0.5 mM S3 strands in NMR buffer A (10 mM sodium cacodylate pH=6.5, 50 mM KCl, 0.05 mM NaN₃ in 90% H₂O/10% D₂O) by heating at 95°C for 2 min and slowly cooling down to RT for 10 min. Complex formation was verified by native gel electrophoresis and 2D ¹H-¹⁵N HSQC [2].

Native gel electrophoresis. Native polyacrylamide gel electrophoresis was carried out on 8% polyacrylamide gels [acrylamide:bisacrylamide (37.5:1)] in Tris-glycine buffer (25 mM Tris-200 mM glycine). The gels were pre-run for 30 min at 100 V, loaded with RNA samples [8 μL RNA at a concentration of 20 μM in 2 mM filtered DTT and 10% (v/v) glycerol] and then run for 2 h at 250 V and 4°C. The gels were stained with a Stains All solution (Sigma-Aldrich).

NMR data collection for resonance assignment and structural restraints. The exchangeable protons of the J345 RNA and their attached nitrogen were assigned at 15°C (Table S4.1 of Supporting Information) from the following experiments collected in 90% H₂O/10% D₂O: 1D ¹H flip-back watergate [3]; 2D imino- and amino-optimized ¹H-¹⁵N HSQC [2]; 2D HNN-COSY [4]; 2D ¹H-¹⁵N CPMG-NOESY [5]; G-specific H(NC)-TOCSY-(C)H [6]; 3D ¹⁵N-edited NOESY-HSQC (mixing time (τ_m) =150 ms; [7]) and 3D ¹⁵N/¹³C-edited NOESY (τ_m=150 ms; [8]) The non-exchangeable protons of the J345 RNA and their attached carbons were assigned at 25°C (Table S4.2 of Supporting Information) from the following experiments collected in 100% D₂O: 2D ¹H-¹³C constant-time-HSQC [9, 10], 2D ¹H-¹³C HMQC [11], 2D ¹H-¹⁵N MQ-(HC)N(C)H [12], 3D CT-HCCH-COSY [13], 3D HCCH-TOCSY [13] and a 3D ¹³C-edited HMQC-NOESY (τ_m=100 ms; [11]) ¹H, ¹³C and ¹⁵N chemical shifts were referenced at 25°C to an external standard of 2,2-dimethyl-2-silapentane-5-sulfonic acid at 0.00 ppm [14]. Due to exceptional chemical shift dispersion, almost complete assignments were obtained for the observable resonances (¹H, ¹⁵N and ¹³C) of the bases and the C1'-H1', C2'-H2' and C3'-H3' resonances of the ribose moieties (Tables S4.1 and S4.2 of Supporting Information).

A large set of restraints was obtained for structure calculation of the J345 RNA. An HNN-COSY [4] spectrum was collected in H₂O at 15°C to detect ²J_{NN} couplings across WC base pairs. Distance restraints were obtained from 2D ¹H-¹⁵N CPMG-NOESY [5] and 3D ¹⁵N-edited NOESY-HSQC [7] spectra collected in H₂O at 15°C, from a 3D ¹⁵N/¹³C-edited

NOESY-HSQC spectrum [8] collected in H₂O at 25°C and from a 3D ¹³C-edited HMQC-NOESY [11] spectrum recorded in D₂O at 25°C.

RNA-metal distance restraints. Manganese (Mn²⁺) titrations were performed by adding small volumes of a concentrated solution (0.5 M) of 99.99% MnCl₂ (Sigma-Aldrich, MO) directly to the RNA sample to achieve final concentrations of 5, 10, 20 and 40 μM MnCl₂. The first Mn²⁺ titration was carried out with a 1.25 mM ¹⁵N-labeled J345 sample in NMR buffer A with 5 mM MgCl₂ and 90%:10% H₂O:D₂O, and Mn²⁺-induced paramagnetic relaxation enhancement (PRE) was monitored by imino- and amino-optimized 2D ¹H-¹⁵N HSQC [2] spectra at 15°C. The second Mn²⁺ titration was carried out with a 2.0 mM ¹³C/¹⁵N-labeled J345 sample in NMR buffer A with 5 mM MgCl₂ and 100% D₂O, and Mn²⁺-induced PRE was monitored by 2D ¹H-¹³C CT-HSQC [9, 10] spectra at 25°C.

The RNA-metal distance restraints were derived from Mn²⁺-induced PRE using the ratio of signal intensity (I_o/I_{Mn}) determined from spectra collected at 0 μM MnCl₂ (I_o) and 10 μM MnCl₂ (I_{Mn}). Signal intensities were determined from the peak intensities of single cross sections of 2D spectra, but it was verified that the use of peak volumes would yield the same results. For each signal in the spectra with $I_o/I_{Mn} \geq 50$, $50 < I_o/I_{Mn} \leq 10$ and $3 \leq I_o/I_{Mn} < 10$, distance restraints were set to 1.8-5.0, 1.8-7.0 and 1.8-8.0 Å, respectively, between a defined Mg²⁺ ion and the pair of atoms giving rise to the observed signal (Table S4.3 of Supporting Information). When the identity of the Mg²⁺ ion could not be defined, ambiguous metal restraints were used with bounds of 1.8-10.0 Å (Table S4.3 of Supporting Information). An inner-sphere restraint of 1.8-2.2 Å was also applied between the non-bonded *pro-Sp* phosphate oxygen of A30 and the Mg²⁺ ion associated with the GAAA tetraloop based on previous studies [15, 16]. The effect of the Mn²⁺-induced PRE on the 5'/3'-end of stem III was not considered in this study [16].

Supporting Figures

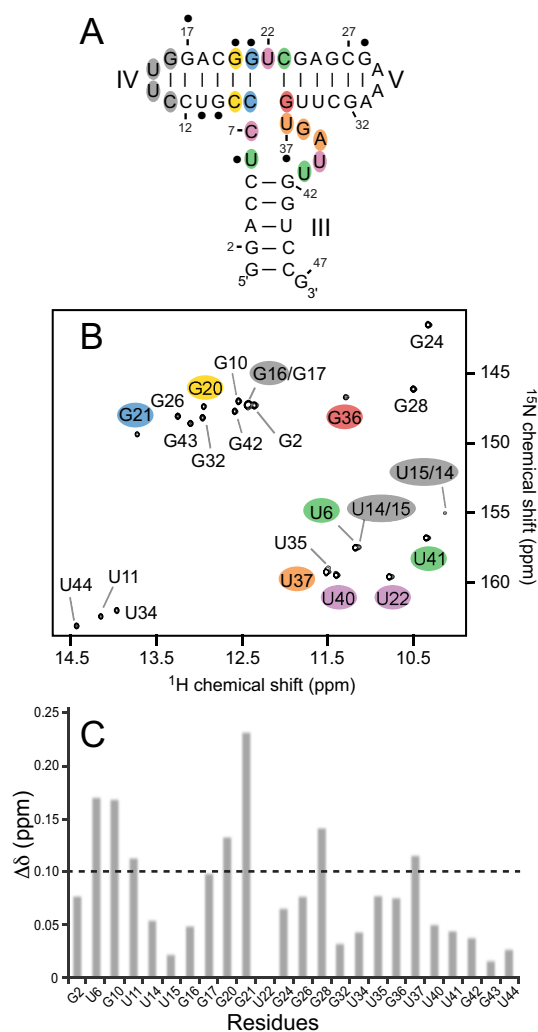


Figure S4.1. The J345 RNA adopts a single stable structure in the absence of Mg^{2+} ions that is similar to that formed in the presence of Mg^{2+} ions. (A) Sequence and proposed secondary structure [17] of the 47-nt J345 RNA used in this study. (B) 2D 1H - ^{15}N HSQC spectrum of the ^{15}N -labeled J345 RNA collected in the absence of Mg^{2+} ions. In (A) and (B) shading of residues is color-coded according to structural elements (Figure 4.3). (C) Histogram of the change in chemical shift [$\Delta\delta$ (ppm)] between spectra collected in the absence (B) and the presence of 5 mM $MgCl_2$ (Figure 4.1C). The value of $\Delta\delta$ was calculated according to the equation: $\Delta\delta = [(\Delta_H)^2 + (0.2 \times \Delta_N)^2]^{1/2}$, where Δ_H and Δ_N are the change in 1H and ^{15}N chemical shifts, respectively. Residues with $\Delta\delta \geq 0.1$ ppm are marked by a black dot in (A).

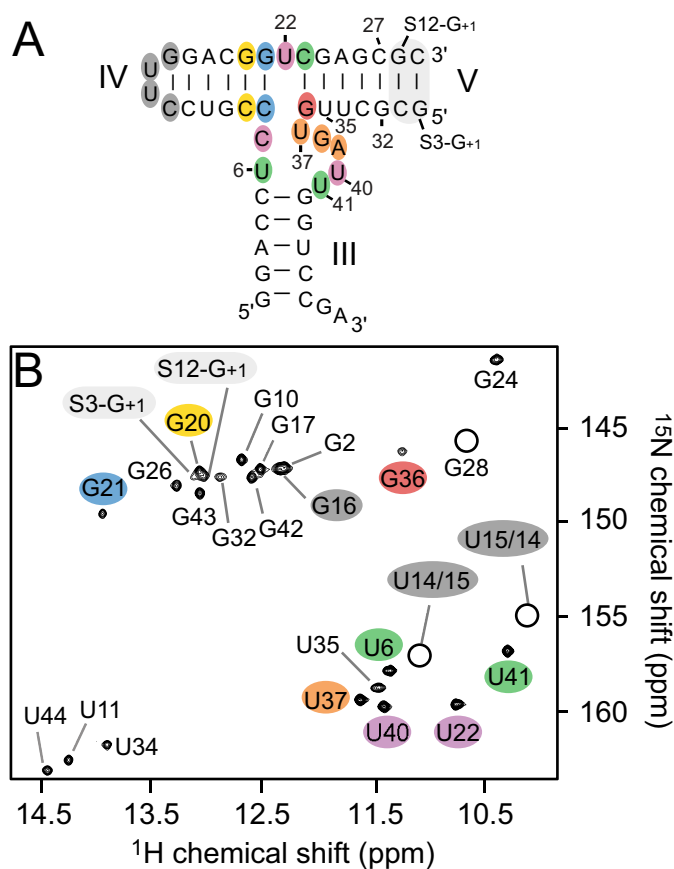


Figure S4.2. The S12/S3 complex adopts a structure similar to that of the J345 RNA. (A) Sequence and secondary structure [17] of the S12/S3 complex. **(B)** 2D ^1H - ^{15}N HSQC spectrum of the [^{15}N]-S12/[^{15}N]-S3 complex in the presence of 5 mM MgCl_2 . In **(A)** and **(B)** shading of residues is color-coded according to structural elements (see Figure 4.3). In addition, nucleotides present in the S12/S3 complex, but not in J345 are highlighted in light gray. Signals that are present in the 2D HSQC spectrum of J345 (Figure 4.1C), but do not appear in **(B)** are indicated with empty ovals. The signals of U_{14} and U_{15} are present in the HSQC spectrum shown in **(B)**, but are too weak to appear at the intensity level displayed here.

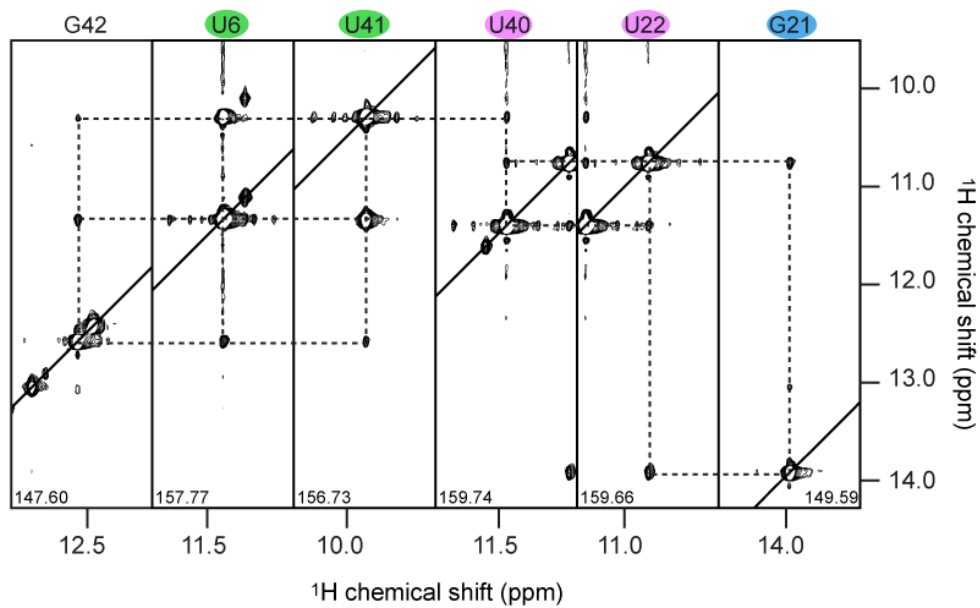


Figure S4.3. NMR evidence for a highly organized core and for helical stacking between stems III and IV. NOEs between the imino protons of the core residues in the 2D ^1H - ^1H slices taken at specific ^{15}N chemical shifts (reported in ppm, bottom left corner) of a 3D ^{15}N -edited NOESY-HSQC ($\tau_m = 150$ ms; [7]) The connections between the NOEs are shown by dashed lines.

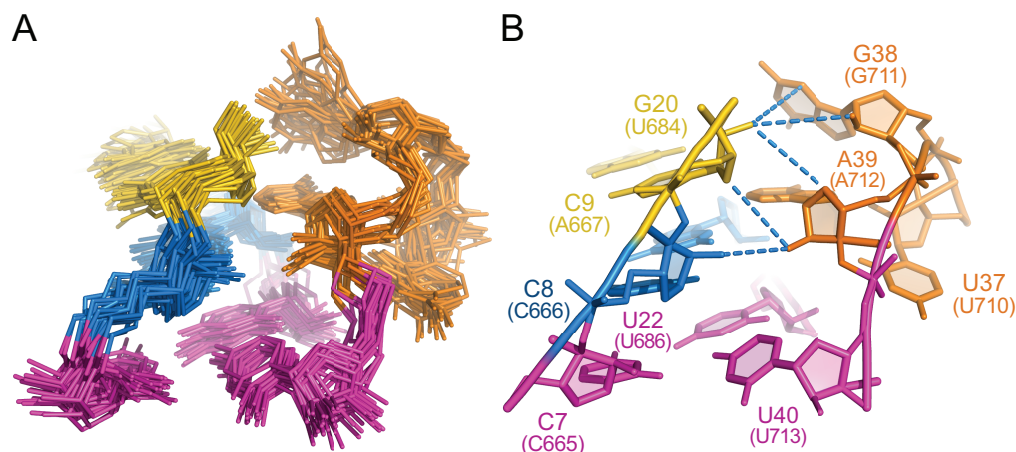


Figure S4.4. A ribose zipper within the core domain of the J345 RNA. (A) Heavy atom superposition of core residues encompassing the ribose zipper (7-9, 20-22, 37-40) within the 20 lowest-energy structures. (B) Stick representation of the same core residues from the minimized average structure. The cyan dashed lines represent hydrogen bonds defining the ribose zipper, as evidenced by short interatomic distances in the ensemble of structures: C₈ O2' to A₃₉ O2' (3.22 – 4.68 Å), C₉ O4' to A₃₉ O2' (3.20 – 5.00 Å), C₉ O2' to A₃₉ O4' (3.57 – 4.84 Å), C₉ O2' to G₃₈ O2' (4.03 – 4.98 Å) and C₉ O2' to G₃₈ N3 (3.87 – 4.96 Å).

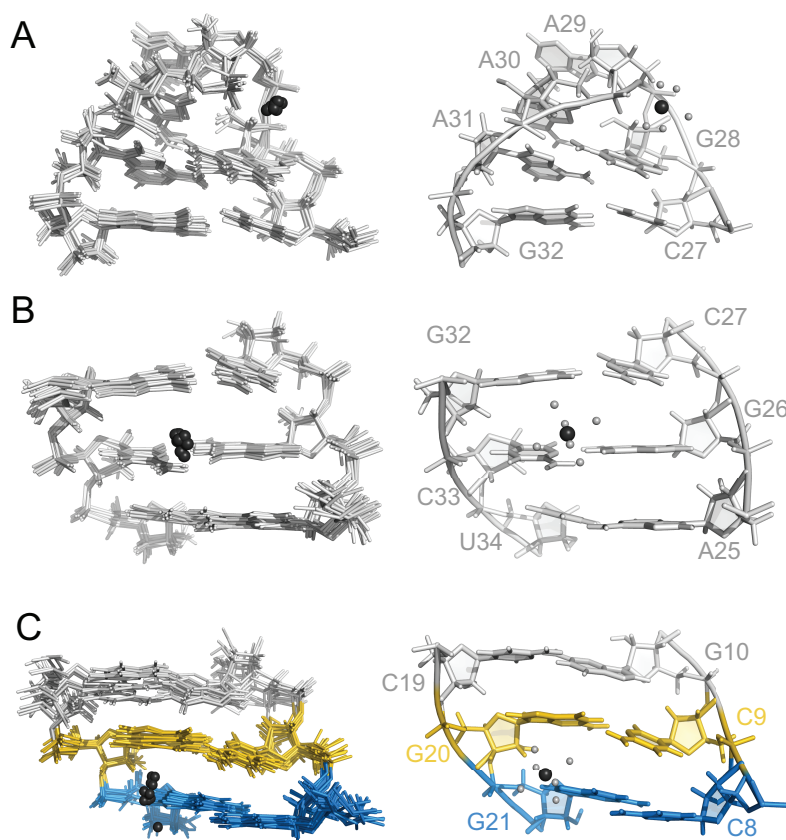


Figure S4.5. Additional $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ -binding sites within the J345^{Mg} RNA. One $\text{Mg}(\text{H}_2\text{O})_5^{2+}$ -binding site is associated with (A) the GAAA tetraloop (site 1; in 20/20 structures) and other $\text{Mg}(\text{H}_2\text{O})_6^{2+}$ -binding sites are associated with (B) the major groove of stem V (site 2; in 20/20 structures) and (C) the major groove of stem IV (site 4; in 14/20 structures). For each $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ -binding site, the superposition of the 10 lowest-energy structures is shown (left panels) along with a stick representation of the lowest-energy structure (right panels). The Mg^{2+} ions are shown in black with their bound water molecules in dark gray.

Supporting Tables

Table S4.1. Resonance assignment of exchangeable ^1H and their attached ^{15}N for the J345 RNA.

BASE EXCHANGEABLE PROTON CHEMICAL SHIFTS (ppm)								
± 0.02 ppm								
	H1	H21	H22	H3	H41	H42	H61	H62
G1								
G2	12.46							
A3								
C4					8.54	6.90		
C5					8.29	6.68		
U6				11.36				
C7					7.20	6.67		
C8					7.95	6.97		
C9					8.28	6.48		
G10	12.73							
U11				14.26				
C12					8.47	6.87		
C13					8.29	6.70		
U14				10.10/11.12				
U15				10.10/11.12				
G16	12.51							
G17	12.55							
A18								
C19					8.10	6.70		
G20	13.10							
G21	13.97	7.88	7.09					
U22				10.77				
C23								
G24	10.41							
A25								
G26	13.31							
C27					8.29	6.50		
G28	10.64							
A29							6.35	6.35
A30							6.33	6.33
A31								
G32	12.98							
C33					8.68	6.75		
U34				13.94				
U35				11.45				
G36	11.25							

U37				11.63				
G38								
A39							6.10	6.10
U40				11.41				
U41				10.30				
G42	12.62							
G43	13.12							
U44				14.46				
C45					8.51	6.93		
C46								
G47								
BASE NITROGEN CHEMICAL SHIFTS (ppm)								
± 0.1 ppm								
	N1	N2	N3	N4	N6			
G1								
G2	147.3							
A3								
C4				99.2				
C5				98.6				
U6			157.8					
C7				97.6				
C8				98.3				
C9				97.3				
G10	146.8							
U11			162.5					
C12				99.2				
C13				97.9				
U14			155.0/157.6					
U15			155.0/157.6					
G16	147.3							
G17	147.3							
A18								
C19				98.8				
G20	147.4							
G21	149.6	75.0						
U22			159.6					
C23								
G24	141.6							
A25								
G26	148.0							
C27				98.9				
G28	146.1							
A29					78.9			
A30					80.6			

A31								
G32	148.1							
C33				98.5				
U34			161.8					
U35			158.7					
G36	146.5							
U37			159.4					
G38								
A39					72.58			
U40			159.7					
U41			156.7					
G42	147.6							
G43	148.6							
U44			163.0					
C45				98.7				
C46								
G47								

Table S4.2. Resonance assignment of non-exchangeable ¹H and their attached ¹³C for the J345 RNA.

NON-EXCHANGEABLE PROTON CHEMICAL SHIFTS (ppm)										
± 0.01 ppm										
	H1'	H2'	H3'	H4'	H5'	H5''	H2	H5	H6	H8
G1	5.98	4.69	4.98	4.45	4.45	4.10				7.54
G2	5.84	4.59	4.69							7.52
A3	6.02	4.51	4.66				7.84			7.88
C4	5.46	4.28	4.39	4.40				5.19	7.50	
C5	5.51	4.57						5.45	7.61	
U6	5.60	4.61	4.34	4.42	4.56	4.12		5.58	7.76	
C7	5.36	3.87	4.58	4.33				5.91	8.24	
C8	4.82	4.15	4.58	4.27				5.64	7.95	
C9	6.21	4.73	4.72	4.67	4.59	4.17		5.45	7.99	
G10	5.65	4.39	4.68	4.34	4.53	4.11				7.56
U11	5.54	4.39	4.45	5.14	4.50	4.10		5.14	7.82	
C12	5.63	4.18	4.67	4.51				5.57	7.81	
C13	5.56	4.38	4.41					5.69	7.65	
U14	5.91	4.38		4.16	4.15	3.96		5.79	7.98	
U15	5.64	4.15	4.52	4.06	3.72	3.39		5.66	7.54	
G16	5.88	4.88	4.62	4.45	4.13	3.97				8.12
G17	5.98	4.82	5.38	4.43	4.55	4.34				8.21
A18	6.01	4.44	4.72				7.77			8.01
C19	5.48	4.42	4.49	4.40	4.54	4.09		5.17	7.57	
G20	5.77	4.78	4.51	4.32	4.51	4.13				7.55
G21	5.97	4.58	4.36	4.57	4.58	4.08				7.24
U22	3.69	4.12	3.70	3.69	4.31	3.82		5.08	6.93	
C23	5.59	4.65	4.38	4.45	4.49	4.18		5.66	7.81	
G24	5.84	4.83	4.68	4.49	4.63	4.26				8.12
A25	5.98	4.65	4.69	4.57	4.51	4.18	7.51			7.68
G26	5.78	4.39	4.51	4.51	4.51	4.10				7.29
C27	5.64	4.47	4.50	4.39	4.54	4.13		5.10	7.45	
G28	5.77	4.32	4.70	4.42	4.55	4.10				7.55
A29	5.67	4.80	4.46	4.24	4.30	3.95	7.74			8.35
A30	5.39	4.36	4.62	4.36	4.34	4.20	7.74			7.94
A31	5.98	4.65	4.93	4.46			8.21			8.22
G32	3.41	4.25	4.14	4.32	4.33	4.23				7.95
C33	5.56	4.44	4.44	4.26	4.58	4.14		5.20	7.72	
U34	5.86	4.77	4.23	4.41	4.57	4.12		5.46	7.89	
U35	5.37	4.36	4.52	4.35	4.31			5.86	7.92	
G36	5.57	4.26	4.68	4.41	4.55	4.17				7.84
U37	5.89	5.18	4.60	4.43				5.17	7.58	
G38	5.45	3.64	4.71	4.20	4.26	4.03				8.19

A39	6.22	4.79	5.33	4.18			9.07			8.51
U40	5.56	4.80	4.70	4.51	4.63	4.16		6.12	8.52	
U41	5.64	4.70	4.36	4.65	4.41	4.33		6.04	7.92	
G42	5.99	4.80	4.68	4.57	4.57	4.26				7.93
G43	5.87	4.58	4.40	4.60	4.45	4.15				7.34
U44	5.66	4.40	4.57					5.13	7.96	
C45	5.62	4.43	4.44					5.49	7.63	
C46	5.45	4.43						5.69	7.91	
G47	5.74	4.05	4.24	4.20	4.40	4.05				7.62
CARBON CHEMICAL SHIFTS (ppm)										
± 0.2 ppm										
	C1'	C2'	C3'	C4'	C5'	C2	C5	C6	C8	
G1	91.8	84.7	78.5	86.2	65.9				138.0	
G2	92.9	75.4	72.7						136.8	
A3	92.7	75.6	72.5			154.3			139.3	
C4	93.5	75.7	72.3	81.9			97.2	140.7		
C5	94.0	75.5					98.1	141.1		
U6	94.7	75.2	71.9	83.0	64.4		104.8	140.6		
C7	93.9	76.2	71.7	82.0			98.5	141.9		
C8	93.3	75.6	71.7	82.1			97.2	141.8		
C9	93.7	74.6	71.7	82.0	64.3		97.6	141.4		
G10	91.9	75.6	72.5	82.0	64.8				136.2	
U11	93.4	75.4	72.4		64.2		102.5	142.2		
C12	93.6	77.0	72.6	82.1			97.6	141.3		
C13	93.9	76.8	72.3				98.4	141.5		
U14	90.4	75.1		85.3	66.6		104.8	143.8		
U15	89.6	76.4	78.7	85.1	67.3		104.8	143.5		
G16	93.1	75.2	73.9	84.4	62.6				138.4	
G17	91.5	76.6	76.2	84.0	69.8				138.8	
A18	92.7	75.5	72.5			154.3			139.7	
C19	93.0	75.5	72.5	81.3	64.3		97.8	139.9		
G20	92.3	75.4	73.3	83.5	65.5				135.9	
G21	93.7	74.7	72.9	82.0	65.0				136.5	
U22	94.5	73.9	72.5	81.3	64.9		102.9	140.9		
C23	93.5	74.9	72.3	82.1	64.7		98.0	141.5		
G24	92.9	75.4	72.1	82.6	65.1				138.8	
A25	93.1	75.6	72.6	81.7	65.0	153.5			139.3	
G26	92.7	75.2	72.1	81.9	64.9				135.8	
C27	93.4	76.0	71.7	81.9	64.7		98.0	139.5		
G28	93.9	75.9	71.7	81.8	65.0				136.5	
A29	93.3	75.4	72.1	83.0	63.5	153.9			142.2	
A30	93.1	75.6	73.1	83.0	69.0	154.5			139.9	
A31	91.8	76.2	72.9	81.9		155.8			140.1	
G32	93.1	74.7	74.6	82.2	69.4				137.3	

C33	93.9	75.4	72.2	82.6	64.9		96.8	141.5		
U34	91.0	75.7	72.3	80.0	64.2		103.6	141.8		
U35	94.3	75.6	72.9	82.8	64.4		105.5	140.9		
G36	92.8	75.4	71.7	82.0	64.8				137.2	
U37	94.3	74.6	71.7	81.8			103.4	141.5		
G38	93.8	76.8	72.1	83.2	63.8				141.1	
A39	91.9	78.2	70.4	82.2		157.8			142.6	
U40	93.7	76.0	72.5	82.7	64.7		104.9	145.2		
U41	94.7	75.6	74.3	82.9	68.6		104.2	144.0		
G42	92.1	76.2	73.5	82.0	65.8				136.8	
G43	93.7	75.2	73.9	82.6	67.3				136.8	
U44	93.8	75.4	71.7				102.3	142.4		
C45	93.7	74.8	72.2				98.2	140.5		
C46	94.1	75.5					97.6	141.7		
G47	91.7	77.6	70.2	84.0	65.5				137.6	

Table S4.3. Distance restraints derived from Mn²⁺-induced PRE used for the J345^{Mg} structure calculation.

Site	Atoms	I/Io	Restraint (Å)
Site 1	G28 N1-H1	63	1.8-5.0
	A30 C8-H8	18	1.8-7.0
	G28 C2'-H2'	3	1.8-8.0
	G28 C8-H8	4	1.8-8.0
Site 1 or 2	G32 C8-H8	14	1.8-10.0
	C27 N4-H41	7	1.8-10.0
	C27 N4-H42	8	1.8-10.0
	C27 C6-H6	8	1.8-10.0
	C27 C5-H5	6	1.8-10.0
	C27 C3'-H3'	3	1.8-10.0
	C27 C2'-H2'	3	1.8-10.0
	G32 C5'-H5'	3	1.8-10.0
	A31 C3'-H3'	3	1.8-10.0
Site 2	G32 C8-H8	14	1.8-7.0
	G26 C8-H8	3	1.8-8.0
	C33 C5-H5	3	1.8-8.0
Site 2 or 6	G24 C8-H8	13	1.8-10.0
	U35 N3-H3	3	1.8-10.0
	A25 C8-H8	6	1.8-10.0
Site 3	U14 N3-H3	15	1.8-7.0
	U15 N3-H3	15	1.8-7.0
	G16 C8-H8	13	1.8-7.0
	G17 C8-H8	7	1.8-8.0
Site 3 or 4	G10 C8-H8	7	1.8-10.0
Site 4	G21 C8-H8	16	1.8-7.0
	C9 N4-H42	5	1.8-8.0
	C9 C5-H5	3	1.8-8.0
	G20 C8-H8	3	1.8-8.0
	G21 C3'-H3'	3	1.8-8.0
Site 5	G38 C8-H8	11	1.8-7.0
	U37 C6-H6	4	1.8-8.0
	U37 C5-H5	4	1.8-8.0
	U37 N3-H3	3	1.8-8.0
	G38 C5'-H5''	3	1.8-8.0
Site 5 or 6	G36 N1-H1	4	1.8-10.0
	G36 C8-H8	3	1.8-10.0
Site 6	G42 C8-H8	12	1.8-7.0
	G43 C8-H8	14	1.8-7.0
	U41 C6-H6	7	1.8-8.0
	U41 C5-H5	6	1.8-8.0
	U41 C2'-H2'	6	1.8-8.0
	G42 C3'-H3'	6	1.8-8.0
	U41 N3-H3	6	1.8-8.0
	G42 N1-H1	5	1.8-8.0
	U6 N3-H3	4	1.8-8.0
	C23 C2'-H2'	3	1.8-8.0
	U41 C3'-H3'	3	1.8-8.0

Table S4.4. Structural characteristics of $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ -binding sites in J345^{Mg}.

Site	Inner-sphere ligands ^a	Outer-sphere ligands ^b	Electrostatic interactions ^c	Cation- π interactions ^d
Site 1 (n=5)	A_{30} 5'- <i>pro</i> -S _p OP	A_{30} 5'- <i>pro</i> -R _p OP A_{30} O5' A_{30} O3'	A_{29} 5'-PO ₄ ⁻ A_{30} 5'-PO ₄ ⁻ A_{31} 5'-PO ₄ ⁻	G ₂₈
Site 2 (n=6)		G ₂₆ O6 G ₃₂ N7 G ₃₂ O6 U ₃₄ O4	G ₃₂ 5'-PO ₄ ⁻ C ₃₃ 5'-PO ₄ ⁻	
Site 3 (n=6)		G ₁₆ N7 G ₁₆ O6 G ₁₇ N7 G ₁₇ O6	G ₁₆ 5'-PO ₄ ⁻	
Site 4 (n=6)		G ₂₀ N7 G ₂₁ N7 G ₂₁ O6	G ₂₀ 5'-PO ₄ ⁻ C ₂₁ 5'-PO ₄ ⁻	
Site 5 (n=6)		G ₃₈ 5'- <i>pro</i> -R _p OP G ₃₈ O5' A ₃₉ 5'- <i>pro</i> -R _p OP A ₃₉ 5'- <i>pro</i> -S _p OP A ₃₉ O5'	U ₃₇ 5'-PO ₄ ⁻ G ₃₈ 5'-PO ₄ ⁻ A ₃₉ 5'-PO ₄ ⁻ U ₄₀ 5'-PO ₄ ⁻	U ₃₇
Site 6 (n=6)		C ₂₃ O2' G ₂₄ 5'- <i>pro</i> -R _p OP G ₂₄ 5'- <i>pro</i> -S _p OP G ₄₂ N7 G ₄₂ O6 G ₄₃ N7 G ₄₃ O6	G ₂₄ 5'-PO ₄ ⁻ G ₄₂ 5'-PO ₄ ⁻	

^aInner-sphere ligands were inferred based on a previous study [16].

^bOuter-sphere ligands were determined by selecting all hydrogen bond acceptors within 4.0 Å of a $\text{Mg}(\text{H}_2\text{O})_n^{2+}$ oxygen and in a favorable orientation for outer-sphere interaction in at least 25% of the lowest-energy structures presenting an occupancy at a site.

^cElectrostatic interactions were determined by selecting 5'-PO₄⁻ within 8.0 Å of a Mg²⁺ ion.

^dCation- π interactions were determined for Mg²⁺ ions within 5.2 Å and at an angle smaller than 52° from the ring centroid [18].

References (for Supplementary Information only)

1. Strobel, S.A. and T.R. Cech, *Tertiary interactions with the internal guide sequence mediate docking of the P1 helix into the catalytic core of the Tetrahymena ribozyme*. *Biochemistry*, 1993. **32**(49): p. 13593-604.
2. Kay, L.E., P. Keifer, and T. Saarinen, *Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity*. *J. Am. Chem. Soc.*, 1992. **114**: p. 10663-10665.
3. Piotto, M., V. Saudek, and V. Skleňár, *Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions*. *J. Biomol. NMR*, 1992. **2**: p. 661-665.
4. Dingley, A.J. and S. Grzesiek, *Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide $^2J_{NN}$ couplings*. *J. Am. Chem. Soc.*, 1998. **120**: p. 8293-8297.
5. Mueller, L., P. Legault, and A. Pardi, *Improved RNA structure determination by detection of NOE contacts to exchange-broadened amino groups*. *J. Am. Chem. Soc.*, 1995. **117**: p. 11043-11048.
6. Simorre, J.P., et al., *Correlation of the guanosine exchangeable and non-exchangeable base protons in $^{13}C/^{15}N$ -labeled RNA with an HNC-TOCSY-CH experiment*. *J. Biomol. NMR*, 1996. **7**: p. 153-156.
7. Zhang, O., et al., *Backbone 1H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques*. *J. Biomol. NMR*, 1994. **4**(6): p. 845-58.
8. Pascal, S.M., et al., *Simultaneous acquisition of ^{15}N and ^{13}C -edited NOE spectra of proteins dissolved in H_2O* . *J. Magn. Res. Series B*, 1994. **103**: p. 197-201.
9. Vuister, G.W. and A. Bax, *Resolution enhancement and spectral editing of uniformly ^{13}C -enriched proteins by homonuclear broadband ^{13}C decoupling*. *J. Magn. Reson.*, 1992. **98**: p. 428-435.
10. Santoro, J. and G.C. King, *A constant-time 2D overboderhausen experiment for inverse correlation of isotopically enriched species*. *J. Magn. Reson.*, 1992. **97**: p. 202-207.

11. Ikura, M., et al., *Three-dimensional NOESY-HMQC spectroscopy of a ^{13}C -labeled protein*. J. Magn. Reson., 1990. **86**: p. 204-209.
12. Skleňár, V., et al., *Optimization of triple-resonance HCN experiments for application to larger RNA oligonucleotides*. J. Magn. Reson., 1998. **130**: p. 119-124.
13. Pardi, A. and E.P. Nikonowicz, *Simple procedure for resonance assignment of the sugar protons in ^{13}C -labeled RNAs*. J. Am. Chem. Soc., 1992. **114**: p. 9202-9203.
14. Wishart, D.S., et al., *^1H , ^{13}C , ^{15}N chemical shift referencing in biomolecular NMR*. J. Biomol. NMR, 1995. **6**: p. 135-140.
15. Maderia, M., T.E. Horton, and V.J. DeRose, *Metal interactions with a GAAA RNA tetraloop characterized by ^{31}P NMR and phosphorothioate substitution*. Biochemistry, 2000. **39**: p. 8193-8200.
16. Bonneau, E. and P. Legault, *NMR localization of divalent cations at the active site of the Neurospora VS ribozyme provides insights into RNA-metal-ion interactions*. Biochemistry, 2014. **53**(3): p. 579-90.
17. Beattie, T.L., J.E. Olive, and R.A. Collins, *A secondary-structure model for the self-cleaving region of Neurospora VS RNA*. Proc. Natl. Acad. Sci. U. S. A., 1995. **92**: p. 4686-4690.
18. McFail-Isom, L., X. Shui, and L.D. Williams, *Divalent cations stabilize unstacked conformations of DNA and RNA by interacting with base pi systems*. Biochemistry, 1998. **37**(49): p. 17105-11.

5. Détermination de la structure RMN de la jonction II-III-VI et identification d'une interaction tertiaire critique

Mise en contexte

L'étude RMN du Chapitre 4 s'est intéressée à la structure de la jonction III-IV-V, un domaine architectural essentiel du ribozyme VS de *Neurospora*. Dans cette quatrième étude, nous nous sommes intéressés au dernier domaine architectural dont la structure à haute résolution n'était pas disponible, soit la jonction II-III-VI. L'ARN J236, long de 62 nucléotides, est formé de trois hélices, chacune possédant un renflement. Les études par spectroscopie RMN de la jonction II-III-VI ont démontré que cet ARN est dépendant de la présence d'ions Mg^{2+} pour se replier. De plus, la mutation du renflement de la tige II (A_{652}) affecte d'une façon significative le repliement de l'ARN J236, suggérant que ce renflement joue un rôle important dans le repliement de la jonction II-III-VI. La structure à haute résolution de la jonction II-III-VI a été obtenue par spectroscopie de RMN et est en accord avec l'ensemble des données biochimiques. Le cœur de la jonction II-III-VI est constitué de deux triplets de bases ainsi que de la paire de base *cis* WC/WC A-G. De plus, nous avons identifié la formation d'une interaction tertiaire entre les renflements des tiges II et VI, caractérisée par un motif *S-turn* et une paire de bases *cis Sugar edge*/WC A-A qui n'avait pas été identifiée précédemment. Finalement, la localisation des ions divalents a été réalisée par l'approche de relaxation paramagnétique induite par le manganèse (*Mn²⁺-induced paramagnetic relaxation enhancement*). Ainsi, deux sites de liaison d'ions Mg^{2+} ont été identifiés dans le cœur de la jonction tandis que trois ions Mg^{2+} se lient à proximité de l'interaction tertiaire entre les tiges II et VI, en accord avec la dépendance au Mg^{2+} de la jonction II-III-VI. Finalement, cette structure semble correspondre à la structure pré-catalytique de la jonction II-III-VI, permettant de former une plateforme d'interaction qui favorise la formation du site actif en positionnant correctement la boucle A730.

Dans ce quatrième article, j'ai contribué avec le Dr. Pascale Legault à la conception du projet. J'ai réalisé plusieurs synthèses d'échantillons avec l'aide de Dominique Chaussé, une technicienne du laboratoire, ainsi que d'Aurélie Tréfier, une stagiaire française au 1er cycle que j'ai eu l'honneur de diriger. De plus, la collecte et l'analyse des données de RMN, le calcul et l'analyse des structures ainsi que la rédaction de l'article ont été réalisées par le Dr. Pascale Legault et moi-même.

The NMR structure of the II-III-VI three-way junction from the
Neurospora VS ribozyme reveals a critical long-range tertiary
interaction

Eric Bonneau and Pascale Legault*

Département de Biochimie et Médecine Moléculaire, Université de Montréal, C.P. 6128,
Succursale Centre-Ville, Montréal, QC, Canada, H3C 3J7

This work was prepared for future submission in RNA.

ABSTRACT

As part of a global effort to structurally characterize the *Neurospora* VS ribozyme, NMR solution structures of several subdomains have been determined, including the internal loops of domains I and VI, a stem-loop derived from domain V, the I/V kissing-loop interaction and the III-IV-V junction. Here, we expand this work by determining the NMR structure of J236, a 62-nt RNA that encompasses the VS ribozyme II-III-VI three-way junction and its adjoining stems. Furthermore, we localize Mg^{2+} -binding sites within the J236 structure using Mn^{2+} -induced paramagnetic line broadening. The NMR structure of the J236 RNA displays a family C topology with a compact core stabilized by continuous stacking of stems II and VI, a *cis* WC-WC G•A base pair, two base triples and Mg^{2+} -ion binding. Moreover, it reveals a long-range interaction between the adenine bulges of stems II and VI. Complementary NMR studies demonstrate that both this bulge-bulge interaction and Mg^{2+} ions are critical for the stable folding of the II-III-VI junction. The NMR structure of J236 is in general agreement with biochemical data on the VS ribozyme, but not with biophysical studies performed with a minimal II-III-VI junction that does not incorporate the II-VI bulge-bulge interaction. Furthermore, it allows us to propose that a bulged out adenine in the II-III-VI junction (A_{657}) forms a tertiary interaction with another domain of the VS ribozyme to enhance cleavage activity. Together with previous NMR studies, our findings provide significant new insights into the overall three-dimensional structure of this unique ribozyme.

INTRODUCTION

The Varkud Satellite (VS) ribozyme is a member of the family of small nucleolytic ribozymes that also contains the hairpin, hammerhead, hepatitis delta virus, *glms* and twister ribozymes [1-6]. Like several members of this family, the VS ribozyme is derived from a satellite RNA, the non-coding VS RNA, found in the mitochondria of certain natural isolates of *Neurospora* filamentous fungi [1]. It catalyzes self-cleavage and self-ligation at a specific phosphodiester bond, and both of these transesterification reactions are critical for the replication cycle of the VS RNA [1, 7]. *In vitro*, these reactions require the presence of metal cations, which are important for both the structure and the chemical reaction [8-16]. The minimal contiguous VS ribozyme is composed of six helical domains, numbered I-VI [9]. Stem-loop I (SLI) forms the substrate domain since its internal loop contains the cleavage site. Helical domains II-VI constitute the catalytic domain, also termed the *trans*-acting ribozyme, which is organized around two three-way junctions: the II-III-VI and the III-IV-V junctions (Figure 5.1A; [9]). The SLI substrate is recognized by stem-loop V (SLV) through the formation of a highly stable kissing-loop interaction, which has been extensively characterized [17-21]. In particular, this kissing-loop interaction is associated with a conformational change in the SLI internal loop that activates the substrate for catalysis [18, 19, 21]. Similarly to the hairpin ribozyme, the VS ribozyme active site is formed by the interaction of two internal loops: the SLI internal loop and the A730 loop of stem-loop VI (SLVI; [4, 6, 22, 23]). The internal loops of SLI and SLVI respectively contain the key catalytic residues G₆₃₈ and A₇₅₆ that contribute to the proposed general acid-base mechanism [16, 22, 24-29].

Although no complete high-resolution structure has been reported for the VS ribozyme, low-resolution models have been derived from biochemical, fluorescence resonance energy transfer (FRET) and small-angle X-ray scattering (SAXS) studies that provide insights into its global structure [13, 30-32]. Moreover, high-resolution NMR structures of isolated domains have been determined for almost all of the structural and functional domains of the VS ribozyme, including the SLI substrate in its active and inactive conformations [33-35], the SLV terminal loop in the presence and absence of magnesium ions (Mg²⁺; [36, 37]), the I/V kissing-loop interaction [21], the A730 loop of SLVI [23, 38, 39] and more recently, the III-IV-V three-way junction [40]. Our laboratory has significantly contributed to this

divide-and-conquer approach, with the ultimate goal of determining a complete high-resolution solution structure of the *Neurospora* VS ribozyme. At this point, the only significant domain that remains to be structurally characterized is the II-III-VI three-way junction.

The II-III-VI junction is a key architectural domain (Figure 5.1A) that is essential for the activity of the VS ribozyme; it orients the A730 loop of SLVI in such a way that it can form the active site with SLI [14, 30, 31]. Substitutions and chemical modifications of several residues within the junction decrease the catalytic activity of the VS ribozyme [10, 30] and several of these residues are protected from chemical modifications under native conditions [9, 12-15], which indicates that this junction adopts a compact structure. Interestingly, substitution of the II-III-VI junction by a ribosomal junction of similar sequence lead to a 10-fold reduction in cleavage rate, suggesting that these junctions may share structural characteristics [30]. Moreover, bioinformatics [41, 42] and biophysical [30] studies of the II-III-VI junction suggest that it adopts a structure where stem III and VI are coaxially stacked, fitting a topological family A structure [30, 41, 42]. In contrast, another bioinformatics study based on a data mining approach proposed coaxial stacking of stems II and III with a family C topology [43]. Thus, there is a lack of consensus at this time on the basic topology of the II-III-VI junction within the VS ribozyme. In addition, although the adenine bulges of the adjoining stems have been shown to contribute to the activity and likely the global structure of the VS ribozyme, it is not clear if they play a role in defining the II-III-VI junction [14, 25, 31, 44]. Similarly, although divalent metal ions are known to stabilize the II-III-VI junction, their precise role remains to be determined [9, 12-15, 30].

In this study, we investigated the NMR solution structure of J236 (Figure 5.1B), a 62-nucleotide (nt) RNA that encompasses the VS ribozyme II-III-VI junction and its three adjoining stems, each one containing its natural adenine bulge. Furthermore, we localized Mg^{2+} -binding sites within the J236 RNA using manganese (Mn^{2+})-induced paramagnetic relaxation enhancement (PRE; [38]). The NMR structure of J236 reveals that the II-III-VI junction belongs to the family C of three-way junctions, with a complex network of interactions at the junction and helical stacking of stems II and III. We also identified a long-range interaction between the adenine bulges of stems II and VI that stabilizes the structure of the II-III-VI junction. In combination with the wealth of information available from previous NMR studies of isolated domains, the NMR structure of J236 provides

significant new insights into the global three-dimensional structure of the *Neurospora* VS ribozyme.

EXPERIMENTAL PROCEDURES

Plasmids. Double-stranded PCR fragments coding for the J236-VS, J236-A₆-bp-VS and J236-ΔA₆-VS RNAs and flanked by a T7 promoter, were inserted into the HindIII/EcoRI sites of the pTZ19R-derived pTR-4 vector [45] to generate the pJ236, pJ236-A₆-bp and pJ236-ΔA₆ plasmids. These plasmids were fully linearized using EcoRI (New England Biolabs) and used for transcription of the J236, J236-A₆-bp and J236-ΔA₆ RNAs (Figure 5.1B and Figure S5.1 of Supporting Information) with a Varkud Satellite (VS) ribozyme substrate at their 3'-end [45].

RNA synthesis and purification. Unlabeled, ¹⁵N-labeled and ¹³C/¹⁵N-labeled RNAs (J236-VS, J236-A₆-bp-VS and J236-ΔA₆-VS) were synthesized *in vitro* with the T7 RNA polymerase, as previously described [46]. Nucleotide-specific ¹³C/¹⁵N-labeled J236-VS RNAs were also synthesized using purified ¹³C/¹⁵N ATP, ¹³C/¹⁵N CTP or ¹³C/¹⁵N GTP for preparation of A^{CN}-J236, C^{CN}-J236 and G^{CN}-J236 RNAs [47]. Following RNA synthesis, the VS ribozyme substrate was cleaved using a *trans*-acting VS ribozyme to produce a homogeneous 3'-end [40, 45, 48]. The RNAs were then purified by denaturing gel electrophoresis, treated with calf intestinal alkaline phosphatase (CIP; Roche Diagnostics) to remove their 5'-phosphates and further purified by DEAE-Sepharose chromatography [38, 49]. The purified RNAs were concentrated and exchanged in NMR buffer A [10 mM sodium cacodylate (pH 6.5), 50 mM KCl, 0.05 mM NaN₃ and 90%:10% H₂O:D₂O] with Amicon Ultra-4 centrifugation filter devices (Millipore). The RNAs were then heated at 37°C for 2 min and cooled in ice water for 5 min before changing to the final NMR buffer using the same filter device [NMR buffer A with 5 mM MgCl₂ 99.995% (Sigma-Aldrich)]. For NMR studies in D₂O, the samples were obtained by four cycles of lyophilization and resuspension in 99.996% D₂O.

NMR spectroscopy. All NMR experiments were collected on a Varian ^{Unity}INOVA 600 MHz spectrometer and details of data collection, assignment and structural restraints are

provided as Supporting Information. Briefly, resonance assignment of the J236 RNA was achieved as previously described [40]. In addition, an A-specific H(NC)-TOCSY-(C)H spectrum [50] was collected for unambiguous assignment of the adenine protons in the junction. Experimental restraints for structure calculation of the J236 RNA were derived from an HNN-COSY spectrum and several NOESY-type spectra [40]. Of note, 3D CT-HCCH-COSY [51], 3D HCCH-TOCSY [51] and ^{13}C -edited HMQC-NOESY [52] spectra collected on nucleotide-specific $^{13}\text{C}/^{15}\text{N}$ -labeled J236 RNAs (A^{CN} -J236, C^{CN} -J236 and G^{CN} -J236 RNAs) significantly contributed to the unambiguous assignment of NMR signals from fully $^{13}\text{C}/^{15}\text{N}$ -labeled J236. NOE-derived distance restraints were separated in four classes based on NOE crosspeak intensities, and dihedral angle restraints for the sugar puckers (δ) and other backbone dihedral angles (α , γ , χ and ζ) were defined based on comparative NOE analyses [53]. Canonical distance restraints and backbone torsion angles were used to define helical regions in agreement with the NMR data.

Metal-ion binding studies. Manganese (Mn^{2+}) titrations were performed with two J236 samples (1.5-mM ^{15}N -labeled J236 and 2.0-mM $^{13}\text{C}/^{15}\text{N}$ -labeled J236) in NMR buffer A containing 5 mM MgCl_2 , as previously described [38, 40]. RNA-metal distance restraints were derived from Mn^{2+} -induced paramagnetic relaxation enhancement (PRE) using the ratio of signal intensity (I_0/I_{Mn}) determined from spectra collected at 0 μM MnCl_2 (I_0) and 20 μM MnCl_2 (I_{Mn} ; [38, 40]).

Structure calculation. Three-dimensional structures of J236 were calculated with restrained molecular dynamics and simulated annealing in X-PLOR-NIH version 2.1.9 [54] from structures with randomized backbone angles, as previously described [37, 40]. A force field was used that included bond, angle, improper and repulsive van der Waals energy terms as well as NOE and torsion-angle pseudoenergy terms, but no electrostatic terms. Three-dimensional structures of J236 bound to eight magnesium complexes ($\text{Mg}(\text{H}_2\text{O})_6^{2+}$; [55]), termed J236^{Mg}, were calculated as described for J236, but using additional metal-RNA restraints (Table S5.3 of Supporting Information; [38, 40]). For both J236 and J236^{Mg}, the 20 lowest-energy structures that satisfied all the experimental restraints (all distance violations <0.2 Å and all torsion angle violations $<5^\circ$) were selected for analysis and used to calculate average structures that were minimized against all experimental restraints.

Structures were visualized with PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC and analyzed with PyMOL and Curves+ [56]. Reported values of rmsd, interatomic distances and interhelical angles are given as average values with standard deviations for the 20 lowest-energy structures. To calculate the interhelical angles, the helical segments were defined as residues 7-10 and 53-56 for stem II, residues 12-14 and 26-28 for stem III and residues 29-31 and 46-48 for stem VI.

RESULTS

The J236 RNA adopts a stable structure in the presence of Mg^{2+} ions. The J236 RNA is a 62-nucleotide (nt) RNA that incorporates all the single-stranded residues of the II-III-VI junction from the *Neurospora* VS ribozyme and parts of stems II, III and VI, including their A-rich bulges that were previously found to be important for the VS ribozyme cleavage activity (Figure 5.1; [14, 31, 44]). Two terminal loops were also inserted into J236 to stabilize the RNA and facilitate the NMR assignment; a CUUG tetraloop closed by a G-C base pair (gCUUGc) caps stem III and a UUCG tetraloop closed by a C-G base pair (cUUCGg) caps stem VI (Figure 5.1B).

We first investigated the effect of Mg^{2+} ions on the structure of the J236 RNA by 1H - ^{15}N heteronuclear NMR methods. In the presence of Mg^{2+} ions (5 mM $MgCl_2$), the 2D 1H - ^{15}N HSQC spectrum (Figure 5.1C) is well dispersed with detectable signals for 25 out of the 28 imino groups, in agreement with formation of a unique stable structure for J236. By comparison, the 2D 1H - ^{15}N HSQC spectrum of J236 collected in the absence of Mg^{2+} ions is significantly different and contains fewer high-intensity signals (Figure S5.1C of Supporting Information). These results indicate that formation of a stable II-III-VI junction is dependent on Mg^{2+} ions, in agreement with previous biochemical and FRET studies [8-15, 30, 44]. Interestingly, an imino signal with an unusual ^{15}N chemical shift is observed for G_{53} only in the presence of Mg^{2+} ions, and a 2D HNN-COSY spectrum reveals that this residue forms a WC/WC G•A base pair within the junction (Figure S5.2A of Supporting Information). Furthermore, a network of NOEs is observed between G_{53} and both G_9 of stem II and G_{28} of stem III, which is indicative of continuous stacking between stems II and III (Figure S5.2B of Supporting Information) and provides evidence for a highly organized junction in the J236 RNA.

Three-way junctions are most often stabilized by remote interactions that involve residues other than those that directly link the helical domains [57]. To test for the possibility of a remote interaction in the II-III-VI junction, we examined the role of the stem II bulge in stabilizing the J236 RNA (Figure S5.1D-G of Supporting Information). Interestingly, the deletion of this bulge (J236- ΔA_6 RNA) or the insertion of a U residue on the opposite strand to form an A-U base pair (J236- A_6 -bp RNA) destabilizes the J236 structure, since 2D ^1H - ^{15}N HSQC spectra of these variants in the presence of Mg^{2+} ions are similar to that of J236 in the absence of Mg^{2+} ions (Figure S5.1 of Supporting Information). Evidently, the A_6 bulge of stem II plays an important role in stabilizing the structure of the II-III-VI junction, and we postulated that it could mediate a long-range interaction with another bulge. Thus, we did not attempt to further reduce the size of the J236 RNA, even though it represented a challenge for NMR structure determination. Prior to depositing the J236 structures, only five NMR structures of RNAs larger than 20 kDa (or > 60 nt) were available in the PDB [58-62].

The overall NMR structure of the J236 RNA. The high-resolution solution structure of the J236 RNA was determined in the presence of Mg^{2+} ions using heteronuclear NMR spectroscopy of uniformly-labeled (^{15}N or $^{13}\text{C}/^{15}\text{N}$) and selectively-labeled (A^{CN} , C^{CN} and G^{CN}) J236 RNAs. This labeling strategy allowed for almost complete assignments of the observable resonances (^1H , ^{15}N and ^{13}C) of the bases and the C1'-H1', C2'-H2', C3'-H3' and C4'-H4' resonances of the ribose moieties (Tables S5.1 and S5.2 of Supporting Information).

Three-dimensional structures of J236 were first determined using NOE-derived distance restraints and dihedral angle restraints. Subsequently, three-dimensional structures of J236 were calculated with bound Mg^{2+} ions (J236 $^{\text{Mg}}$) by adding metal-RNA restraints derived from Mn^{2+} -induced PRE to the existing set of experimental restraints (Table S5.3 of Supporting Information; [38, 40]). Due to these additional restraints, the J236 $^{\text{Mg}}$ structure (Figure 5.2) is better defined than the original J236 structure (Figure S5.3 of Supporting Information) with heavy atom rmsd values that are smaller for the overall structure ($2.95 \pm 0.73 \text{ \AA}$ versus $3.97 \pm 1.51 \text{ \AA}$) as well as for local structural elements (Table 5.1). Thus, the J236 $^{\text{Mg}}$ structure is presented herein (Figures 5.2-5.5), although similar observations are made from the J236 structure (Figures S5.3-S5.6 of Supporting Information).

The J236^{Mg} RNA adopts a Y-shaped fold that belongs to the family C of three-way junctions [41], with a well-defined core domain (rmsd of 0.67 ± 0.10 Å; Table 5.1 and Figure 5.2) that orients stems II and VI side-by-side and away from stem III. Stems II and VI define an acute interhelical angle ($\phi_{II-VI} = 77.5^\circ \pm 10.0^\circ$), whereas the other two stem pairs define obtuse interhelical angles ($\phi_{II-III} = 127.0^\circ \pm 7.9^\circ$ and $\phi_{III-VI} = 149.3^\circ \pm 9.6^\circ$). The orientation of the stems II and VI allows for the formation of a long-range interaction between the A₆ bulge of stem II and the A₃₂A₃₃ bulge of stem VI.

The structure also incorporates a total of eight Mg²⁺-binding sites that were identified from Mn²⁺-induced PRE and modeled with hexahydrated Mg²⁺ complexes [Mg(H₂O)₆²⁺] (Figure 5.2 and Figure S5.7 and Table S5.4 of Supporting Information). As detailed below, several of these sites are associated with structural elements that stabilize the II-III-VI junction. It is important to note that the PRE studies do not provide information about the occupancy of the Mg²⁺ ions at these sites, but simply reveal preferential sites for Mg²⁺ binding [40]. Accordingly, several residues within the J236 core and bulge-bulge interaction display ¹H-¹³C and ¹H-¹⁵N correlation signals of either relatively low intensity (C₅, A₆, A₁₀, C₁₂, U₁₃, C₃₀, U₃₁, C₄₅, A₅₁, A₅₂ and G₅₃) or that correspond to two populations in slow exchange (G₇), suggestive of conformational exchange associated with Mg²⁺-ion binding.

The core of the J236 RNA adopts a compact structure. The core of J236^{Mg} adopts a well-defined structure, in which stems II and III appear to form a continuous helical segment that excludes A₁₁ (Figure 5.2). In contrast, the stacking is disrupted between stems III and VI, with G₂₈ and G₂₉ being stacked on their respective stem, but in a splayed conformation relative to each other and defining an abrupt turn in the phosphate backbone (Figures 5.2 and 5.3). The longest single-stranded region in the junction connects stems II and VI and contains a backbone turn involving residues A₄₉, C₅₀, A₅₁ and A₅₂ (Figure 5.3). Within this ACAA turn, there is continuous stacking between C₄₈ from stem VI, A₄₉ and C₅₀, whereas the backbone is reversed after C₅₀ to allow continuous stacking between A₅₁, A₅₂, G₅₃ and stem II. Although reminiscent of the U-turn motif, the ACAA turn does not display as sharp of a backbone reversal as the U-turn and is stabilized by only one hydrogen bond between C₅₀ 2'-OH and A₅₁ N7 that has no equivalent in the U-turn (Figure 5.3; [37]) Nevertheless, the ACAA turn of J236 positions the bases of A₅₁ and A₅₂ to interact with the

minor groove of stem III, and this positioning of bases for tertiary contacts is typical of the U-turn motif [21, 40, 57, 63, 64].

The core of the three-way junction within J236 is stabilized by a *cis* WC/WC G•A base pair and by two minor groove base triples (Figure 5.4 and Figure S5.5 of Supporting Information). The (U₁₃-A₂₇)•A₅₁ base triple (rmsd of 0.51 ± 0.10 Å) is formed by the interaction of the WC edge of A₅₁ with the minor groove of the WC U₁₃-A₂₇ base pair, whereas the (C₁₂-G₂₈)•A₅₂ base triple (rmsd of 0.53 ± 0.17 Å) involves the WC edge of A₅₂ with the minor groove of the WC C₁₂-G₂₈ base pair. The (C₁₂-G₂₈)•A₅₂ base triple stacks with the (U₁₃-A₂₇)•A₅₁ base triple on one side and with the well-defined *cis* WC/WC A₁₀•G₅₃ base pair (rmsd of 0.33 ± 0.15 Å; Figure 5.4C) on the other side to allow continuous stacking between stems II and III.

Two Mg²⁺ ions interact with the core domain (Figure 5.3, Figure S5.7D and G and Table S5.4 of Supporting Information) that are both associated with the A₁₀•G₅₃ base pair. The Mg²⁺ ion at Site 4 is located at the A₄₉C₅₀A₅₁A₅₂ turn, making outer-sphere interactions with both A₄₉ and G₅₃ and stabilized by electrostatic interactions with the 5'-phosphates of A₅₂, G₅₃ and C₅₄. The Mg²⁺ ion at Site 7 is located within the major groove of stem II, making outer-sphere contacts with G₉, A₁₀, G₂₈ and G₅₃ and stabilized by electrostatic interactions with the 5'-phosphates of G₉, A₁₀ and G₂₉. Although the Mg²⁺ ion at Site 7 may play a specific role in stabilizing the sharp backbone turn at the G₂₈-G₂₉ dinucleotide step, both of these Mg²⁺ ions are likely important for the compact structure at the II-III-VI junction.

A long-range interaction between stems II and VI. It was previously demonstrated that the adenine bulges in stems II (A₆) and VI (A₃₂A₃₃) play critical roles in the VS ribozyme, since their deletion, strand inversion or substitution by other standard bases reduces the cleavage activity by up to a thousand fold [14, 25, 44]. The NMR structure of J236^{Mg} reveals that A₆ is extruded from stem II and is inserted into the A₃₂A₃₃ bulge of stem VI to participate in a long-range interaction (Figure 5.5A, left panel). In this bulge-bulge interaction, A₆ shares two hydrogen bonds with A₃₂ to form a *cis* Sugar edge/WC A₆•A₃₂ base pair (rmsd of 0.80 ± 0.24 Å; Figure 5.5B), in agreement with NOE data (Figure S5.8 of Supporting Information). In addition, the base of A₆ is sandwiched between the bases of residues C₄₅ and A₄₆ in stem VI (Figure 5.5A, right panel). This tertiary interaction may be facilitated by the S-turn motif centered at the A₃₂A₃₃ bulge (Figure 5A, right panel), which

involves the backbone reversal and the 2'-endo conformation of the A₃₂ ribose and the exclusion of A₃₃ from stem VI [23, 65].

Three Mg²⁺-binding sites were identified near the bulge-bulge interaction. The Mg²⁺ ions at sites 2 and 3 interact with residues of both stems II and VI, whereas the Mg²⁺ ion at Site 8 interacts with stem II near the A₆ bulge (Figure 5.5A, Figure S5.7 and Table S5.4 of Supporting Information). It is likely that these three Mg²⁺ ions facilitate the formation of the II-VI bulge-bulge interaction by counter-acting the negative charges carried by the backbones of stems II and VI. Furthermore, it is likely that Mg²⁺ binding at these sites is largely responsible for the Mg²⁺-dependent folding of the II-III-VI junction.

DISCUSSION

The NMR studies of the J236 RNA reveal that the II-III-VI junction of the VS ribozyme forms a compact network of interactions at the core of the junction as well as a tertiary interaction between the adenine bulges of stems II and VI. Moreover, these studies demonstrate that the structure of the II-III-VI junction depends on both the formation of this long-range tertiary interaction and the presence of Mg²⁺ ions. These new findings are discussed below in light of previous biochemical and biophysical studies of the VS ribozyme. The II-III-VI junction being the last piece of the puzzle in our quest to characterize isolated domains of the *Neurospora* VS ribozyme by NMR spectroscopy, its structure allows us to gain further insights into the overall three-dimensional structure of the complete ribozyme.

The NMR structure of the II-III-VI junction is consistent with biochemical data. The core of the J236 RNA (Figure 5.1B) relies on an extended network of interactions that is compatible with previous biochemical data (Figures 5.3 and 5.4). First, the importance of specific residues was previously revealed by kinetic studies of VS ribozyme variants [9, 25, 30]. In agreement with our observation that the *cis* WC-WC A₁₀•G₅₃ base pair is important for the stability of the II-III-VI junction, single base substitutions of residues A₁₀ and G₅₃, but not the A₁₀•G₅₃ base-pair inversion, greatly decrease (up to 500 fold) the catalytic activity of the VS ribozyme [25, 30]. Similarly, a 20-fold reduction in the cleavage activity was observed for the A to U substitution of both A₅₁ and A₅₂ within the ACAA turn,

consistently with their participation in base triples (Figure 5.4; [30]). In contrast, when A₄₉ and C₅₀ of the ACAA turn were substituted by a uridine, no significant decrease in activity was observed, which is consistent with the lack of interaction of their bases in the structure [25, 30]. Likewise, the individual inversions of the proximal base pairs of the II-III-VI junction [G₉-C₅₄ in stem II, C₁₂-G₂₈ and U₁₃-A₂₇ in stem III and G₂₉-C₄₈ in stem VI] did not significantly alter the catalytic activity [9, 30], which is consistent with the lack of tertiary interaction involving the closing base pairs of stems II and VI. Moreover, it suggests that the inversions of the C₁₂-G₂₈ and U₁₃-A₂₇ base pairs in stem III also support the formation of base triples equivalent to the (U₁₃-A₂₇)•A₅₁ and (C₁₂-G₂₈)•A₅₂ base triples within the II-III-VI junction. Using WebFR3D, we found several examples of such inverted minor groove (G-C)•A and (A-U)•A base triples with the same topology as those observed in J236 (results not shown; [66]).

Chemical probing and interference data obtained under native conditions can also be reexamined in light of the J236 structure. Briefly, the WC edges of all core residues are protected from chemical probing in the presence of Mg²⁺ ions, with the exception of A₁₁ and C₅₀, in agreement with the compact structure of the J236 core [9, 14, 15]. Moreover, chemical modification data are consistent with impairment of critical interactions observed in the NMR structure; this includes carboxyethylation at the N7 position of A₂₇, G₂₈, A₄₉, A₅₁, A₅₂ and G₅₃ by diethyl pyrocarbonate (DEPC) and removal of the C₁₂ and U₁₃ bases by hydrazine [10]. Also, substitution of G₂₈ by an N²-methylguanosine decreases the ligation rate, supporting its role in the formation of the (C₁₂-G₂₈)•A₅₂ base triple (Figure 5.4; [24]). Moreover, the 5'-phosphates of A₁₀, A₁₁ and G₂₉ are protected from ethylnitrosourea (ENU) modification [12] and the riboses of A₁₀, A₂₇, G₂₈ and G₂₉ are protected from hydroxyl radical footprinting [13], in agreement with backbone distortion near these residues, reduced accessibility of ribose C4' within base triples (for A₂₇ and G₂₈; Figure 5.4) and binding of a Mg²⁺ ion at site 7 (for A₁₀ and G₂₉; Figure 5.3 and Figure S5.7G of Supporting Information). Furthermore, the importance of the 2'-OH of U₁₃ revealed by site-specific 2'-deoxy modification [67] is consistent with its role in the formation of the (U₁₃-A₂₇)•A₅₁ base triple (Figure 5.4A). Thus, the NMR structure of the J236 core is in general agreement with mutagenesis, chemical probing and interference data of the VS ribozyme.

Exceptionally, A₁₁ adopts a bulged out position in the NMR structure that appears incompatible with biochemical data (Figure 5.3). Base substitution of this residue by a U

was previously shown to lead to a 15-fold reduction in the cleavage activity [30], similarly to the effects of carboxyethylation at the N7 position [10] and base substitution by a purine [24]. Taken together, these results suggest that the A₁₁ modifications prevent the formation of a tertiary interaction that is important for the cleavage activity of the VS ribozyme but not present in J236.

The II-VI bulge-bulge interaction is supported by biochemical data. The NMR structure of J236 reveals a bulge-bulge interaction between stems II and VI that involves a *cis* Sugar edge/WC A₆•A₃₂ base pair (Figure 5.5 and Figure S5.6 of Supporting Information). Such long-range tertiary interaction between two helical domains of a three-way junction is extremely common in RNA and is shared by almost all junctions from family C topology [41, 57]. For example, long-range tertiary interactions have been shown to stabilize the hammerhead ribozyme [68], purine riboswitches [69, 70] and the viral Xrn-1 resistant sfRNA [71].

Although this bulge-bulge interaction was not previously identified in the VS ribozyme, its structure is in agreement with available biochemical data. Notably, the deletion of A₆ and the insertion of one base pair between the A₆ bulge and the junction reduce the cleavage activity by 77 fold [14, 25, 44] and 180 fold [44], respectively, which is consistent with the requirement of this bulge for proper folding of the II-III-VI junction (Figure S5.1 of Supporting Information). Substitutions of A₆ by a U, a C or a G reduce the cleavage activity by 4, 6 and 100 fold, respectively, in line with the chemical compatibility of a pyrimidine (carbonyl at C2) but not a guanine (NH₂ at C2) for hydrogen bonding with the NH₂ group of A₃₂ (Figure 5.5; [44]). Although most single base substitutions of A₃₂ and A₃₃ do not significantly affect the catalytic activity, the deletion of the A₃₂A₃₃ bulge reduces the cleavage activity of the VS ribozyme by 140 fold [14, 44]. These results are in agreement with the bulged out position of A₃₃ and a large subset of isosteric base pairs for the *cis* Sugar edge/WC A₆•A₃₂ base pair (Figure 5.5; [72]).

Previous chemical probing experiments performed under native conditions demonstrated that the WC and Hoogsteen edges of A₆ and A₃₂ but not A₃₃ are protected from chemical modifications, consistent with the *cis* Sugar edge/WC A₆•A₃₂ base pair being nested between the minor grooves of stems II and VI and the exclusion of A₃₃ within the S-turn (Figure 5.5; [9, 14, 15]). Similarly, carboxyethylation at the N7 position of A₆ interferes

with the cleavage activity [10]. The 5'-phosphates of G₇ and G₃₄ are protected from ENU modification and phosphorothioate substitution of G₇ interferes with the cleavage activity, in agreement with the 5'-phosphates of G₇ and G₃₄ contributing to electrostatic stabilization of the Mg²⁺ ion at site 2 (Figure 5.5, Figure S5.7 and Table S5.4 of Supporting Information; [12]). Similarly, residues A₆ and A₃₂ as well as those of the C₅-G₅₇ and G₇-C₅₆ base pairs are protected from hydroxyl radical footprinting [13], consistent with a reduced accessibility of their C4' within the II-VI bulge-bulge interaction. Furthermore, a UV-inducible crosslink was reported between A₆ and A₄₆ [14], and, accordingly, these residues are stacked on each other in the NMR structure of J236. Hence, the II-VI bulge-bulge interaction revealed by the NMR structure of J236 correlates very well with biochemical data.

The structure of the II-III-VI junction depends on the II-VI bulge-bulge interaction and the binding of Mg²⁺ ions. The II-III-VI junction adopts a Y-shaped fold with stems II and VI oriented side-by-side and away from stem III (Figure 5.2), which is typical of a family C topology [41]. Although two different helical stacking schemes have been predicted for this three-way junction, involving either stems II and III [43] or stems III and VI [41, 42], the NMR structure of J236 confirms continuous helical stacking of stems II and III. Such helical stacking was not observed in a three-way junction with a similar sequence derived from the 23S rRNA of the 50S ribosomal subunit of *Haloarcula marismortui* and used to model the II-III-VI junction within the SAXS model of the VS ribozyme [30, 32]. This may explain why a hybrid VS ribozyme in which the core of the junction was replaced by this rRNA three-way junction has a 10-fold lower cleavage rate [30]. Similarly, continuous stacking between stems II and III is incompatible with comparative gel electrophoresis and FRET studies of the II-III-VI junction [30], which provided evidence for near-coaxial alignment of stems III and VI. However, in these studies [30], a minimal II-III-VI junction was used that did not incorporate the adenine bulges in stems II and VI, and thus did not allow for formation of the II-VI bulge-bulge interaction revealed here. Here, NMR studies of J236 variants clearly demonstrate that the II-VI bulge-bulge interaction is required for stable folding of the II-III-VI junction. Thus, it is not surprising that the FRET-derived global structure of the II-III-VI junction [30], although it displays an acute angle between stems II and VI, is not otherwise consistent with the NMR structure determined here.

The overall fold of the II-III-VI junction not only depends on the II-VI bulge-bulge interaction, but also on the presence of Mg^{2+} ions. In fact, the NMR spectra of J236 variants that do not form the II-VI bulge-bulge interactions are very similar to those of J236 collected in the absence of Mg^{2+} ions; in all cases these data reveal a destabilized junction core. The stabilizing effect of Mg^{2+} ions is further supported by the localization of two Mg^{2+} ions in the J236 core and three Mg^{2+} ions within the II-VI bulge-bulge interaction. In particular, we noted that the Mg^{2+} ions at sites 2 and 3 within the bulge-bulge interaction and site 7 at the junction core help bridge phosphates that are remote in the primary structure but come in proximity in the tertiary structure. In agreement with a Mg^{2+} -dependent structure for the II-III-VI junction, the Hoogsteen and WC edges of several purines, including A_6 and A_{32} of the bulges, are more protected from chemical modification in the presence of Mg^{2+} ions [9, 10]. Furthermore, the critical role of Mg^{2+} ions in stabilizing the II-III-VI junction is supported by a single-molecule FRET study of the VS ribozyme that revealed important conformational changes induced by Mg^{2+} ions [73].

Novel insights into the global structure of the VS ribozyme. Over the years, NMR structures of several isolated subdomains of the *Neurospora* VS ribozyme have been determined, including that of the SLI substrate in its inactive [33, 34] and active forms [35], the terminal loop of SLV [36, 37], the I/V kissing-loop interaction [21], the A730 internal loop of SLVI [23, 38, 39], and the III-IV-V junction [40]. With the NMR structure of the II-III-VI junction presented here, all non-helical domains of the minimal VS ribozyme have now been structurally characterized. These available three-dimensional NMR structures of isolated domains provide insights into the global organization of the complete VS ribozyme (Figure 5.6). Similarly to the II-III-VI three-way junction that adopts a family C topology with helical stacking of stems II and III and an ACAA turn connecting stems II and VI, the III-IV-V three-way junction adopts a family C topology with helical stacking of stems III and IV and a U-turn connecting stems III and V [40]. Thus, we propose that these two junctions allow stacking of stems II, III and IV to create an extended, continuous helical segment from which stems V and VI project alongside stems III and II, respectively. Extensive base stacking at the I/V kissing junction also creates a more or less continuous helical region encompassing SLV and SLI [21], and bringing the internal loop of SLI in proximity of the A730 internal loop of SLVI to create the active site [23, 74].

Complementary structural studies are currently underway to incorporate these NMR structures of isolated domains within a global three-dimensional solution structure of the *Neurospora* VS ribozyme.

REFERENCES

1. Saville, B.J. and R.A. Collins, *A site-specific self-cleavage reaction performed by a novel RNA in Neurospora mitochondria*. *Cell*, 1990. **61**: p. 685-696.
2. Collins, R.A., *The Neurospora Varkud satellite ribozyme*. *Bioch. Soc. Trans. Rev.*, 2002. **30**: p. 1122-1126.
3. Lilley, D.M., *The Varkud satellite ribozyme*. *RNA*, 2004. **10**: p. 151-158.
4. Lilley, D.M.J., *The hairpin and Varkud satellite ribozymes*, in *Ribozymes and RNA catalysis*, D.M.J. Lilley and F. Eckstein, Editors. 2008, Royal Society of Chemistry: Cambridge. p. 66-91.
5. Cochrane, J.C. and S.A. Strobel, *Catalytic strategies of self-cleaving ribozymes*. *Acc. Chem. Res.*, 2008. **41**(8): p. 1027-35.
6. Wilson, T.J. and D.M. Lilley, *Do the hairpin and VS ribozymes share a common catalytic mechanism based on general acid-base catalysis? A critical assessment of available experimental data*. *RNA*, 2011. **17**(2): p. 213-21.
7. Saville, B.L. and R.A. Collins, *RNA-mediated ligation of self-cleavage products of a Neurospora mitochondrial plasmid transcript*. *Proc. Natl. Acad. Sci. U. S. A.*, 1991. **88**: p. 8826-8830.
8. Collins, R.A. and J.E. Olive, *Reaction conditions and kinetics of self-cleavage of a ribozyme derived from Neurospora VS RNA*. *Biochemistry*, 1993. **32**: p. 2795-2799.
9. Beattie, T.L., J.E. Olive, and R.A. Collins, *A secondary-structure model for the self-cleaving region of Neurospora VS RNA*. *Proc. Natl. Acad. Sci. U. S. A.*, 1995. **92**: p. 4686-4690.
10. Beattie, T.L. and R.A. Collins, *Identification of functional domains in the self-cleaving Neurospora VS ribozyme using damage selection*. *J. Mol. Biol.*, 1997. **267**: p. 830-840.
11. Murray, J.B., et al., *The hammerhead, hairpin and VS ribozymes are catalytically proficient in monovalent cations alone*. *Chem. Biol.*, 1998. **5**: p. 587-595.

12. Sood, V.D., T.L. Beattie, and R.A. Collins, *Identification of phosphate groups involved in metal binding and tertiary interactions in the core of the Neurospora VS ribozyme*. J. Mol. Biol., 1998. **282**: p. 741-750.
13. Hiley, S.L. and R.A. Collins, *Rapid formation of a solvent-inaccessible core in the Neurospora Varkud satellite ribozyme*. EMBO J., 2001. **20**(19): p. 5461-5469.
14. Sood, V.D. and R.A. Collins, *Identification of the catalytic subdomain of the VS ribozyme and evidence for remarkable sequence tolerance in the active site loop*. J. Mol. Biol., 2002. **320**: p. 443-454.
15. Maguire, J.L. and R.A. Collins, *Effects of cobalt hexammine on folding and self-cleavage of the Neurospora VS ribozyme*. J. Mol. Biol., 2001. **309**: p. 45-56.
16. Smith, M.D., et al., *The ionic environment determines ribozyme cleavage rate by modulation of nucleobase pKa*. RNA, 2008. **14**(9): p. 1942-9.
17. Rastogi, T., et al., *A long-range pseudoknot is required for activity of the Neurospora VS ribozyme*. EMBO J., 1996. **15**(11): p. 2820-2825.
18. Andersen, A. and R.A. Collins, *Rearrangement of a stable RNA secondary structure during VS ribozyme catalysis*. Mol. Cell, 2000. **5**: p. 469-478.
19. Andersen, A.A. and R.A. Collins, *Intramolecular secondary structure rearrangement by the kissing interaction of the Neurospora VS ribozyme*. Proc. Natl. Acad. Sci U. S. A., 2001. **98**(14): p. 7730-7735.
20. Bouchard, P. and P. Legault, *A remarkably stable kissing-loop interaction defines substrate recognition by the Neurospora VS Ribozyme*. RNA, 2014. **20**(9): p. 1451-64.
21. Bouchard, P. and P. Legault, *Structural insights into substrate recognition by the Neurospora Varkud satellite ribozyme: importance of U-turns at the kissing-loop junction*. Biochemistry, 2014. **53**(1): p. 258-69.
22. Wilson, T.J., A.C. McLeod, and D.M. Lilley, *A guanine nucleobase important for catalysis by the VS ribozyme*. EMBO J., 2007. **26**(10): p. 2489-500.
23. Desjardins, G., et al., *NMR structure of the A730 loop of the Neurospora VS ribozyme: insights into the formation of the active site*. Nucleic Acids Res., 2011. **39**(10): p. 4427-37.

24. Jones, F.D. and S.A. Strobel, *Ionization of a critical adenosine residue in the Neurospora Varkud satellite ribozyme active site*. *Biochemistry*, 2003. **42**: p. 4265-4276.
25. McLeod, A.C. and D.M. Lilley, *Efficient, pH-dependent RNA ligation by the VS ribozyme in trans*. *Biochemistry*, 2004. **43**(4): p. 1118-25.
26. Zhao, Z.Y., et al., *Nucleobase participation in ribozyme catalysis*. *J. Am. Chem. Soc.*, 2005. **127**(14): p. 5026-7.
27. Smith, M.D. and R.A. Collins, *Evidence for proton transfer in the rate-limiting step of a fast-cleaving Varkud satellite ribozyme*. *Proc. Natl. Acad. Sci. U. S. A.*, 2007. **104**(14): p. 5818-23.
28. Jaikaran, D., et al., *An important role of G638 in the cis-cleavage reaction of the Neurospora VS ribozyme revealed by a novel nucleotide analog incorporation method*. *RNA*, 2008. **14**(5): p. 938-49.
29. Wilson, T.J., et al., *Nucleobase-mediated general acid-base catalysis in the Varkud satellite ribozyme*. *Proc. Natl. Acad. Sci. U. S. A.*, 2010. **107**(26): p. 11751-6.
30. Lafontaine, D.A., D.G. Norman, and D.M. Lilley, *Structure, folding and activity of the VS ribozyme: importance of the 2-3-6 helical junction*. *EMBO J.*, 2001. **20**(6): p. 1415-1424.
31. Lafontaine, D.A., D.G. Norman, and D.M. Lilley, *The global structure of the VS ribozyme*. *EMBO J.*, 2002. **21**(10): p. 2461-2471.
32. Lipfert, J., et al., *The complete VS ribozyme in solution studied by small-angle X-ray scattering*. *Structure*, 2008. **16**(9): p. 1357-67.
33. Michiels, P.J.A., et al., *Structure of the ribozyme substrate hairpin of Neurospora VS RNA: A close look at the cleavage site*. *RNA*, 2000. **6**: p. 1821-1832.
34. Flinders, J. and T. Dieckmann, *A pH controlled conformational switch in the cleavage site of the VS ribozyme substrate RNA*. *J. Mol. Biol.*, 2001. **308**: p. 665-679.
35. Hoffmann, B., et al., *NMR structure of the active conformation of the Varkud satellite ribozyme cleavage site*. *Proc. Natl. Acad. Sci. U. S. A.*, 2003. **100**(12): p. 7003-7008.

36. Campbell, D.O. and P. Legault, *NMR structure of the Varkud satellite ribozyme stem-loop V RNA and magnesium-ion binding from chemical-shift mapping*. *Biochemistry*, 2005. **44**: p. 4157-4170.
37. Campbell, D.O., et al., *NMR structure of Varkud satellite ribozyme stem-loop V in the presence of magnesium ions and localization of metal-binding sites*. *Biochemistry*, 2006. **45**(35): p. 10591-605.
38. Bonneau, E. and P. Legault, *NMR localization of divalent cations at the active site of the Neurospora VS ribozyme provides insights into RNA-metal-ion interactions*. *Biochemistry*, 2014. **53**(3): p. 579-90.
39. Flinders, J. and T. Dieckmann, *The solution structure of the VS ribozyme active site loop reveals a dynamic "hot-spot"*. *J. Mol. Biol.*, 2004. **341**: p. 935-949.
40. Bonneau, E. and P. Legault, *Nuclear Magnetic Resonance Structure of the III-IV-V Three-Way Junction from the Varkud Satellite Ribozyme and Identification of Magnesium-Binding Sites Using Paramagnetic Relaxation Enhancement*. *Biochemistry*, 2014. **53**(39): p. 6264-75.
41. Lescoute, A. and E. Westhof, *Topology of three-way junctions in folded RNAs*. *RNA*, 2006. **12**(1): p. 83-93.
42. Tyagi, R. and D.H. Mathews, *Predicting helical coaxial stacking in RNA multibranch loops*. *RNA*, 2007. **13**(7): p. 939-51.
43. Laing, C., et al., *Predicting coaxial helical stacking in RNA junctions*. *Nucleic Acids Res.*, 2012. **40**(2): p. 487-98.
44. Lafontaine, D.A., et al., *The A730 loop is an important component of the active site of the VS ribozyme*. *J. Mol. Biol.*, 2001. **312**: p. 663-674.
45. Rastogi, T. and R.A. Collins, *Smaller, faster ribozymes reveal the catalytic core of Neurospora VS RNA*. *J. Mol. Biol.*, 1998. **277**: p. 215-224.
46. Nikonowicz, E.P., et al., *Preparation of ¹³C and ¹⁵N labelled RNAs for heteronuclear multidimensional NMR studies*. *Nucleic Acids Res.*, 1992. **20**: p. 4507-4513.
47. Dagenais, P. and P. Legault, *Preparative separation of ribonucleoside monophosphates by ion-pair reverse-phase HPLC*. *Methods Mol. Biol.*, 2012. **941**: p. 247-56.

48. Guo, H.C.T. and R.A. Collins, *Efficient trans-cleavage of a stem-loop RNA substrate by a ribozyme derived from Neurospora VS RNA*. EMBO J., 1995. **14**(12): p. 368-376.
49. Delfosse, V., et al., *Riboswitch structure: an internal residue mimicking the purine ligand*. Nucleic Acids Res., 2010. **38**(6): p. 2057-68.
50. Simorre, J.P., et al., *Triple-resonance experiments for assignment of adenine base resonances in ¹³C/¹⁵N-labeled RNA*. J. Am. Chem. Soc., 1996. **118**: p. 5316-5317.
51. Pardi, A. and E.P. Nikonowicz, *Simple procedure for resonance assignment of the sugar protons in ¹³C-labeled RNAs*. J. Am. Chem. Soc., 1992. **114**: p. 9202-9203.
52. Ikura, M., et al., *Three-dimensional NOESY-HMQC spectroscopy of a ¹³C-labeled protein*. J. Magn. Reson., 1990. **86**: p. 204-209.
53. Wijmenga, S.S., M.M.W. Mooren, and C.W. Hilbers, *NMR of nucleic acids; from spectrum to structure*, in *NMR of macromolecules: a practical approach.*, G.C.K. Roberts, Editor. 1993, Oxford University Press: New York. p. 217-288.
54. Schwieters, C.D., et al., *The Xplor-NIH NMR molecular structure determination package*. J. Magn. Reson., 2003. **160**: p. 66-74.
55. Kleywegt, G.T. and T.A. Jones, *Databases in protein crystallography*. Acta Crystallogr. D, 1998. **54**: p. 1119-11131.
56. Lavery, R., et al., *Conformational analysis of nucleic acids revisited: Curves+*. Nucleic Acids Res., 2009. **37**(17): p. 5917-29.
57. de la Pena, M., D. Dufour, and J. Gallego, *Three-way RNA junctions with remote tertiary contacts: a recurrent and highly versatile fold*. RNA, 2009. **15**(11): p. 1949-64.
58. Lukavsky, P.J., et al., *Structure of HCV IRES domain II determined by NMR*. Nat. Struct. Biol., 2003. **10**(12): p. 1033-8.
59. D'Souza, V., et al., *NMR structure of the 101-nucleotide core encapsidation signal of the Moloney murine leukemia virus*. J. Mol. Biol., 2004. **337**(2): p. 427-42.
60. Miyazaki, Y., et al., *Structure of a conserved retroviral RNA packaging element by NMR spectroscopy and cryo-electron tomography*. J. Mol. Biol., 2010. **404**(5): p. 751-72.
61. Burke, J.E., et al., *Structure of the yeast U2/U6 snRNA complex*. RNA, 2012. **18**(4): p. 673-83.

62. Miller, S.B., et al., *A structure-based mechanism for tRNA and retroviral RNA remodelling during primer annealing*. *Nature*, 2014.
63. von Ahsen, U., et al., *Identification of 2'-hydroxyl groups required for interaction of a tRNA anticodon stem-loop region with the ribosome*. *RNA*, 1997. **3**(1): p. 49-56.
64. Lambert, D., J.E. Heckman, and J.M. Burke, *Three conserved guanosines approach the reaction site in native and minimal hammerhead ribozymes*. *Biochemistry*, 2006. **45**(23): p. 7140-7.
65. Correll, C.C., et al., *Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain*. *Cell*, 1997. **91**: p. 705-712.
66. Petrov, A.I., C.L. Zirbel, and N.B. Leontis, *WebFR3D--a server for finding, aligning and analyzing recurrent RNA 3D motifs*. *Nucleic Acids Res.*, 2011. **39**(Web Server issue): p. W50-5.
67. Sood, V.D., S. Yekta, and R.A. Collins, *The contribution of 2'-hydroxyls to the cleavage activity of the Neurospora VS ribozyme*. *Nucleic Acids Res.*, 2002. **30**(5): p. 1132-1138.
68. Martick, M. and W.G. Scott, *Tertiary contacts distant from the active site prime a ribozyme for catalysis*. *Cell*, 2006. **126**(2): p. 309-20.
69. Batey, R.T., S.D. Gilbert, and R.K. Montange, *Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine*. *Nature*, 2004. **432**(7015): p. 411-5.
70. Lemay, J.F., et al., *Folding of the adenine riboswitch*. *Chem. Biol.*, 2006. **13**(8): p. 857-68.
71. Chapman, E.G., et al., *The structural basis of pathogenic subgenomic flavivirus RNA (sfRNA) production*. *Science*, 2014. **344**(6181): p. 307-10.
72. Leontis, N.B., J. Stombaugh, and E. Westhof, *The non-Watson-Crick base pairs and their associated isostericity matrices*. *Nucleic Acids Res.*, 2002. **30**(16): p. 3497-531.
73. Pereira, M.J., et al., *Single VS ribozyme molecules reveal dynamic and hierarchical folding toward catalysis*. *J. Mol. Biol.*, 2008. **382**(2): p. 496-509.
74. Hiley, S.L., et al., *4-thio-U cross-linking identifies the active site of the VS ribozyme*. *EMBO J.*, 2002. **21**(17): p. 4691-4698.

TABLE

Table 5.1. Structural statistics for J236 and J236^{Mg}.

	J236	J236 ^{Mg}
Number of experimental restraints		
NOE-derived distance restraints	2341	2341
Inter-nucleotide	1090	1090
Intra-nucleotide	1244	1244
Ambiguous	7	7
Hydrogen-bond restraints	120	120
Base-pair planarity restraints	42	42
Mg ²⁺ -RNA distance restraints	0	95
Dihedral angle restraints	193	193
Total number of restraints	2696	2791
Rmsd from experimental restraints		
NOE (Å) (none > 0.2)	0.0088 ± 0.0004	0.0117 ± 0.0003
Dihedral (°) (none > 5°)	0.06 ± 0.01	0.07 ± 0.01
Rmsd from idealized geometry		
Bonds (Å)	0.00375 ± 0.00003	0.00382 ± 0.00002
Angles (°)	1.017 ± 0.001	1.004 ± 0.001
Impropers (°)	0.379 ± 0.002	0.386 ± 0.002
Heavy-atom rmsd to the minimized average structure (Å)		
Overall (residues 2-60)	3.97 ± 1.51	2.95 ± 0.73
Stem II (residues 2-9 and 54-60)	1.80 ± 0.72	1.38 ± 0.51
Stem-loop III (residues 13-27)	1.47 ± 0.61	0.97 ± 0.32
Stem-loop VI (residues 30-32, 34-47)	1.50 ± 0.63	1.21 ± 0.35
Core (residues 10-13, 27-29 and 48-53)	0.83 ± 0.14	0.67 ± 0.10
(U ₁₃ -A ₂₇)•A ₅₁ base triple	0.59 ± 0.09	0.51 ± 0.10
(C ₁₂ -G ₂₈)•A ₅₂ base triple	0.69 ± 0.15	0.53 ± 0.17
A ₁₀ •G ₅₃ <i>cis</i> WC/WC base pair	0.58 ± 0.21	0.33 ± 0.15
A ₆ •A ₃₂ <i>cis</i> Sugar edge/WC base pair	1.20 ± 0.25	0.80 ± 0.24

FIGURES

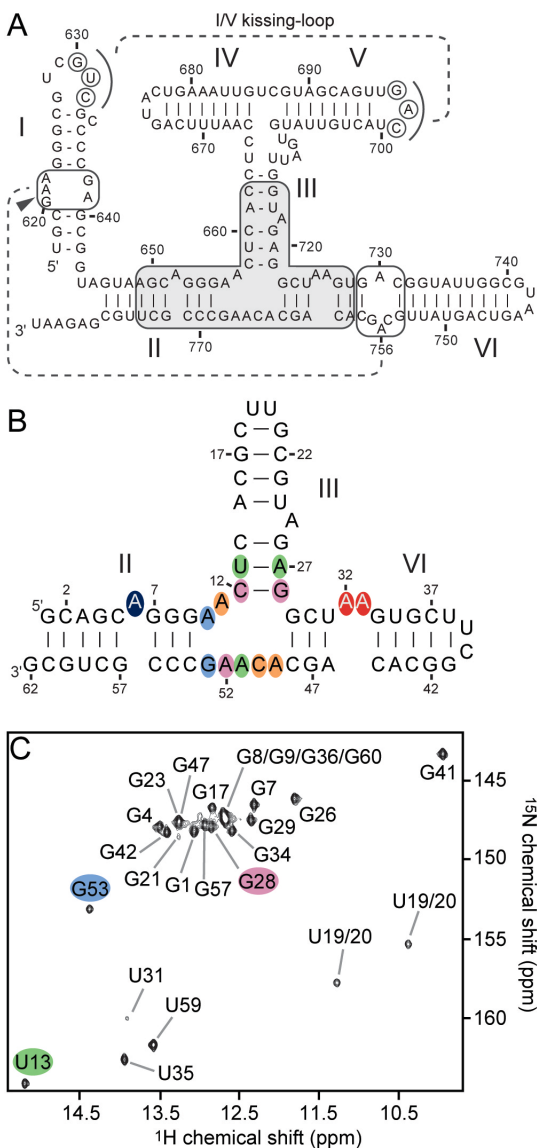


Figure 5.1. The J236 RNA adopts a stable structure in the presence of Mg^{2+} ions. (A) Primary and secondary structures of a *cis*-cleaving VS ribozyme (including residues 617-783; [9]). The substrate is shown in its active conformation [18], with the cleavage site indicated by a gray arrowhead. Dashed lines represent tertiary interactions involving the SLI substrate, either the I/V kissing-loop interaction or the I/VI loop-loop interaction that forms the active site. The nucleotides involved in WC base pairs at the I/V kissing-loop junction are circled. Gray shading is used to highlight the residues from the II-III-VI junction that are present in the J236 RNA. (B) Sequence and proposed secondary structure of the 62-nt J236 RNA used in this study. (C) 2D ^1H - ^{15}N HSQC spectrum of the ^{15}N -labeled J236 RNA in the presence of 5 mM MgCl_2 . In (B) and (C) shading of residues is color-coded according to structural elements present in the NMR structure (see Figure 5.2).

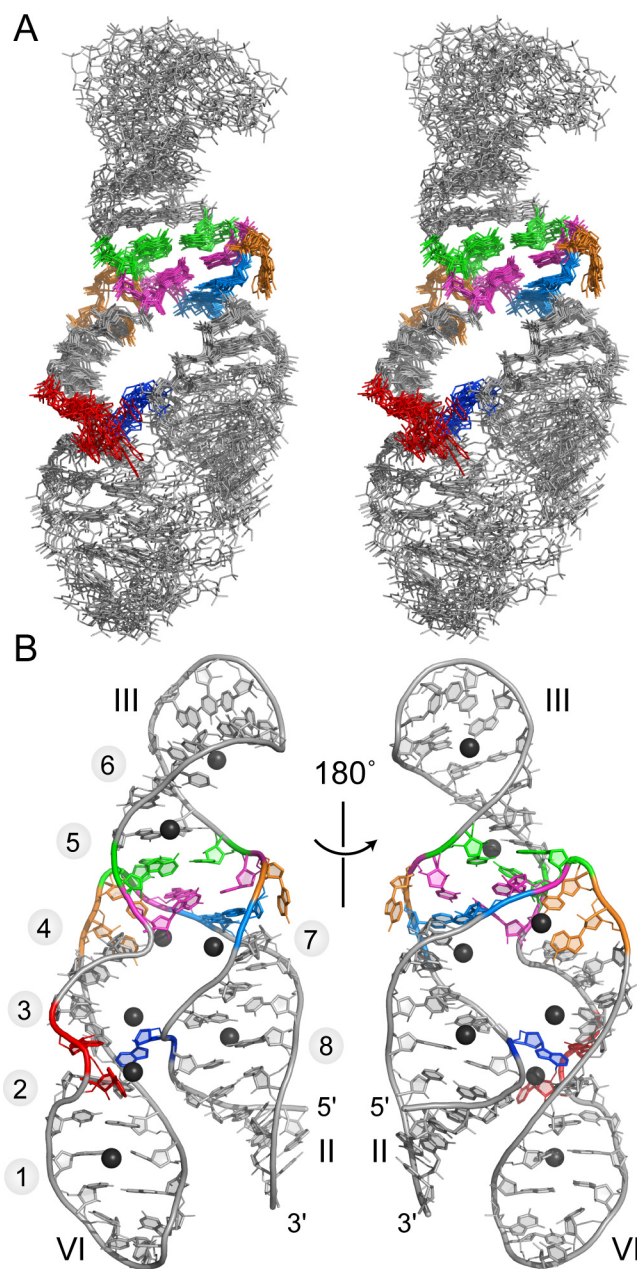


Figure 5.2. NMR solution structure of the J236^{Mg} RNA. (A) Stereoview of the 10 lowest-energy structures of J236^{Mg}. Only heavy atoms of residues 2-60 were used for the superposition and are shown. For clarity, the Mg²⁺ ions were omitted. (B) Stick representations of the lowest-energy structure of the J236^{Mg} RNA (residues 1-61), with the Mg²⁺ ions (dark gray spheres) numbered 1-8. In (A) and (B), residues are color-coded according to structural elements: the (U₁₃-A₂₇)•A₅₁ base triple in green, the (C₁₂-G₂₈)•A₅₂ base triple in magenta, the A₁₀•G₅₃ base pair in light blue, the unpaired residues of the core (A₁₁, A₄₉ and C₅₀) in orange, the A₆ bulge of stem II in dark blue and the A₃₂A₃₃ bulge of stem VI in red.

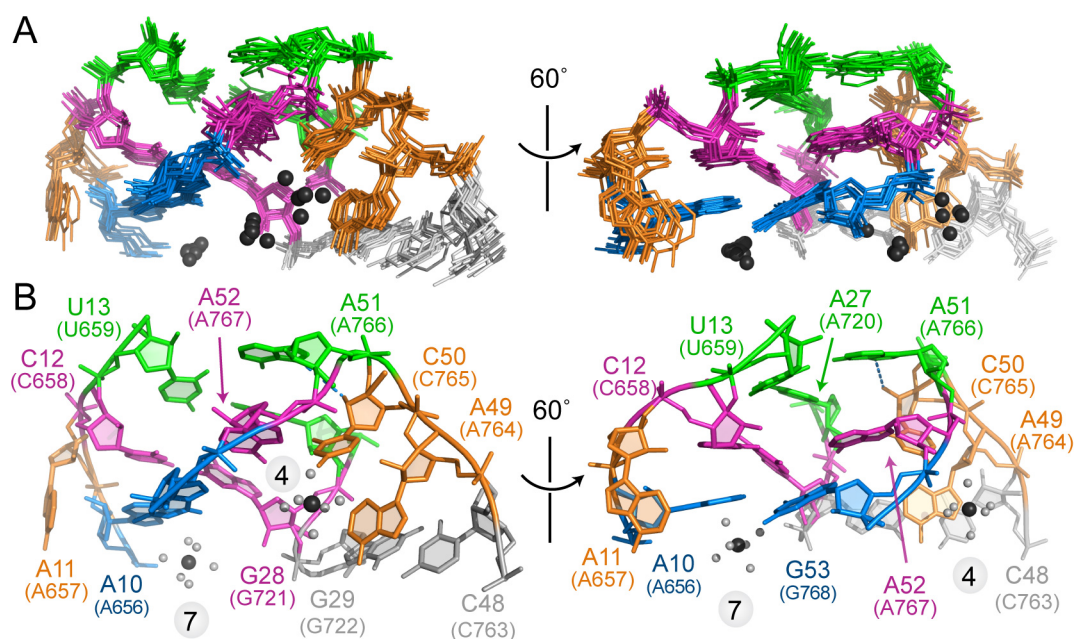


Figure 5.3. NMR structure of the core domain of J236^{Mg}. (A) Superposition of the 10 lowest-energy structures. Only heavy atoms of core residues (10-13, 27-29, 48-53) were used for the superposition and are shown. (B) Stick representations of the lowest-energy structure of the core domain. The Mg²⁺ ions at sites 4 and 7 are shown in dark gray, with their bound water molecules in lighter gray. The cyan dashed line represents a hydrogen bond defined on the basis of a short distance (≤ 4.0 Å) between C₅₀ 2'-OH and A₅₁ N7 in the ensemble of structures.

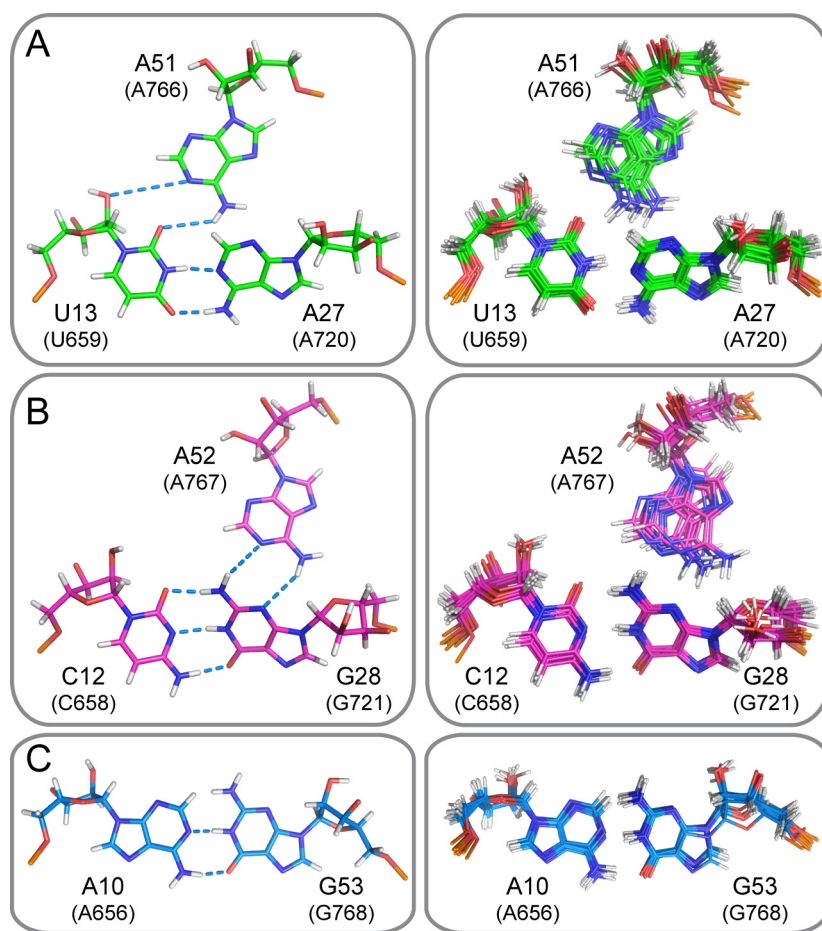


Figure 5.4. Base pairing within the core domain of J236^{Mg}. The two minor groove base triples [(U₁₃-A₂₇)•A₅₁] and (C₁₂-G₂₈)•A₅₂] and the *cis* WC/WC A₁₀•G₅₃ base pair located in the core of J236^{Mg} are shown on the lowest-energy structure (left panels) and on the superposition of the 10 lowest-energy structures (right panels; only the residues shown were used for the superposition). Cyan dashed lines represent hydrogen bonds that are defined on the basis of short distances (≤ 4.0 Å) between heavy atoms in the ensemble of structures.

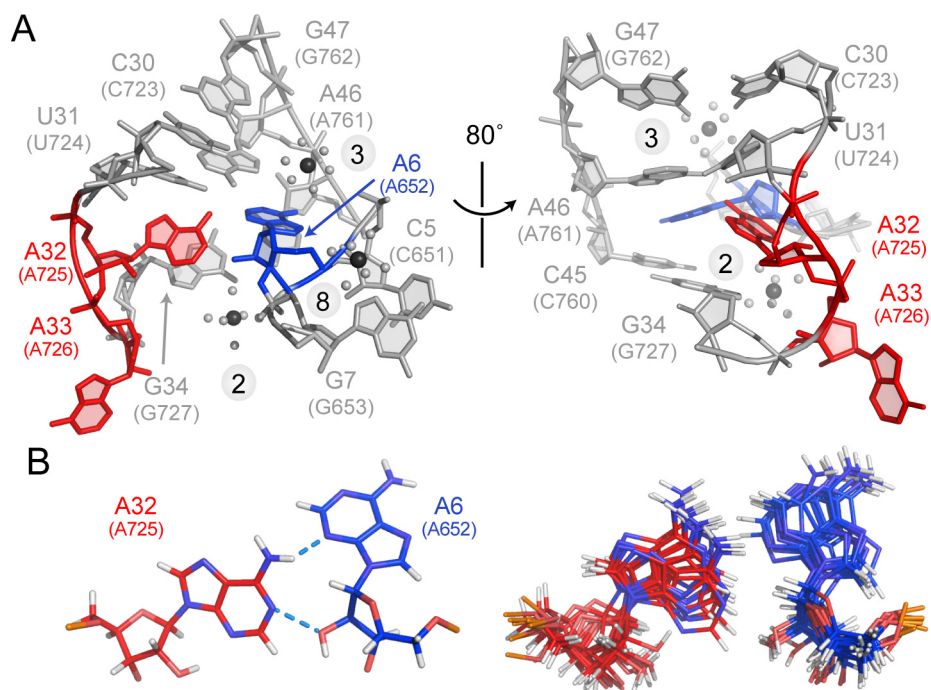


Figure 5.5. NMR structure of the II-VI bulge-bulge interaction of the J236^{Mg} RNA. (A) Stick representations of one of the lowest-energy structure of the II-VI bulge-bulge interaction showing the A₆•A₃₂ base pair (left panel) and the S-turn at the A₃₂A₃₃ bulge (right panel). Only heavy atoms from the bulges, the closing base pairs of stems II and VI and the S-turn at the A₃₂A₃₃ bulge (residues 5-7, 30-34, 45-47) are shown. The Mg²⁺ ions at sites 2, 3 and 8 are shown in dark gray, with their bound water molecules in lighter gray. (B) The *cis* Sugar edge/WC A₆•A₃₂ base pair is shown on one of the lowest-energy structure (left panel) and on the superposition of the 10 lowest-energy structures (right panel; only the residues shown were used for the superposition). Cyan dashed lines represent hydrogen bonds that are defined on the basis of short distances ($\leq 4.0 \text{ \AA}$) between heavy atoms in the ensemble of structures.

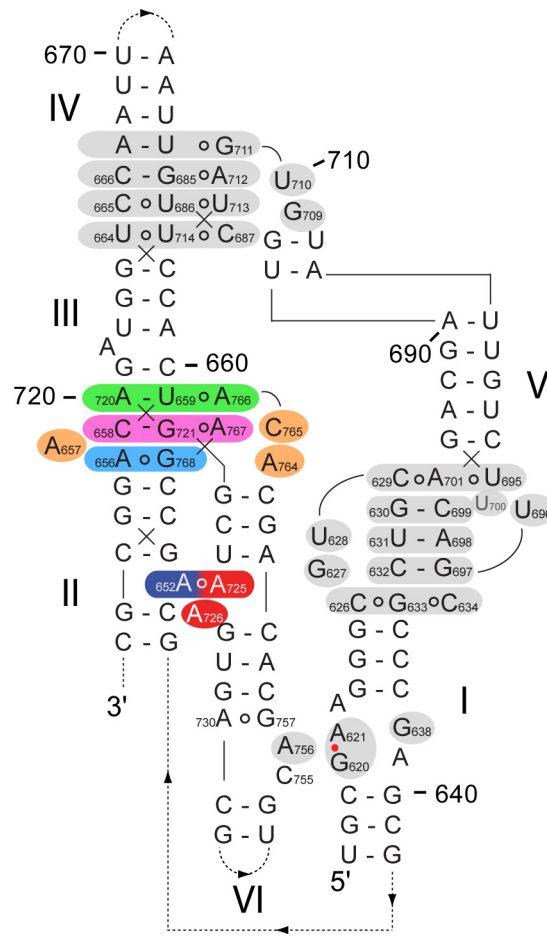


Figure 5.6. Schematic of the tertiary structure of a *cis*-cleaving VS ribozyme derived from NMR solution structures of isolated domains. The shaded residues are particularly important for the structure and activity of the VS ribozyme, including those of the I/V kissing-loop interaction [21], the III-IV-V three-way junction [40], the II-III-VI three-way junction and the active site formed by the association of domains I and VI [23]. For the II-III-VI junction, shaded residues are color-coded according to the structural elements identified in the NMR structure presented here (Figure 5.2). Within the secondary structure, non-essential paired regions are omitted [45], whereas canonical and non-canonical base pairing is represented by solid line and open circles, respectively. The small red circle indicates the cleavage site.

SUPPORTING INFORMATION

Supporting Experimental Procedures

NMR data collection, assignment and structural restraints for the J236 RNA. All NMR experiments were collected on a Varian ^{Unity}INOVA 600 MHz spectrometer equipped with a pulse-field gradient unit and an actively-shielded z-gradient ¹H/¹³C/¹⁵N triple resonance probe. The exchangeable protons of the J236 RNA and their attached nitrogens were assigned at 15°C (Table S5.1 of Supporting Information) from the following experiments in 90% H₂O/10% D₂O: 1D ¹H flip-back watergate [1]; 2D imino- and amino-optimized ¹H-¹⁵N HSQC [2]; 2D HNN-COSY [3]; 2D ¹H-¹⁵N CPMG-NOESY [4]; G-specific H(NC)-TOCSY-(C)H [5]; A-specific H(NC)-TOCSY-(C)H [6]; 3D ¹⁵N-edited NOESY-HSQC (mixing time (τ_m) =150 ms; [7]) and 3D ¹⁵N/¹³C-edited NOESY-HSQC (τ_m =150 ms; [8]). The non-exchangeable protons of the J236 RNA and their attached carbons were assigned at 25°C (Table S5.2 of Supporting Information) from the following experiments in 100% D₂O: 2D ¹H-¹³C constant-time (CT)-HSQC [9, 10], 2D ¹H-¹³C HMQC [11], 2D ¹H-¹⁵N MQ-(HC)N(C)H [12], 3D CT-HCCH-COSY [13], 3D HCCH-TOCSY [13] and 3D ¹³C-edited HMQC-NOESY (τ_m =100 ms; [11]). ¹H, ¹³C and ¹⁵N chemical shifts were referenced at 25°C to an external standard of 2,2-dimethyl-2-silapentane-5-sulfonic acid at 0.00 ppm. NMR data were processed using the NMRPipe/NMRDraw package [14] and analyzed with the CCPNMR suite [15].

A large set of restraints was defined for structure calculation of the J236 RNA (Table 5.1). A HNN-COSY [3] spectrum was collected in H₂O at 15°C to detect ²J_{NN} couplings across WC base pairs. Distance restraints were obtained from 2D ¹H-¹⁵N CPMG-NOESY [4] and 3D ¹⁵N-edited NOESY-HSQC [7] spectra collected in H₂O at 15°C, from a 3D ¹³C/¹⁵N-edited NOESY-HSQC spectrum [8] collected in H₂O at 25°C and from a 3D ¹³C-edited HMQC-NOESY [11] spectrum recorded in D₂O at 25°C. The NOE-derived distance restraints were separated in four classes: strong (1.8-3.3 Å), medium (1.8-4.5 Å), weak (1.8-5.5 Å) and very weak (2.8-7.5 Å) based on NOE crosspeak intensities. Based on NMR evidence from NOESY and 2D HNN-COSY spectra, canonical distance restraints were employed to define the hydrogen-bond pattern and planarity of the WC base pairs in helical regions (SLII: 1-5/57-61 and 7-9/54-56, SLIII: 12-14/26-28 and 16-18/21-23, SLVI: 29-31/46-48 and 34-37/42-45) as well as the *cis* WC-WC A₁₀•G₅₃ base pair. Dihedral angle

restraints for the sugar puckers (δ) and other backbone dihedral angles (α , γ , χ and ζ) were defined from comparative NOE analyses [16]. Based on NMR evidence, backbone torsion angles of residues in helical regions were restrained to A-form values ($\pm 15^\circ$).

RNA-metal distance restraints. Manganese (Mn^{2+}) titrations were performed by adding small volumes of a concentrated solution (0.5 M) of 99.99% MnCl_2 (Sigma-Aldrich, MO) directly to the RNA sample to achieve final concentrations of 5, 10, 20 and 40 μM MnCl_2 . The first Mn^{2+} titration was carried out with a 1.5-mM ^{15}N -labeled J236 sample in NMR buffer A with 5 mM MgCl_2 and 90%:10% $\text{H}_2\text{O}:\text{D}_2\text{O}$, and Mn^{2+} -induced paramagnetic relaxation enhancement (PRE) was monitored by imino- and amino-optimized 2D ^1H - ^{15}N HSQC [2] spectra at 15°C . The second Mn^{2+} titration was carried out with a 2.0-mM $^{13}\text{C}/^{15}\text{N}$ -labeled J236 sample in NMR buffer A with 5 mM MgCl_2 and 100% D_2O , and Mn^{2+} -induced PRE was monitored by 2D ^1H - ^{13}C CT-HSQC [9, 10] spectra at 25°C .

The RNA-metal distance restraints were derived from Mn^{2+} -induced PRE using the ratio of signal intensity (I_o/I_{Mn}) determined from 2D ^1H - ^{15}N HSQC [2] and 2D ^1H - ^{13}C CT-HSQC [9, 10] spectra collected at 0 μM MnCl_2 (I_o) and 20 μM MnCl_2 (I_{Mn}). Signal intensities were determined from the peak intensities of single cross sections of 2D spectra, and it was verified that the use of peak volumes would yield the same results. For each signal in the spectra with $I_o/I_{\text{Mn}} \geq 9$, $9 > I_o/I_{\text{Mn}} \geq 7$, $7 > I_o/I_{\text{Mn}} \geq 4$ and $4 > I_o/I_{\text{Mn}} \geq 3$ distance restraints were set to 1.8-6.5, 1.8-7.0, 1.8-7.5 and 1.8-8.0 Å, respectively, between a defined Mg^{2+} ion and the atom or the pair of atoms giving rise to the observed signal (Table S5.3 of Supporting Information). The effect of the Mn^{2+} titrations at the 5'/3' end of stem II was not considered in this study [17].

SUPPORTING FIGURES

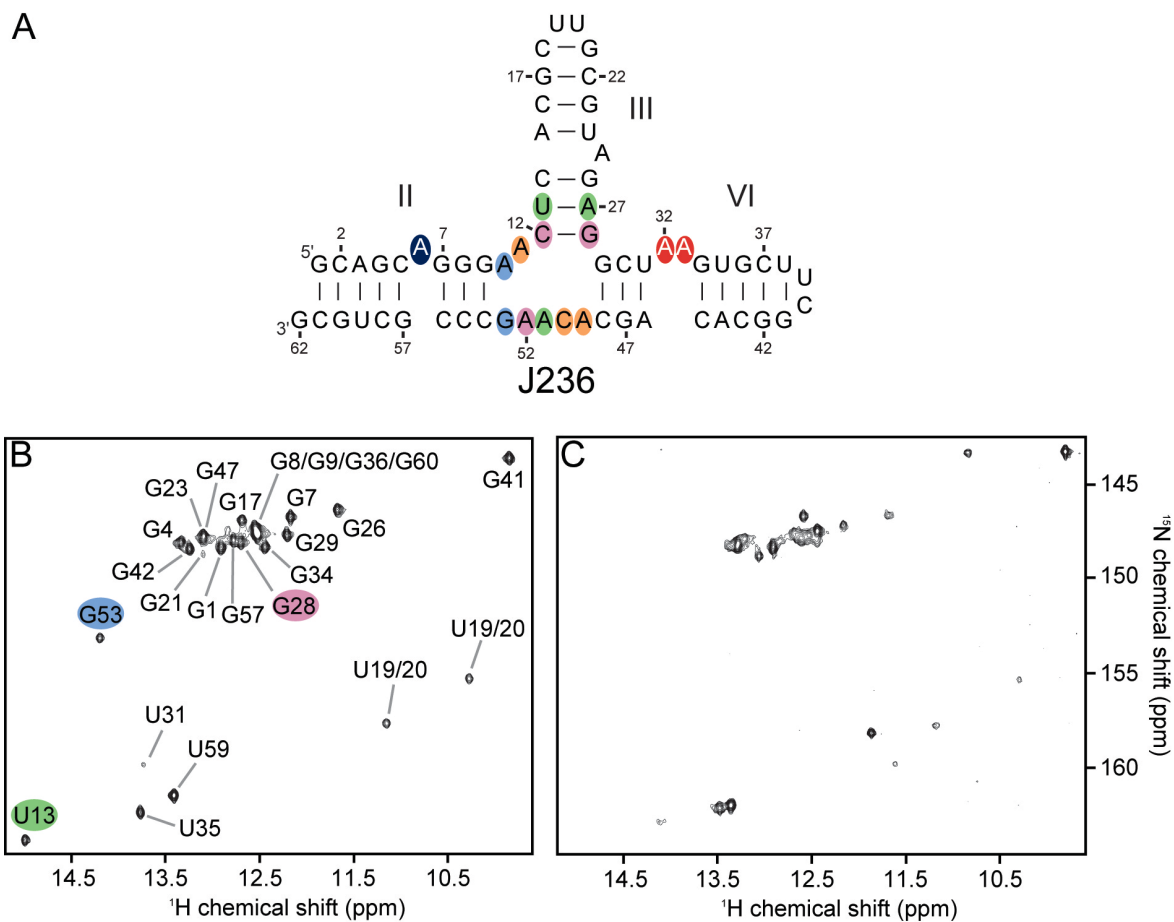


Figure S5.1. The J236 RNA is unstable in the absence of Mg^{2+} ions and when the A_6 bulge of stem II is modified. (A) Primary and secondary structures of the J236 RNA. (B-C) 2D 1H - ^{15}N HSQC spectra of the J236 RNA (B) in the presence of 5 mM $MgCl_2$ and (C) in the absence of Mg^{2+} ions. The shading of residues is color coded according to structural elements present in the NMR structure (see Figure 5.2).

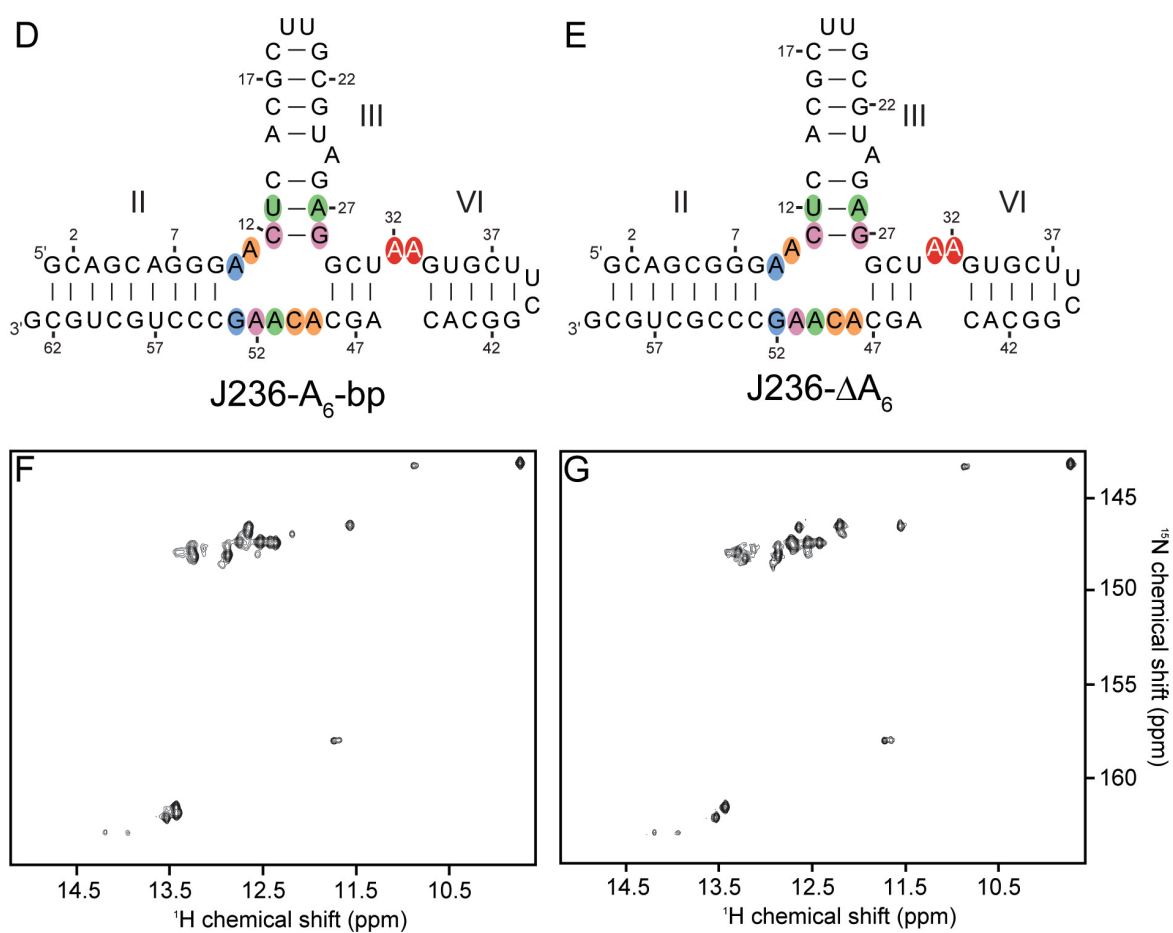


Figure S5.1 (continued). (D-E) Primary and secondary structures of (D) the J236-A₆-bp and (E) the J236-ΔA₆ RNAs. (F-G) 2D ¹H-¹⁵N HSQC spectra of (F) the J236-A₆-bp and (G) the J236-ΔA₆ RNAs, both collected in the presence of 5 mM MgCl₂.

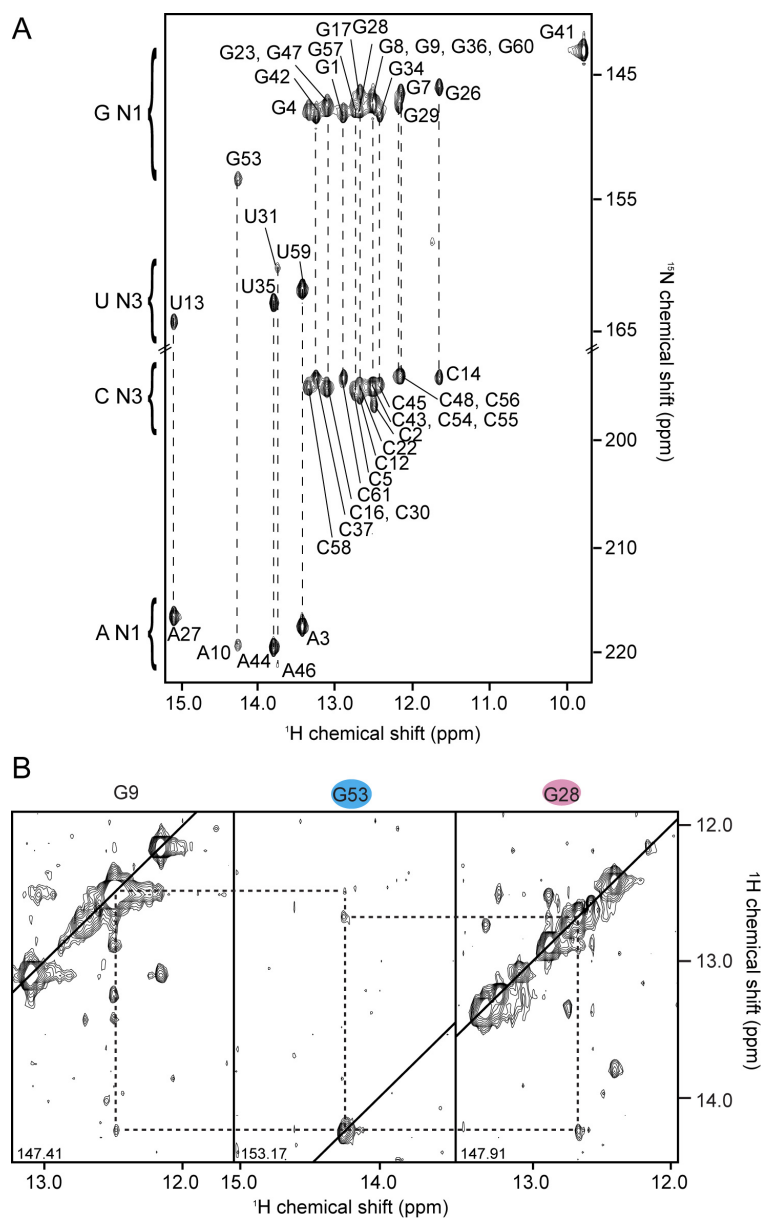


Figure S5.2. NMR evidence of the *cis* WC-WC A₁₀•G₅₃ base pair in stem II and of continuous helical stacking between stems II and III. (A) 2D HNN-COSY spectrum with dashed lines connecting the G N1 – C N3, U N3 – A N1 and G N1 – A N1 correlations [3]. (B) NOE signals between the imino protons of the core residues in 2D ^1H - ^1H slices taken at specific ^{15}N chemical shifts (reported in ppm, bottom left corner) of a 3D ^{15}N -edited NOESY-HSQC spectrum ($\tau_m = 150$ ms; [7]). The dashed lines represent the connection between the NOE signals.

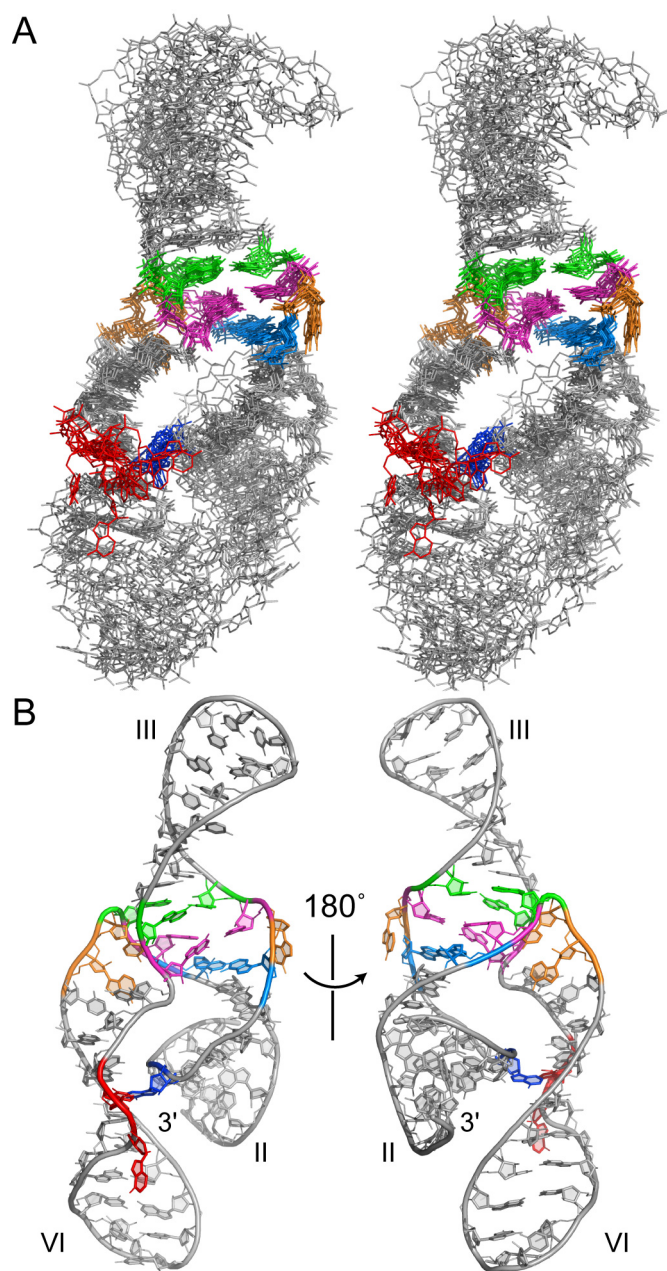


Figure S5.3. NMR solution structure of the J236 RNA. (A) Stereoview of the 10 lowest-energy structures of J236. Only heavy atoms of residues 2-60 were used for the superposition and are shown. **(B)** Stick representations of the lowest-energy structure of the J236 RNA (residues 1-61).

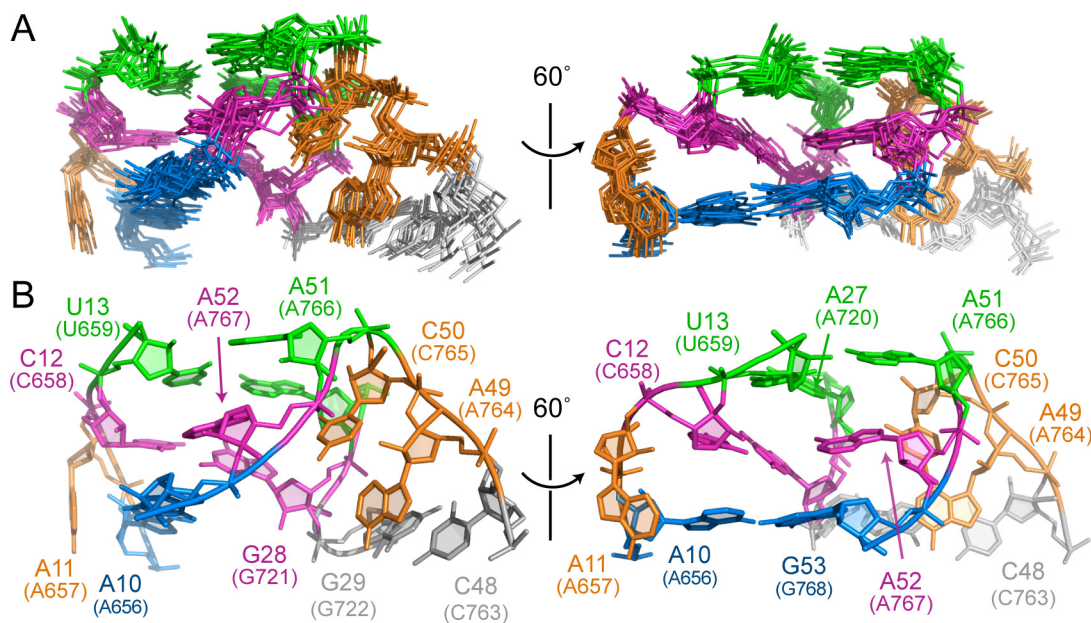


Figure S5.4. NMR structure of the core domain of J236. (A) Superposition of the 10 lowest-energy structures. Only heavy atoms of core residues (10-13, 27-29, 48-53) were used for the superposition and are shown. (B) Stick representations of the lowest-energy structure of the core domain. The cyan dashed line represents a hydrogen bond defined on the basis of a short distance (≤ 4.0 Å) between C₅₀ 2'-OH and A₅₁ N7 in the ensemble of structures.

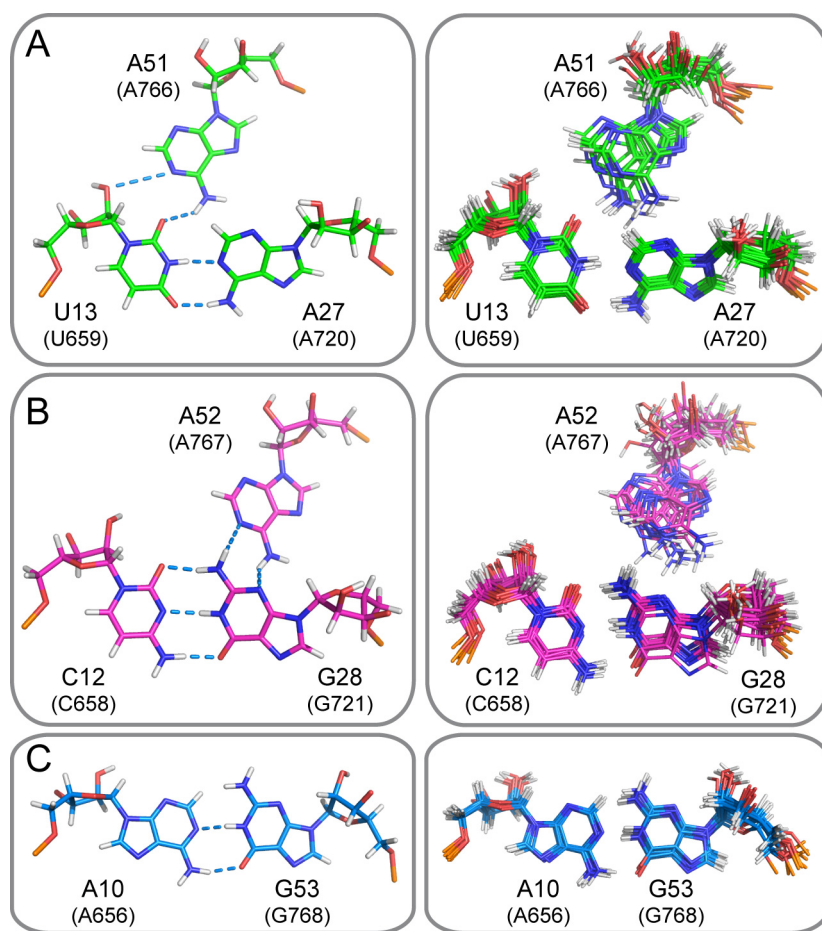


Figure S5.5. Base pair and base triples within the core domain of J236. The two minor base triples [(U₁₃-A₂₇)•A₅₁ and (C₁₂-G₂₈)•A₅₂] and the *cis* WC/WC A₁₀•G₅₃ base pair located in the core of J236 are shown with their hydrogen bonds on the lowest-energy structure (left panels) and on the superposition of the 10 lowest-energy structures (right panels; only the residues shown were used for the superposition). Cyan dashed lines represent hydrogen bonds that are defined on the basis of short distances (≤ 4.0 Å) between heavy atoms in the ensemble of structures.

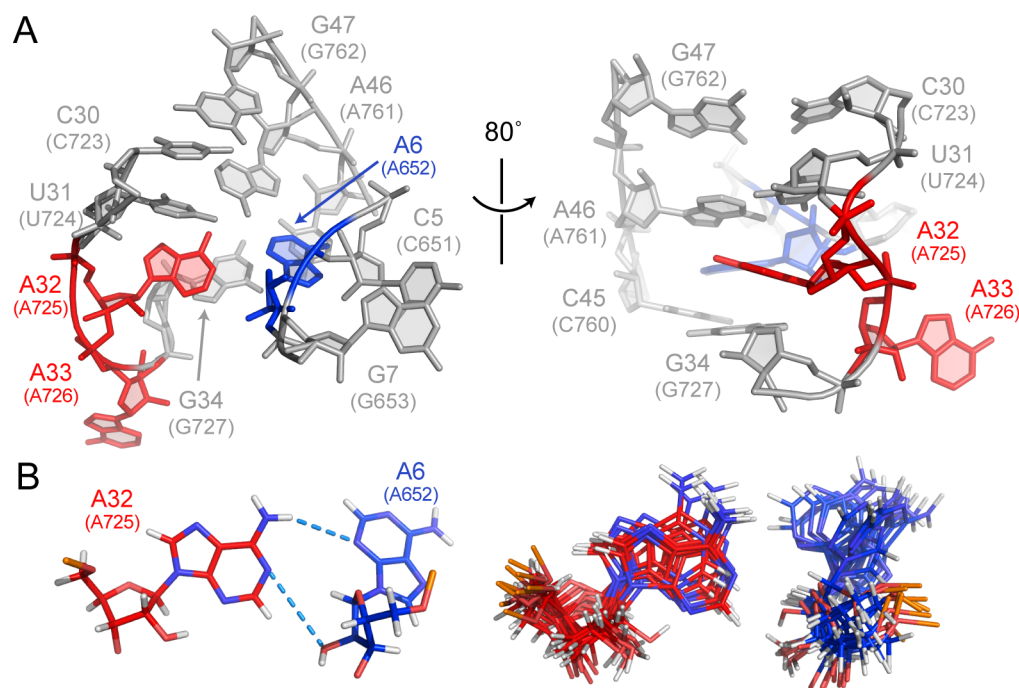


Figure S5.6. NMR structure of the II-VI bulge-bulge interaction of the J236 RNA. (A) Stick representations of one of the lowest-energy structure of the II-VI bulge-bulge interaction showing the $A_6 \bullet A_{32}$ base pair and the S-turn at the $A_{32}A_{33}$ bulge. Only heavy atoms from the bulges, the closing base pairs of stems II and VI and the S-turn at the $A_{32}A_{33}$ bulge (residues 5-7, 30-34, 45-47) are shown. (B) The *cis* Sugar edge-WC $A_6 \bullet A_{32}$ base pair with its hydrogen-bonding pattern shown on one of the lowest-energy structure (left panel) and the superposition of the 10 lowest-energy structures (right panel; only the residues shown were used for the superposition). Cyan dashed lines represent hydrogen bonds that are defined on the basis of short distances (≤ 4.0 Å) between heavy atoms in the ensemble of structures.

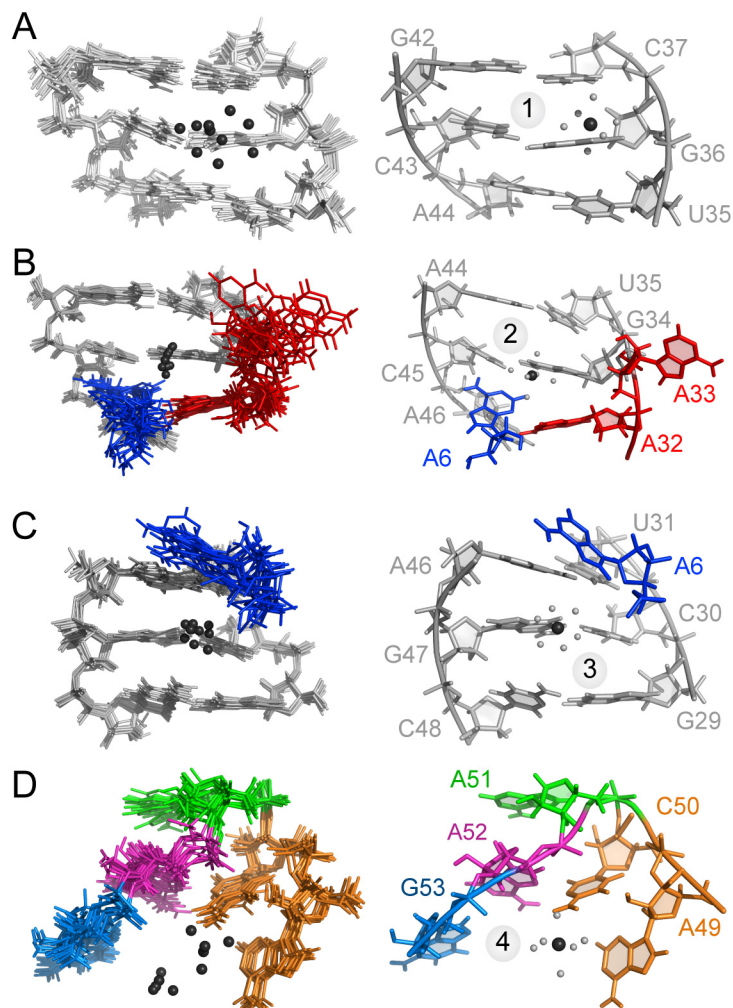


Figure S5.7. Preferential $\text{Mg}(\text{H}_2\text{O})_6^{2+}$ -binding sites within the J236^{Mg} RNA. $\text{Mg}(\text{H}_2\text{O})_6^{2+}$ -binding sites are associated with (A) the major groove of stem VI near the UUCG tetraloop (site 1), (B) the $\text{A}_{32}\text{A}_{33}$ bulge in stem VI (site 2), (C) the major groove of stem VI near the core domain (site 3) and (D) the ACAA turn in the core domain (site 4). For each $\text{Mg}(\text{H}_2\text{O})_6^{2+}$ -binding site (A-H), the superposition of the 10 lowest-energy structures is shown (left panels; only the residues shown were used for the superposition) along with a stick representation of the lowest-energy structure (right panels). The Mg^{2+} ions are shown in dark grey with their bound water molecules in lighter gray.

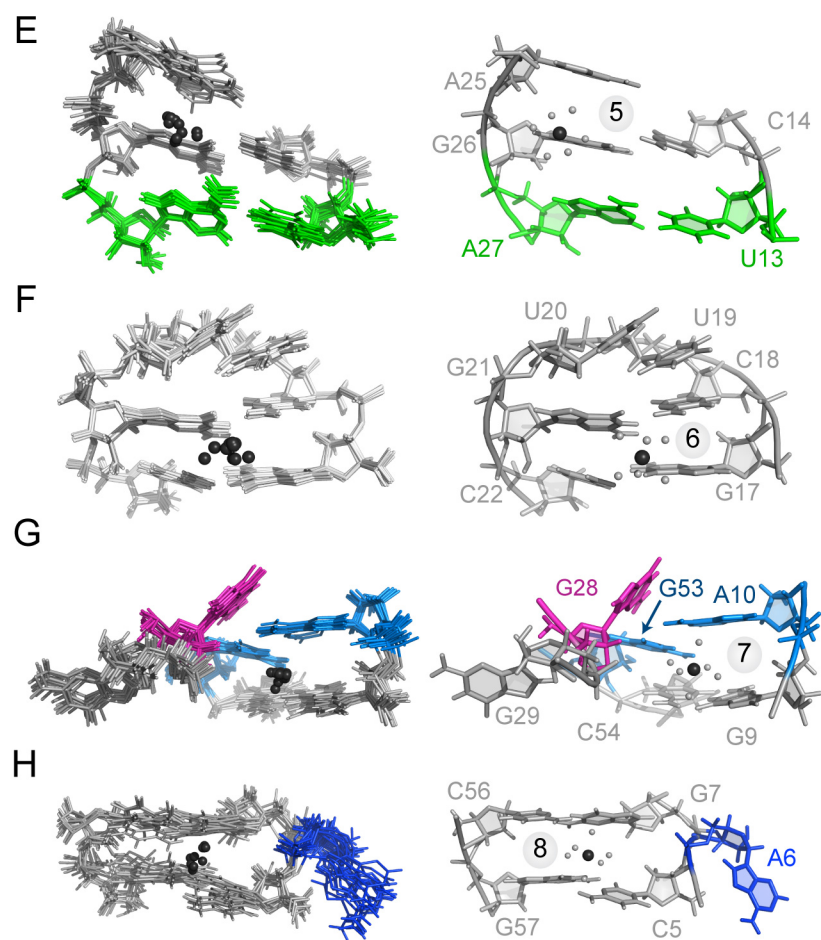


Figure S5.7 (continued). $\text{Mg}(\text{H}_2\text{O})_6^{2+}$ -binding sites are also associated with **(E)** the major groove of stem III near the $(\text{U}_{13}\text{-A}_{27})\cdot\text{A}_{51}$ base triple (site 5), **(F)** the CUUG tetraloop (site 6), **(G)** the major groove of stems II near the $\text{A}_{10}\cdot\text{G}_{53}$ base pair (site 7) and **(H)** the major groove of stem II near the A_6 bulge (site 8).

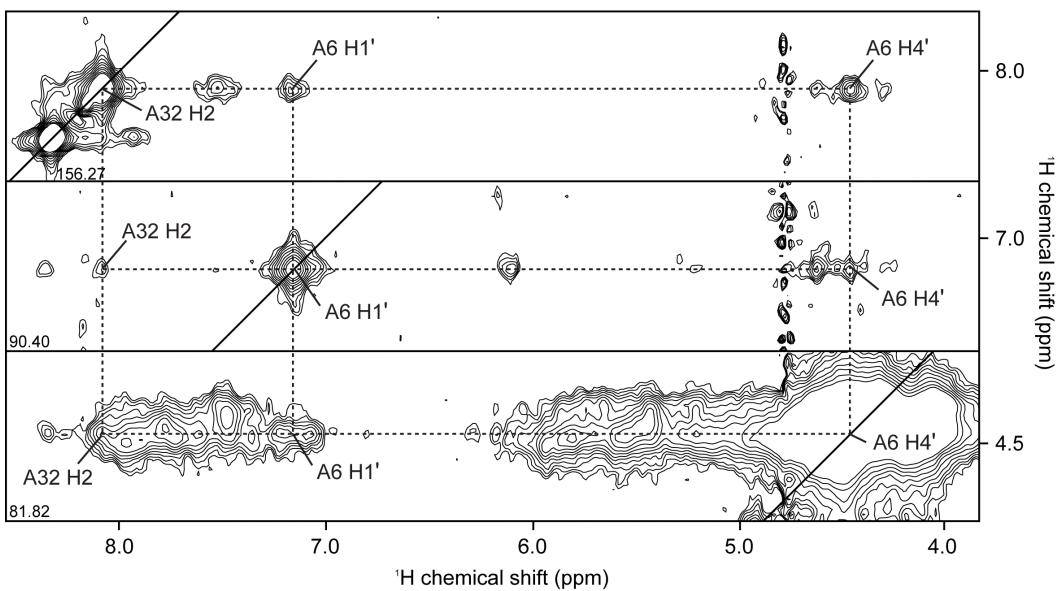


Figure S5.8. NMR evidence of the *cis* Sugar edge/WC A₆•A₃₂ base pair defining the II-VI bulge-bulge interaction. NOE signals between the A₆ and A₃₂ protons are shown from 2D ¹H-¹H slices taken at specific ¹⁵N chemical shifts (reported in ppm, bottom left corner) of a 3D ¹³C-edited HMQC-NOESY spectrum ($\tau_m = 100$ ms; [11]). The connections between the NOE signals are shown by dashed lines.

Table S5.1. Chemical shifts (ppm) of base exchangeable ¹H and their attached ¹⁵N for the J236 RNA.^a

	G H1 U H3	G N1 U N3	G H21/H22 C H41/H42 A H61/H62	G N2 C N4 A N6
G1	12.98	148.1		
C2			8.40/6.73	98.8
A3				
G4	13.42	148.0		
C5			8.32/6.47	97.9
A6				
G7	12.24	146.6		
G8	12.60	147.2		
G9	12.60	147.2		
A10			6.44/6.44	71.9
A11				
C12			8.53/7.27	97.2
U13	15.11	164.1		
C14			8.08/6.86	98.5
A15				
C16			8.08/6.43	97.6
G17	12.78	146.8		
C18			6.95/6.95	95.8
U19	10.30/11.20	155.3/157.8		
U20	10.30/11.20	155.3/157.8		
G21	13.19	148.6		
C22			8.16/6.36	96.2
G23	13.19	147.6		
U24				
A25			6.62/6.62	78.7
G26	11.72	146.2	5.60/5.60	71.3
A27				
G28	12.77	147.9	8.50/7.94	78.1
G29	12.26	147.6		
C30			8.26/6.40	97.1
U31	13.83	160.0		
A32				
A33				
G34	12.51	148.2		
U35	13.86	162.6		
G36	12.60	147.4		
C37			8.40/6.64	99.4
U38	11.70/11.81	158.1/160.2		
U39	11.70/11.81	158.1/160.2		
C40			7.04/6.18	93.7
G41	9.86	143.4		

G42	13.34	148.3		
C43			8.53/6.72	97.8
A44				
C45			7.43/6.43	95.6
A46				
G47	13.19	147.6		
C48			7.48/6.45	97.7
A49				
C50			6.58/6.58	100.5
A51				
A52			6.80/6.80	78.2
G53	14.30	153.1	7.79/6.31	81.4
C54			8.21/6.61	98.0
C55			8.40/6.62	97.7
C56			8.05/6.42	98.0
G57	12.85	147.8		
C58			8.38/6.60	99.4
U59	13.50	161.7		
G60	12.60	147.6		
C61			8.06/6.49	98.5
G62				

^a ¹H resonances were assigned with an error of ± 0.02 ppm and ¹⁵N resonances were assigned with a an error of ± 0.1 ppm

Table S5.2. Chemical shifts (ppm) of non-exchangeable ¹H and their attached ¹³C for the J236 RNA.

Non-exchangeable proton chemical shifts (ppm) ± 0.01 ppm										
	H1'	H2'	H3'	H4'	H5'	H5''	H2	H5	H6	H8
G1	5.84	4.75	4.58	4.43	4.14	3.96				8.13
C2	5.59	4.57	4.62	4.49	4.62	4.10		5.28	7.90	
A3	5.93	4.68	4.64	4.48	4.65	4.16	7.04			7.99
G4	5.60	4.30	4.45	4.26	4.53	4.14				7.22
C5	5.71	4.27	4.73	4.46				5.13	7.59	
A6	7.16	4.62	5.21	4.46	4.31	4.26	6.96			8.36
G7	5.86	4.51	4.26							7.54
G8	5.74	4.59	4.32							7.09
G9	5.59	4.26	4.45							7.07
A10	6.05	4.14	4.85	4.65	4.35	4.21	7.61			8.29
A11	6.32	5.33	4.66	4.77	4.39	4.35	8.15			8.83
C12	5.54	4.89	4.84	4.44	4.65	4.31		6.00	8.41	
U13	5.91	4.78	4.65	4.71				5.75	8.24	
C14	5.21	4.21	4.38	3.37	4.55	4.07		5.68	7.89	
A15	5.75	4.38	4.63	4.32	4.38	3.98	7.15			7.97
C16	5.21	4.04	4.42	4.29	4.50	4.11		5.28	7.48	
G17	6.02	4.68	4.96	4.44	4.41	4.08				7.52
C18	5.82	4.20	4.64	4.44	4.39	4.04		5.54	7.50	
U19	6.11	4.49	4.56	4.37	4.21	4.07		6.01	8.08	
U20	4.97	3.99	4.47	4.30	4.05	3.90		5.35	7.75	
G21	5.63	4.98	4.85	4.46	4.05	3.98				8.02
C22	5.75	4.86	4.64	4.64	4.64	4.35		5.38	7.58	
G23	5.80	4.39	4.58	4.52	4.54	4.18				7.62
U24	5.85	4.32	4.69	4.47				5.12	7.66	
A25	6.16	4.85	4.85	4.58	4.33	4.25	8.14			8.39
G26	5.60	4.79	4.58	4.49	4.57	4.14				7.74
A27	5.87	4.52	4.78	4.52	4.60	4.15	7.47			7.94
G28	6.01	4.20	4.70	4.46	4.62	4.13				7.36
G29	5.93	4.68	4.24	4.42	4.47	4.23				8.36
C30	5.55	4.58	4.57	4.43	4.52	4.09		4.99	7.77	
U31	5.81	4.49	4.56					4.65	7.58	
A32	6.31	4.50	4.55	4.61	4.44	4.07	8.08			8.84
A33	6.26	4.94	4.96	4.75	4.44	4.28	8.32			8.52
G34	4.66	4.30	4.45	4.45	4.53	4.06				7.20
U35	5.54	4.56	4.62	4.48				5.08	7.93	
G36	5.84	4.48	4.60	4.48	4.56	4.14				7.73
C37	5.45	4.25	4.38	4.34	4.51	4.06		5.20	7.47	
U38	5.64	3.77	4.53	4.36	4.54	4.08		5.71	7.77	
U39	6.10	4.66	4.03	4.48	4.24	4.03		5.86	8.02	
C40	5.96	4.11	4.50	3.81	3.62	2.72		6.12	7.70	
G41	5.98	4.84	5.64	4.43	4.42	4.21				7.87

G42	4.43	4.42	4.51	4.44						8.34
C43	5.48	4.28	4.53	4.40				5.26	7.80	
A44	5.94	4.49	4.61	4.55	4.58	4.19	7.28			8.01
C45	5.44	4.09	4.49	4.49	4.60	4.18		4.93	7.37	
A46	5.74	4.76	4.67	4.50	4.62	4.21	7.70			7.91
G47	5.74	4.49	4.64	4.31	4.49	4.05				7.52
C48	5.43	4.44	4.23	4.44	4.51	4.09		5.02	7.47	
A49	5.86	4.46	4.36	4.44	4.47	4.04	6.87			7.77
C50	5.61	4.08	4.54	4.37	4.43	4.02		5.29	7.38	
A51	5.93	4.94	4.58	4.54	4.55	4.20	7.74			8.30
A52	5.87	4.43	4.72	4.52	4.58	4.18	7.94			8.03
G53	5.76	4.53	4.61	4.49	4.54	4.14				7.36
C54	5.60	4.29	4.52	4.46	4.60	4.09		5.16	7.65	
C55	5.47	4.73	4.55					5.32	7.87	
C56	5.42	4.54	4.36	4.36				5.42	7.79	
G57	5.67	4.47	4.65	4.47	4.52	4.14				7.62
C58	5.51	4.53	4.59	4.43	4.53	4.09		5.27	7.79	
U59	5.54	4.56	4.63	4.48				5.41	7.94	
G60	5.79	4.48	4.59	4.48	4.56	4.13				7.72
C61	5.80	4.37	4.43	4.33				5.19	7.45	
G62	5.77	4.03	4.23	4.18						7.61

**Carbon chemical shifts (ppm)
± 0.2 ppm**

	C1'	C2'	C3'	C4'	C5'	C2	C5	C6	C8	
G1	93.3	75.0	73.8	84.7	62.3				138.6	
C2	93.7	75.2	72.1	81.8	64.2		97.0	141.5		
A3	93.1	75.6	72.1	81.9	64.9	152.5			139.3	
G4	92.9	75.2	72.5	82.4	64.9				135.5	
C5	92.3	78.1	72.7	81.7			97.4	140.9		
A6	90.4	77.7	73.9	81.8	65.7	152.7			143.0	
G7	93.3	74.5	75.2						136.6	
G8	92.6	75.2	73.2						136.3	
G9	93.7	74.8	72.5						136.6	
A10	93.1	76.2	74.7	82.6	67.5	154.1			140.5	
A11	88.5	76.0	78.7	87.5	68.4	155.4			143.3	
C12	94.9	75.2	72.5	85.1	65.0		97.5	143.2		
U13	91.7	74.8	71.8	82.4			102.6	142.7		
C14	93.8	74.6	72.3	81.6	64.2		97.4	141.5		
A15	92.7	75.4	72.1	81.8	64.2	153.7			139.2	
C16	93.3	75.2	71.7	81.5	64.2		97.6	139.9		
G17	91.5	84.7	78.4	85.9	65.9				138.0	
C18	89.5	75.8	72.1	81.8	67.1		99.1	141.9		
U19	89.3	76.4	79.9	86.9	67.8		105.5	144.3		
U20	89.4	77.7	80.5	85.3	68.4		104.4	144.2		
G21	86.9	74.2	77.5	85.3	69.8				137.8	
C22	94.2	72.5	83.2	83.6	70.0		98.0	141.9		
G23	93.3	75.4	73.3	82.4	66.1				136.4	

U24	91.0	76.4	76.0	85.3			103.8	142.3		
A25	90.4	76.0	76.0	84.7	67.5	155.4			142.2	
G26	93.7	74.8	72.5	81.8	64.9				137.8	
A27	92.9	75.6	72.3	81.9	64.9	153.2			139.5	
G28	91.5	77.0	71.9	81.2	64.9				136.1	
G29	94.7	75.6	74.1	83.2	71.0				139.1	
C30	93.6	75.3	71.9	81.6	64.3		96.3	141.9		
U31	92.6	76.0	72.3				102.3	141.9		
A32	88.7	75.6	72.5	82.4	65.3	156.3			141.9	
A33	89.1	76.2	78.4	86.7	68.6	155.8			142.5	
G34	93.1	74.6	72.5	82.0	64.8				138.0	
U35	93.3	75.2	72.2	82.0			102.5	142.1		
G36	92.4	75.4	72.5	81.8	64.9				136.2	
C37	93.7	75.6	72.1	82.0	64.5		97.3	140.3		
U38	94.5	76.0	73.1	82.4	64.5		105.0	140.7		
U39	89.3	74.6	77.8	86.9	67.8		105.5	144.6		
C40	89.2	77.6	80.3	84.5	67.3		98.5	142.8		
G41	94.6	77.2	75.8	83.1	69.0				143.0	
G42	93.3	75.1	73.7	82.4					139.0	
C43	93.7	75.6	71.9	81.7			97.0	141.3		
A44	92.9	75.5	72.5	81.8	65.1	152.9			139.3	
C45	93.5	75.6	71.9	81.8	64.6		96.6	140.5		
A46	92.8	76.2	73.9	81.8	64.6	153.7			140.3	
G47	92.7	75.5	72.3	82.0	64.2				136.2	
C48	93.7	75.4	71.8	81.7	64.2		97.0	139.9		
A49	93.7	75.2	72.9	82.4	65.3	153.7			138.9	
C50	91.6	76.4	76.2	84.2	65.9		97.6	141.5		
A51	93.9	75.1	74.5	83.0	64.5	152.9	141.5			
A52	91.8	76.4	72.7	81.9	64.9	154.7			139.3	
G53	92.2	75.6	72.5	82.0	65.0				136.2	
C54	94.5	75.2	71.9	82.4	64.2		97.0	140.9		
C55	94.5	75.0	71.9				97.0	141.5		
C56	93.9	75.6	72.2	81.8			97.8	141.3		
G57	92.9	75.1	72.2	81.8	64.2				136.0	
C58	93.8	75.2	71.9	81.8	64.2		97.0	140.7		
U59	93.3	75.2	72.2	82.0			103.4	141.6		
G60	92.7	75.4	72.5	81.8	65.0				136.2	
C61	92.7	74.7	71.8	81.6			97.3	139.9		
G62	91.7	77.8	70.3	83.9					137.5	

Table S5.3. Distance restraints derived from Mn²⁺-induced PRE and used for the J236^{Mg} structure calculation.

Site	Atoms ^a	I/lo	Restraint
Site 1	C37 N4-H42	9	1.8-6.5 Å
	C37 H41	4	1.8-8.0 Å
	C43 N4-H41	4	1.8-8.0 Å
Site 1 or 2	U35 C5-H5	3	1.8-8.0 Å
Site 2	G34 C8-H8	5	1.8-7.5 Å
	C45 N4-H41	5	1.8-7.5 Å
	C45 H42	4	1.8-8.0 Å
	A32 C1'-H1'	3	1.8-8.0 Å
	A33 C5'-H5''	3	1.8-8.0 Å
	C45 C5-H5	3	1.8-8.0 Å
Site 2 or 3	A46 C8-H8	3	1.8-8.0 Å
Site 3	C30 N4-H42	8	1.8-7.0 Å
	C30 H41	6	1.8-7.5 Å
	C48 N4-H41	5	1.8-7.5 Å
	C48 H42	4	1.8-8.0 Å
	C30 C5-H5	3	1.8-8.0 Å
Site 3 or 7	A6 C8-H8	3	1.8-8.0 Å
	A6 C5'-H5'	3	1.8-8.0 Å
	A6 H5''	3	1.8-8.0 Å
Site 4	C50 N4-H41 or H42	8	1.8-7.0 Å
	A53 C3'-H3'	3	1.8-8.0 Å
Site 5 or 6	C16 N4-H41	4	1.8-8.0 Å
	C16 H42	4	1.8-8.0 Å
	U24 C2'-H2'	3	1.8-8.0 Å
Site 5	G26 C8-H8	10	1.8-6.5 Å
	C14 N4-H42	7	1.8-7.0 Å
	C14 H41	5	1.8-7.5 Å
Site 5 or 8	A27 C8-H8	3	1.8-8.0 Å
Site 6	U19 N3-H3	6	1.8-7.5 Å
	U20 N3-H3	5	1.8-7.5 Å
	G21 N1-H1	3	1.8-8.0 Å
	G21 C8-H8	3	1.8-8.0 Å
Site 7	C12 N4-H42	5	1.8-7.5 Å
	G28 N2-H21	5	1.8-7.5 Å
	G28 H22	4	1.8-8.0 Å
	G28 C2'-H2'	3	1.8-8.0 Å
Site 7 or 8	C58 N4-H42	9	1.8-6.5 Å
	C54 N4-H41	7	1.8-7.0 Å
	C55 N4-H41	7	1.8-7.0 Å
	C56 N4-H42	6	1.8-7.5 Å
	G9 C8-H8	6	1.8-7.5 Å
	G8 C8-H8	5	1.8-7.5 Å
	C58 H41	4	1.8-8.0 Å

	C54 H42	4	1.8-8.0 Å
	G53 N2-H21	4	1.8-8.0 Å
	G53 H22	3	1.8-8.0 Å
Site 7 or 8	A10 C8-H8	3	1.8-8.0 Å
	G9 C2'-H2'	3	1.8-8.0 Å
	G8 C3'-H3'	3	1.8-8.0 Å
	G7 N1-H1	11	1.8-6.5 Å
Site 8	C5 C2'-H2'	3	1.8-8.0 Å
	C5 C3'-H3'	3	1.8-8.0 Å

^aRestraints are defined between the listed atoms and the magnesium ion at this site. For those cases where both the heavy atom and the attached proton are listed, they were both used to define two individual restraints to the metal at this site. For those cases where the metal-binding site is not well defined (e.g. Site 1 or 2), ambiguous restraints were defined.

Table S5.4. Structural characteristics of Mg(H₂O)₆²⁺-binding sites in J236^{Mg}.

Site	Outer-sphere ligands ^a	Electrostatic interactions ^b	Cation- π interactions ^c
Site 1	G ₃₆ O6	G ₃₆ 5'-PO ₄ ⁻	
	G ₃₆ N7		
	G ₃₆ 5'- <i>pro-R</i> _p OP		
Site 2	A ₆ O2'	G ₇ 5'-PO ₄ ⁻	
	G ₃₄ O6	G ₃₄ 5'-PO ₄ ⁻	
	G ₃₄ N7		
Site 3	A ₄₆ N7	C ₅ 5'-PO ₄ ⁻	
	G ₄₇ O6	A ₆ 5'-PO ₄ ⁻	
	G ₄₇ N7	A ₄₆ 5'-PO ₄ ⁻	
		G ₄₇ 5'-PO ₄ ⁻	
Site 4	A ₄₉ O2'	A ₅₂ 5'-PO ₄ ⁻	
	G ₅₃ 5'- <i>pro-R</i> _p OP	G ₅₃ 5'-PO ₄ ⁻	
		C ₅₄ 5'-PO ₄ ⁻	
Site 5	G ₂₆ 5'- <i>pro-R</i> _p OP	A ₂₅ 5'-PO ₄ ⁻ G ₂₆ 5'-PO ₄ ⁻	
	G ₂₆ O6		
	G ₂₆ N7		
	A ₂₇ N7		
Site 6	G ₁₇ O6		
	G ₂₁ O6		
	G ₂₁ N7		
Site 7	G ₉ 5'- <i>pro-R</i> _p OP	G ₉ 5'-PO ₄ ⁻ A ₁₀ 5'-PO ₄ ⁻ G ₂₉ 5'-PO ₄ ⁻	
	G ₉ N7		
	A ₁₀ N7		
	G ₂₈ O2'		
	G ₅₃ O6		
Site 8	G ₇ O6	A ₆ 5'-PO ₄ ⁻	C ₅
	G ₅₇ O6		

^aOuter-sphere ligands were determined by selecting all hydrogen-bond acceptors within 4.0 Å of a Mg(H₂O)₆²⁺ oxygen and in an favorable orientation for outer-sphere interaction in at least 25% of the lowest-energy structures presenting an occupancy at a site.

^bElectrostatic interactions were determined by selecting 5'-PO₄⁻ within 8.0 Å of a Mg²⁺ ion.

^cCation- π interactions were determined for Mg²⁺ ions within 5.2 Å and at an angle smaller than 52° from the ring centroid [18].

REFERENCES (For supporting information only)

1. Piotto, M., V. Saudek, and V. Skleňár, *Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions*. J. Biomol. NMR, 1992. **2**: p. 661-665.
2. Kay, L.E., P. Keifer, and T. Saarinen, *Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity*. J. Am. Chem. Soc., 1992. **114**: p. 10663-10665.
3. Dingley, A.J. and S. Grzesiek, *Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide $^2J_{NN}$ couplings*. J. Am. Chem. Soc., 1998. **120**: p. 8293-8297.
4. Mueller, L., P. Legault, and A. Pardi, *Improved RNA structure determination by detection of NOE contacts to exchange-broadened amino groups*. J. Am. Chem. Soc., 1995. **117**: p. 11043-11048.
5. Simorre, J.P., et al., *Correlation of the guanosine exchangeable and non-exchangeable base protons in $^{13}C/^{15}N$ -labeled RNA with an HNC-TOCSY-CH experiment*. J. Biomol. NMR, 1996. **7**: p. 153-156.
6. Simorre, J.P., et al., *Triple-resonance experiments for assignment of adenine base resonances in $^{13}C/^{15}N$ -labeled RNA*. J. Am. Chem. Soc., 1996. **118**: p. 5316-5317.
7. Zhang, O., et al., *Backbone 1H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques*. J. Biomol. NMR, 1994. **4**(6): p. 845-58.
8. Pascal, S.M., et al., *Simultaneous acquisition of ^{15}N and ^{13}C -edited NOE spectra of proteins dissolved in H_2O* . J. Magn. Res. Series B, 1994. **103**: p. 197-201.
9. Vuister, G.W. and A. Bax, *Resolution enhancement and spectral editing of uniformly ^{13}C -enriched proteins by homonuclear broadband ^{13}C decoupling*. J. Magn. Reson., 1992. **98**: p. 428-435.
10. Santoro, J. and G.C. King, *A constant-time 2D overboderhausen experiment for inverse correlation of isotopically enriched species*. J. Magn. Reson., 1992. **97**: p. 202-207.
11. Ikura, M., et al., *Three-dimensional NOESY-HMQC spectroscopy of a ^{13}C -labeled protein*. J. Magn. Reson., 1990. **86**: p. 204-209.

12. Skleňár, V., et al., *Optimization of triple-resonance HCN experiments for application to larger RNA oligonucleotides*. J. Magn. Reson., 1998. **130**: p. 119-124.
13. Pardi, A. and E.P. Nikonowicz, *Simple procedure for resonance assignment of the sugar protons in ¹³C-labeled RNAs*. J. Am. Chem. Soc., 1992. **114**: p. 9202-9203.
14. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes*. J. Biomol. NMR, 1995. **6**: p. 277-293.
15. Vranken, W.F., et al., *The CCPN data model for NMR spectroscopy: development of a software pipeline*. Proteins, 2005. **59**(4): p. 687-96.
16. Wijmenga, S.S., M.M.W. Mooren, and C.W. Hilbers, *NMR of nucleic acids; from spectrum to structure*, in *NMR of macromolecules: a practical approach.*, G.C.K. Roberts, Editor. 1993, Oxford University Press: New York. p. 217-288.
17. Bonneau, E. and P. Legault, *NMR localization of divalent cations at the active site of the Neurospora VS ribozyme provides insights into RNA-metal-ion interactions*. Biochemistry, 2014. **53**(3): p. 579-90.
18. McFail-Isom, L., X. Shui, and L.D. Williams, *Divalent cations stabilize unstacked conformations of DNA and RNA by interacting with base pi systems*. Biochemistry, 1998. **37**(49): p. 17105-11.

6. Discussion

Le ribozyme VS, qui catalyse des réactions de clivage et de ligation essentielles au mécanisme de réplication de l'ARN VS, possède plusieurs caractéristiques qui le distinguent des autres petits ribozymes nucléolytiques. Entre autre, il possède des structures primaire, secondaire et tertiaire différentes des autres ribozymes naturels lui permettant de reconnaître et cliver son substrat. L'activité du ribozyme est dépendante de la formation de deux interactions tertiaires importantes : une interaction *kissing-loop* entre les boucles terminales de SLI et SLV qui permet la reconnaissance du substrat ainsi qu'une interaction entre la boucle interne de SLI et la boucle A730 de SLVI qui permet la formation du site actif. Comme le ribozyme VS est le seul ribozyme de la classe des petits ribozymes nucléolytiques à reconnaître un substrat en tige-boucle, il pourrait être exploité pour cibler des ARN repliés pour diverses applications. Cependant, bien que le ribozyme VS ait été grandement caractérisé biochimiquement, il est le seul ribozyme de sa famille pour lequel aucune structure complète à haute résolution n'a encore été publiée. Ainsi, la compréhension des mécanismes inhérents au fonctionnement du ribozyme VS est limitée par l'absence d'une telle structure tridimensionnelle.

Précédemment, une approche structurale de type *divide-and-conquer* a été initiée sur le ribozyme VS. Dans le cadre de cette approche, des études visant à déterminer la structure à haute résolution par spectroscopie RMN de certains sous-domaines importants ont révélé les éléments structuraux intrinsèques du substrat sous sa forme *unshifted* et *shifted* [228, 253, 257], de l'hélice SLV en absence et en présence de Mg^{2+} [230, 231] ainsi que de l'interaction *kissing-loop* I/V [227]. Bien que ces structures aient contribué à l'élucidation des bases moléculaires impliquant la réorganisation conformationnelle et la reconnaissance du substrat, très peu de données structurales étaient disponibles sur le domaine catalytique du ribozyme VS. Au début de mon doctorat, plusieurs questions subsistaient donc au niveau de la structure du domaine catalytique, de la formation du site actif et du rôle structural des ions métalliques. Afin de répondre à ces questions, j'ai caractérisé par RMN les structures de la boucle A730 de SLVI, de la jonction III-IV-V et de la jonction II-III-VI du domaine catalytique. Ces structures ont révélé la contribution de ces trois sous-domaines à l'architecture du domaine catalytique ainsi qu'à la formation des interactions tertiaires clés du ribozyme VS de *Neurospora*.

6.1. Retour sur les résultats principaux de l'étude

6.1.1. Le site actif du ribozyme VS est pré-organisé en présence d'ions Mg^{2+}

Nos études par spectroscopie de RMN de la boucle A730 ont mené à une compréhension plus détaillée de l'organisation du site actif du ribozyme VS. En effet, la boucle A730 adopte une structure stable et unique qui est caractérisée par la formation de deux éléments structuraux essentiels : une paire de bases *cis* WC/WC A-G et un motif *S-turn* (Chapitre 2). Le motif *S-turn* crée une large distorsion dans la chaîne ribose-phosphate qui exclut le A₇₅₆ de l'hélice, en accord avec le rôle de la boucle A730 dans le positionnement du A₇₅₆ et dans la formation du site actif [220, 223, 224, 237-239, 244-246, 248, 251, 289].

De plus, nous avons déterminé que la structure de la boucle A730 requiert la présence d'ions Mg^{2+} pour être stabilisée (Chapitre 2). Afin de caractériser plus en détail la dépendance aux ions de la boucle A730, nous avons localisé les ions Mg^{2+} dans cette structure (Chapitre 3). Cette étude a été réalisée en combinant deux approches RMN soit la relaxation paramagnétique induite par le Mn^{2+} (*Mn²⁺-induced paramagnetic relaxation enhancement*) et la perturbation de déplacement chimique induite par le Cd^{2+} (*Cd²⁺-induced chemical shift perturbation*) d'ARN ayant une modification phosphorothioate. Au total, cinq sites de liaison de métaux ont été identifiés dans l'ARN SLVI (Chapitre 3). Plus particulièrement, deux sites de liaison de métaux ont été identifiés dans la boucle A730. L'un d'eux est coordonné au motif *S-turn*; il forme une interaction *inner-sphere* avec un groupement 5'-phosphate du *S-turn* ainsi que des interactions *outer-sphere* avec deux groupements 5'-phosphates avoisinants. La localisation de cet ion Mg^{2+} au motif *S-turn* explique donc la dépendance de la boucle A730 aux ions divalents puisqu'en absence d'ions, les groupements phosphates du *S-turn* ne pourraient vraisemblablement pas se rapprocher suffisamment pour former un motif stable.

D'un point de vue fonctionnel, la boucle A730 est essentielle à l'activité du ribozyme VS puisqu'elle forme le site actif en s'associant avec la boucle interne de SLI [220, 223, 224, 237-239, 244-246, 248, 251, 289]. Nos études RMN ont démontré que la boucle A730 est pré-organisée en présence d'ions Mg^{2+} , adoptant une architecture précise et

ce, préalablement à la formation du site actif. De plus, la structure de la boucle A730 concorde avec les données biochimiques et mécanistiques du ribozyme VS puisque la formation du motif *S-turn* permet de positionner le A₇₅₆ dans le sillon mineur élargi de l'hélice, le rendant ainsi disponible pour participer à l'activité catalytique [239, 248]. En effet, il a été proposé que la position N1 protonée du A₇₅₆ participe directement au mécanisme enzymatique du ribozyme VS en jouant le rôle d'acide général [220, 239, 245]. La structure RMN de la boucle A730 est aussi en accord avec les données biophysiques du ribozyme VS puisque, l'angle interhélicale entre les hélices de part et d'autre de la boucle A730 a été déterminé à ~150° à partir de la structure RMN tandis que l'angle correspondant dans le modèle de SAXS est de ~130°, suggérant que la structure RMN de la boucle A730 est similaire à la structure de la boucle A730 dans le cadre du ribozyme VS complet. [260]. Par contre, la structure RMN de la boucle A730 présentée dans cette thèse diffère de la structure RMN déterminée précédemment pour cette boucle (Section 1.5.4; [258]). Il est à noter que cette structure précédente n'était pas en accord avec les données biochimiques et ne permettait donc pas de rationaliser le rôle du A₇₅₆ dans la catalyse (voir section 1.5.4).

La détermination de la structure de la boucle A730 a aussi offert de nouvelles bases comparatives avec le ribozyme *hairpin*. En effet, plusieurs similarités étaient déjà répertoriées entre les sites actifs des ribozymes *hairpin* et VS [205, 206]. Ces deux ribozymes utilisent les mêmes nucléobases catalytiques, soit une guanine comme base générale (G₈ [*hairpin*] et G₆₃₈ [VS]) et une adénine comme acide général (A₃₈ [*hairpin*] et A₇₅₆ [VS]). Leur site actif est essentiellement formé de la même manière, soit par l'association de deux boucles internes (boucle A et B [*hairpin*] et boucle interne de SLI et boucle A730 [VS]). De plus, l'architecture de la boucle interne de SLI et de la boucle A présente des similarités structurales puisque dans chacun des cas, la base générale (G₆₃₈ ou G₈) forme un empilement de purines inter-brin avec le résidu en amont du site de clivage. La structure RMN de la boucle A730 a aussi permis de découvrir une nouvelle similarité dans l'organisation des sites actifs de ces deux ribozymes. En effet, la structure de la boucle B du ribozyme *hairpin*, contenant le A₃₈, repose également sur la formation d'un motif *S-turn* qui permet de positionner le A₃₈ pour la catalyse [203]. De plus, ce motif *S-turn* chez le ribozyme *hairpin* forme aussi un site de liaison de métaux connu [50]. En se basant sur ces similarités et sur la structure cristallographique de l'état de transition du ribozyme *hairpin*, il fut donc possible de modéliser, par homologie, la structure du site actif du ribozyme VS

[203]. Très peu de changements ont dû être apportés aux structures de la boucle interne *shifted* de SLI et de la boucle A730 pour former un site actif similaire à celui du ribozyme *hairpin*, si ce n'est que de positionner les résidus G₆₂₀ et A₆₂₁ dans une conformation étendue (*splayed*) qui est partagée par tous les autres petits ribozymes nucléolytiques à l'état pré-clivage [111]. En plus de fournir une base rationnelle cohérente aux données biochimiques concernant le site actif du ribozyme VS, ce modèle renforce l'idée que l'organisation du site actif des ribozymes VS et *hairpin* est similaire. Toutefois, des études supplémentaires seront nécessaires pour confirmer ce modèle.

6.1.2. La jonction III-IV-V est un domaine architectural qui facilite la formation de l'interaction *kissing-loop* I/V

L'architecture du domaine catalytique est déterminée par le repliement de deux jonctions à trois voies dont les structures n'étaient précédemment pas connues. Dans cette optique, nous avons déterminé la structure RMN de la jonction III-IV-V qui est nécessaire pour la formation de l'interaction *kissing-loop* I/V (Chapitre 4). La jonction III-IV-V adopte un repliement typique de la famille C de jonction à trois voies. Sa structure globale est caractérisée par un empilement coaxial des tiges III et IV tandis que les tiges III et V définissent un angle aigu. Elle implique la formation d'un motif *U-turn* qui oriente SLV au niveau de la jonction, tout en positionnant les deux résidus suivant le U du *U-turn* dans une orientation leur permettant d'interagir avec le sillon mineur de la tige IV. De plus, le cœur de la jonction est formé par l'empilement consécutif de quatre triplets de bases qui favorisent l'empilement coaxial des tiges III et IV. La jonction est également stabilisée par la formation d'un motif *ribose zipper*. Ce réseau d'interaction permet donc de définir l'architecture de la jonction III-IV-V.

La localisation des sites de liaison de métaux dans la jonction III-IV-V a été réalisée par l'approche de relaxation paramagnétique induite par le Mn²⁺ (Chapitre 3). Six sites de liaison de métaux ont été identifiés dont deux se situent au cœur de la jonction. Un des ions Mg²⁺ interagit avec le motif *U-turn* tandis que l'autre contacte le sillon majeur de la tige III et la chaîne ribose-phosphate de SLV. Ces données suggèrent que la liaison de ces métaux contribue à l'architecture de la jonction III-IV-V.

La structure RMN de la jonction III-IV-V est en accord avec les différentes données biochimiques et structurales disponibles [215, 218, 222, 223, 233, 242, 243, 246, 247, 260, 290, 291], ce qui permet d'en dériver des implications fonctionnelles. Entre autre, la jonction III-IV-V est la seule jonction de la famille C à ne pas former d'interaction tertiaire entre les deux tiges positionnées côte-à-côte. Cet élément semble être important pour l'activité du ribozyme VS en affectant le positionnement de SLV et donc, de la reconnaissance du substrat. L'ensemble de structures RMN a d'ailleurs révélé que la jonction III-IV-V forme une charnière dynamique qui permet à SLV d'explorer un certain espace conformationnel. L'interaction *kissing-loop* I/V est vraisemblablement compatible avec toutes les conformations échantillonnées par SLV. De plus, il est probable que suite à la formation de l'interaction *kissing-loop* I/V, le caractère dynamique de la jonction III-IV-V permette de faciliter le positionnement de la boucle interne de SLI dans le domaine catalytique et son association avec la boucle A730 pour former le site actif (Chapitre 4). La jonction III-IV-V joue donc un rôle architectural dans l'organisation du domaine catalytique en plus de contribuer à la formation de l'interaction *kissing-loop* I/V et du site actif du ribozyme VS.

6.1.3. La jonction II-III-VI requiert la formation d'une interaction tertiaire critique pour se replier

Nous avons également déterminé par spectroscopie de RMN la structure de la jonction II-III-VI, soit le dernier sous-domaine du ribozyme VS dont la structure était manquante (Chapitre 5). Nos études RMN ont révélé que la jonction II-III-VI adopte une structure stable où les tiges II et III forment un empilement continu de paires de bases tandis que les tiges II et VI forment un angle aigu. Cet agencement des tiges est caractéristique des jonctions à trois voies de la famille C. La structure de la jonction II-III-VI dépend de la présence d'ions Mg^{2+} et d'une interaction tertiaire impliquant la formation d'une paire de bases non-canonique *cis Sugar edge/WC A-A* entre les renflements des tiges II et VI. En effet, l'absence d'ions Mg^{2+} ou la perturbation de cette interaction tertiaire engendre une altération importante de l'architecture de la jonction II-III-VI, observable par des changements significatifs des profils spectraux des protons imino en RMN. En présence d'ions Mg^{2+} , le cœur de la jonction II-III-VI est stabilisé par la formation de deux triplets de

bases et d'une paire de bases *cis* WC/WC A-G. De plus, le renflement AA de la tige VI possède les caractéristiques d'un motif *S-turn* qui exclut une adénine (A_{726}/A_{33}) hors de l'hélice et permet l'interaction avec la tige II.

La localisation des sites de liaison de métaux dans la jonction II-III-VI a été déterminée par l'approche de relaxation paramagnétique induite par le Mn^{2+} (Chapitre 3). Huit sites de liaison de métaux ont été identifiés dont deux se situent au cœur de la jonction tandis que trois sites se situent à proximité de l'interaction tertiaire entre les renflements des tiges II et VI. Ces données supportent le rôle architectural important des ions Mg^{2+} dans l'adoption de la structure tridimensionnelle de la jonction II-III-VI.

La structure RMN de la jonction II-III-VI est en accord avec toutes les données biochimiques disponibles sur le ribozyme VS [215, 218, 220, 222-224, 233, 246, 249, 290], exception faite du résidu A_{657} (A_{11}) dont la nucléobase est exposée au solvant mais pour lequel les données biochimiques suggèrent un rôle important pour l'activité catalytique. Des études supplémentaires seront nécessaires pour identifier le rôle de ce résidu. La structure RMN de la jonction II-III-VI est aussi en accord avec la plupart des données structurales du ribozyme VS provenant d'études de SAXS et de FRET [249, 250, 260]. La seule étude divergente, réalisée par FRET [249], propose un agencement entre les tiges différent de celui présenté dans cette thèse. Cette différence s'explique par l'absence de renflements dans les tiges II et VI dans l'étude de FRET qui, selon nos résultats, sont essentiels pour définir pour l'architecture de la jonction II-III-VI. Parmi les études qui sont en accord avec la structure RMN de la jonction II-III-VI, une étude de FRET à molécule simple (*single molecule* FRET) a démontré que la forme *cis* du ribozyme VS peut adopter deux états différents avant de former l'interaction *kissing-loop*; une dont la jonction II-III-VI est relâchée et une où elle est repliée [250]. Dans cette étude, l'adoption de l'état replié de la jonction II-III-VI est grandement favorisée par la présence d'ions Mg^{2+} et est requise pour la formation de l'interaction *kissing-loop* et donc, de l'état actif pré-clivage. Il semble donc que la structure RMN de la jonction II-III-VI corresponde à sa forme active où les tiges II et VI forment une plateforme qui facilite la reconnaissance de SLI ainsi que la formation du site actif du ribozyme VS.

6.1.4. Contribution à une meilleure compréhension des interactions ARN-métal

Dans cette thèse, deux approches spectroscopiques principales ont été utilisées pour localiser les ions Mg^{2+} dans les structures d'ARN (Chapitres 3 à 5). L'approche de la relaxation paramagnétique induite par le Mn^{2+} est une méthode rapide qui génère un grand nombre de contraintes structurales entre l'ARN et les ions (Chapitres 3 à 5). Cependant, cette approche est limitée puisqu'elle ne permet pas de discerner le mode de coordination des ions et ne permet généralement pas de sonder l'effet des ions sur les groupements phosphates. L'approche de la perturbation de déplacement chimique induite par le Cd^{2+} sur des ARN ayant des modifications phosphorothioates est quant à elle beaucoup plus informative sur le mode de liaison des ions sur des groupements phosphates spécifiques (Chapitre 3). De plus, elle permet de discerner si un ion métallique est coordonné de façon directe ou indirecte à un groupement phosphate. En combinaison, ces deux approches complémentaires permettent d'évaluer la contribution d'à peu près tous les groupements chimiques d'une molécule d'ARN et ce, en solution. Bien que l'approche de titrages au Cd^{2+} soit plus exigeante techniquement dû à la nécessité de produire des ARN ayant des modifications phosphorothioates stéréospécifiques, les titrages au Mn^{2+} ont l'avantage de pouvoir être réalisés rapidement et d'être hautement informatifs. De plus, la localisation des ions dans une structure RMN permet de raffiner les structures pour obtenir une plus grande précision (meilleur rmsd), ce qui présente un avantage expérimental considérable (Section 6.1.5).

Toutes les structures obtenues et présentées dans cette thèse ont été caractérisées par spectroscopie de RMN ce qui a permis de localiser les ions Mg^{2+} interagissant avec ces ARN. Au total, 19 sites de liaison de métaux ont été localisés dans la boucle A730 et les jonctions III-IV-V et II-III-VI par diverses approches RMN (Chapitres 3 à 5). Ces données ont une valeur considérable puisqu'elles permettent de mieux comprendre les interactions ARN-métal, mais peuvent aussi être utilisées pour introduire des sites de liaison de métaux spécifiques dans certains ARN [292]. La majorité des interactions ARN-métal connues à ce jour ont été précédemment identifiées à partir de structures cristallographiques d'ARN puisque la majorité des structures RMN n'incluent pas la localisation d'ions divalents. Bien

que les données cristallographiques soient informatives sur les tendances de l'ARN à former certains sites de liaisons de métaux, la formation des cristaux (*crystal packing*) peut favoriser ou défavoriser la liaison des cations à certains sites. De plus, les structures cristallographiques nécessitent une résolution suffisamment bonne afin de pouvoir évaluer le mode de coordination d'un métal à un site donné. En effet, si la résolution ne permet pas de positionner les molécules d'eau autour d'un ion métallique, il est alors à peu près impossible de savoir si le mode de coordination est direct (*inner-sphere*) ou indirect (*outer-sphere*). Les approches RMN utilisées pour localiser les ions Mg^{2+} en solution présentent donc des avantages particuliers et auraient avantage à être plus utilisées.

6.1.5. Implémentation d'approches RMN pour l'étude d'ARN de grande taille

Les études structurales d'ARN par spectroscopie RMN sont importantes et comptent pour près de la moitié des structures disponibles dans la base de données des structures de macromolécules (RCSB). Cependant, les études de RMN sont généralement limitées par la taille des molécules étudiées, due entre autre à l'empilement des signaux. Bien que la structure de la boucle A730 ne corresponde qu'à un ARN de 26 nucléotides, les jonctions III-IV-V et II-III-VI sont quant à elles formées de 47 et 62 nucléotides, respectivement, et ont représenté des défis particuliers en raison de leur taille.

Pour rendre possible la détermination de la structure de la jonction III-IV-V, une approche de marquage isotopique (^{15}N) par segment a été implémentée, permettant d'attribuer les signaux des protons imino des uridines du cœur de la jonction (Chapitre 4). Des approches similaires de marquage par segment sont typiquement utilisées pour introduire des nucléotides modifiés ou des fluorophores à des sites spécifiques [293-295], pour faciliter la résolution de phase en cristallographie [296, 297] et pour permettre les études RMN d'ARN de grande taille. Aussi, l'attribution des résonances de la jonction II-III-VI a été facilitée par l'utilisation de marquage isotopique par résidu ($^{13}C/^{15}N$; [275]). Cette approche relativement simple qui a déjà été utilisée [298], a l'avantage de réduire de façon générale l'empilement des signaux, facilitant ainsi l'attribution et l'obtention de contraintes de distance pouvant être incorporées dans le calcul de structure. La détermination des structures des jonctions III-IV-V et II-III-VI n'aurait pas été possible

sans l'utilisation de ces approches de marquages isotopiques spécifiques qui représentent des outils très importants dans l'étude de molécules d'ARN de grande taille par RMN.

Dans le cadre de cette thèse, nous avons aussi démontré que la localisation des ions Mg^{2+} permet de raffiner les structures RMN, ce qui représente un atout pour les ARN de grande taille. En effet, les structures RMN d'ARN où des contraintes ARN-métal ont été utilisées ont des valeurs de rmsd plus basses que les structures sans contraintes ARN-métal (Chapitres 2 à 5). Les contraintes ARN-métal peuvent permettre de mieux définir un élément de structure local lorsque celui-ci est moins bien défini par l'ensemble des autres contraintes. De plus, les contraintes ARN-métal peuvent également permettre de raffiner la structure globale d'un ARN lorsqu'un métal connecte différents éléments de structure. Par exemple, le site 6 de liaison de métaux de la jonction III-IV-V est localisé entre les tiges III et V et sa localisation a permis de mieux définir l'orientation de la tige V par rapport aux autres tiges (Chapitre 4). Bien que la localisation d'ions Mg^{2+} ne remplace pas l'utilisation de données de RDC qui offrent des contraintes à longue distance, la localisation des métaux est une approche complémentaire qui permet tout de même d'améliorer les ensembles de structures. Ainsi, la localisation de métaux est une stratégie efficace et généralement rapide qui contribue au raffinement de structure RMN d'ARN de grande taille.

6.2. Contribution aux bases de données et prédictions de structure d'ARN

Depuis plusieurs années, le nombre d'ARN fonctionnels découverts ne cesse d'augmenter. Ces ARN ont souvent des séquences de plus en plus longues, rendant difficile la détermination de la structure de tous ces ARN, sans compter que ces systèmes plus gros peuvent présenter de la dynamique intrinsèque. En effet, il devient difficile d'un point de vue technique de déterminer les structures à haute résolution de tous ces ARN par RMN et par cristallographie. Pour compenser ces limitations, différentes techniques bioinformatiques sont disponibles afin de prédire la structure tridimensionnelle d'ARN à partir de leur séquence et/ou de leur structure secondaire.

Deux approches principales existent. La première, parfois nommée approche *ab initio* ou approche sans gabarit, se base sur l'hypothèse thermodynamique statuant que la

structure native d'une molécule correspond à la structure ayant la plus basse énergie libre [299]. Les travaux initiaux basés sur cette approche ont démontré la capacité de repliement de la ribonucléase sous sa forme native suite à sa dénaturation complète [300]. Cette approche peut également être adaptée à la prédiction de structure d'ARN [301, 302]. Ainsi, les changements conformationnels d'un ARN sont simulés et les valeurs d'énergie associées sont enregistrées. La structure ayant l'énergie libre la plus basse est considérée comme la structure native de l'ARN.

La deuxième approche se base sur le principe que des macromolécules évolutivement reliées ont des structures secondaires et tertiaires conservées et que leur structure diverge moins rapidement que leur séquence [303]. Cette méthode est nommée approche avec gabarit puisqu'elle requiert la structure d'une molécule de référence. Des alignements de séquences et de la modélisation permettent d'obtenir la structure tridimensionnelle d'une molécule par comparaison avec la molécule de référence [301]. Dans le cas des ARN, le succès de cette approche dépend donc de la disponibilité d'une structure d'ARN de référence [304].

Finalement, les méthodes hybrides, nommées modélisation *de novo*, combinent les approches *ab initio* et avec gabarit [301, 305]. Ces méthodes tirent avantage des forces de chacune de ces approches. Parmi les approches hybrides, les outils de prédiction se basant sur l'assemblage de fragments d'ARN sont populaires et puissants (FARNA [306, 307] et MCSYM [308]). En effet, l'assemblage par fragments utilise des petits blocs d'ARN dont la structure est connue et modélise la structure tridimensionnelle des ARN à partir de l'agencement de ces blocs. Les modèles ayant les énergies libres les plus basses sont alors sélectionnés, résultant en la structure prédite d'un ARN donné.

Le grand nombre de structures obtenues dans cette thèse contribue donc de façon considérable à l'avancement et au raffinement de la prédiction de structure. En effet, les approches bioinformatiques avec gabarit et hybrides sont des outils très utiles mais nécessitent à la base des structures tridimensionnelles connues afin de pouvoir faire la prédiction de structure tertiaire la plus juste possible. Ainsi, le large nombre de motifs structuraux identifiés dans la boucle A730 et dans les jonctions III-IV-V et II-III-VI du ribozyme VS contribuera significativement à la diversification des structures disponibles dans les bases de données.

Similairement, la localisation des ions Mg^{2+} réalisée dans cette thèse contribue à une meilleure compréhension des interactions ARN-métal en plus de garnir les bases de données sur les sites de liaison de métaux. En effet, divers outils informatiques, tel que MINAS ou WebFR3D, sont disponibles pour comparer les sites de liaison de métaux identifiés dans les études structurales mais sont dépendants des bases de données [309-312]. Ainsi, les sites de liaison d'ions Mg^{2+} identifiés dans les boucles GAAA (Chapitre 3 et 4), dans les motifs *S-turn* (Chapitres 3 et 5), dans les boucles CUUG (Chapitres 4 et 5) et dans le cœur des jonctions à trois voies (Chapitres 4 et 5) contribuent de façon significative à la compréhension des interactions entre les métaux et l'ARN tout en nourrissant les bases de données structurales définissant les différents sites de liaison de métaux.

6.3. Nouvelle vision de la structure du ribozyme VS et construction d'un modèle RMN complet à haute résolution

L'objectif principal de cette thèse était de caractériser la structure des jonctions III-IV-V et II-III-VI ainsi que de la boucle A730. Il a été démontré précédemment que l'ensemble des structures RMN des sous-domaines isolés étudiés précédemment ainsi que ceux présentés dans cette thèse adoptent une structure qui est essentiellement en accord avec l'ensemble des données biochimiques et biophysiques obtenues sur la forme active du ribozyme VS. À partir de ces structures, il est maintenant possible de dresser un portrait plus juste de l'architecture du domaine catalytique pour mieux comprendre comment celui-ci arrive à reconnaître son substrat pour former le site actif.

6.3.1. L'architecture globale du ribozyme VS

L'ensemble des données structurales disponibles sur le ribozyme VS permet de mieux comprendre comment chacun des sous-domaines importants contribue à l'adoption de la structure tertiaire active. En effet, l'activité du ribozyme VS dépend de sa capacité à former deux interactions tertiaires clés : l'interaction *kissing-loop* I/V et le site actif. Les éléments structuraux du domaine catalytique doivent donc tous participer à l'adoption de la structure tridimensionnelle du ribozyme VS et ce avec suffisamment de flexibilité pour permettre la formation des interactions tertiaires. Il a été démontré que les différents sous-domaines étudiés précédemment (SLI, SLV et interaction *kissing-loop* I/V [227, 228, 230,

231, 253, 257]) ainsi que ceux présentés dans cette thèse sont pré-organisés de façon isolée, c'est-à-dire qu'ils forment des structures bien définies, suggérant une certaine modularité au niveau de l'organisation de la structure. On note entre autre que l'empilement coaxial de paires d'hélices dans les jonctions III-IV-V et II-III-VI observés dans les sous-domaines isolés fait en sorte que la charpente principale du domaine catalytique soit vraisemblablement formée de l'empilement coaxial des tiges II, III et IV dans le ribozyme complet (Figure 6.1). Il en résulte que les hélices V et VI devraient être orientées parallèlement à cette charpente principale. Dans cette architecture globale, la jonction III-IV-V favorise la formation de l'interaction *kissing-loop* I/V en positionnant SLV tandis que la jonction II-III-VI permet de positionner la boucle A730 de SLVI dans une orientation favorable pour la formation du site actif [250]. La présence d'ions Mg^{2+} permet de définir l'architecture des jonctions III-IV-V et II-III-VI, la formation de l'interaction *kissing-loop* I/V mais aussi d'organiser le site actif du ribozyme VS puisque la boucle A730 peut former son motif *S-turn* et interagir avec la boucle interne de SLI pour former le site actif du ribozyme VS (Chapitres 2, 3 et 5). Par comparaison avec les autres ribozymes nucléolytiques et selon les données biochimiques disponibles du ribozyme VS, nous avons proposé que cette interaction engendre un léger réarrangement conformationnel dans la boucle interne de SLI, lui permettant d'adopter une conformation étendue au site de clivage (Chapitre 2). La formation du site actif devrait permettre de positionner de façon idéale les résidus catalytiques, G₆₃₈ et A₇₅₆, pour mener au clivage ou à la ligation du lien phosphodiester entre les résidus G₆₂₀ et A₆₂₁ (Chapitre 2). Ainsi, les structures RMN obtenues dans le cadre de cette thèse offrent une base rationnelle permettant de mieux comprendre l'adoption de la structure active du ribozyme VS de *Neurospora*.

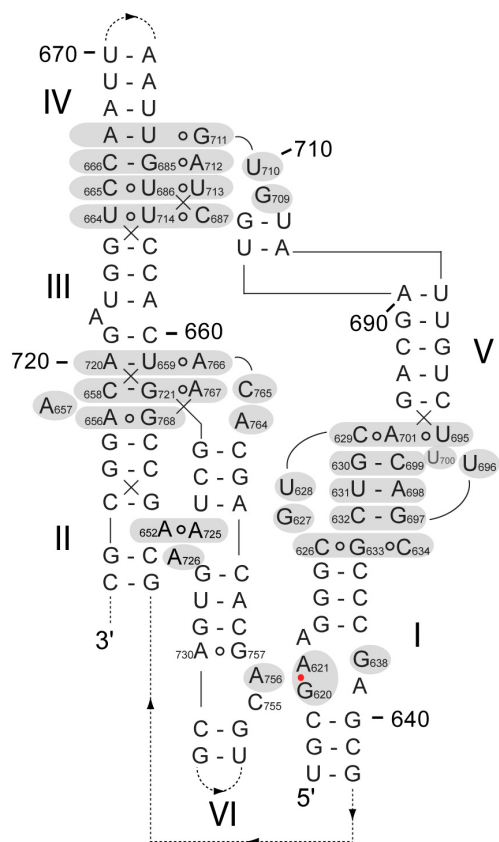


Figure 6.1. Structure secondaire révisée du ribozyme VS de *Neurospora*. Les régions ombragées représentent les jonctions III-IV-V et II-III-VI, l'interaction *kissing-loop* I/V, l'interaction tertiaire A₆₅₂-A₇₂₅ et la formation du site actif entre les boucles internes de SLI et A730. Les cercles ouverts représentent les paires de bases non-canoniques et les triplets de bases.

6.3.2. Éléments structuraux récurrents du ribozyme VS

Plusieurs éléments structuraux sont récurrents chez le ribozyme VS. Le premier élément structural grandement retrouvé est l'empilement de paires de bases qui favorise l'empilement coaxial des hélices. Une étude structurale précédente avait démontré que la formation de l'interaction *kissing-loop* I/V est entre autre stabilisée par la formation et l'empilement de paires de bases à la jonction I/V qui s'étend aux tiges I et V [227]. Les travaux présentés dans cette thèse ont démontré que l'empilement de bases est aussi fortement répandu dans le domaine catalytique du ribozyme VS. Par exemple, la paire de bases *cis* WC/WC A₇₃₀-G₇₅₇ de la boucle A730 est empilée avec les deux paires de bases adjacentes (G₇₂₉-C₇₅₈ et C₇₃₁-G₇₅₄; Chapitre 2) tandis que de l'empilement coaxial est observé entre les tiges III et IV de la jonction III-IV-V (Chapitre 4) et entre les tiges II et III

de la jonction II-III-VI (Chapitre 5). Comme dans d'autres structures d'ARN, l'empilement de paires de bases est donc une interaction stabilisatrice répandue dans l'ensemble du ribozyme VS (Figure 6.1).

Un autre élément structural important est la présence de plusieurs motifs de boucles terminales et de boucles internes récurrents dans plusieurs domaines : il s'agit des motifs *U-turn* et *S-turn*. Dans le ribozyme VS, le motif *U-turn* est retrouvé dans les boucles terminales des hélices I et V, permettant la formation de l'interaction *kissing-loop* I/V, mais est également retrouvé dans la jonction III-IV-V où il facilite la formation d'interactions tertiaires (Chapitre 4; [227]). De plus, un motif s'apparentant à un *U-turn* est aussi retrouvé dans le segment ACAA de la jonction II-III-VI, facilitant la formation d'interactions tertiaires dans cette jonction (Chapitre 5). Aussi, des motifs *S-turn* sont retrouvés dans la boucle A730 ainsi que dans le renflement A₇₂₅A₇₂₆ de SLVI. Chacun de ces motifs *S-turn* favorise la formation d'interaction tertiaires à longue distance, soit la formation du site actif ainsi que la formation de l'interaction entre les renflements des tiges II et VI (Chapitres 2 et 5). Ainsi, ces motifs *U-turn* et *S-turn* récurrents remplissent des rôles cruciaux dans la définition de la structure tridimensionnelle du ribozyme VS.

Le ribozyme VS requiert aussi la formation de plusieurs paires de bases non-canoniques et de triplets de bases qui sont importants pour son repliement (Figure 6.1). Par exemple, une paire de bases *cis* WC/WC G-A est retrouvée dans la boucle A730 et dans la jonction II-III-VI tandis que des triplets de bases sont retrouvés dans les jonctions III-IV-V et II-III-VI ainsi qu'à la jonction I/V de l'interaction *kissing-loop* (Chapitres 2, 4 et 5; [227]). Ces interactions favorisent l'empilement coaxial et l'adoption de structures compactes au cœur des jonctions.

Finalement, le quatrième élément structural récurrent du ribozyme VS est la stabilisation de sa structure tertiaire par la liaison d'ions Mg²⁺. Par exemple, un ion Mg²⁺ est spécifiquement coordonné à la boucle A730, stabilisant la formation du motif *S-turn* (Chapitre 3). De plus, deux ions Mg²⁺ se lient aux cœurs de chacune des jonctions III-IV-V et II-III-VI tandis que trois ions Mg²⁺ se trouvent à proximité de l'interaction tertiaire entre les renflements des tiges II et VI (Chapitres 4 et 5). La coordination de ces ions facilite donc l'adoption d'une structure compacte pour ces domaines.

6.3.3. Vers un modèle RMN complet à haute résolution du ribozyme VS

Ces nouvelles structures rendent possible la génération d'un modèle tridimensionnel complet à haute résolution du ribozyme VS de *Neurospora* basé sur les données de RMN. En effet, les structures RMN de tous les sous-domaines importants sont maintenant disponibles et les travaux effectués dans cette thèse ont d'ailleurs grandement contribué à cet accomplissement. La production d'un modèle tridimensionnel complet basé sur les structures RMN de domaines isolés est essentielle puisqu'actuellement, les seuls modèles complets disponibles sont des modèles à basse résolution provenant des données biochimiques, de FRET et de SAXS qui incorporent peu ou pas de données à haute résolution [224, 243, 260]. L'information provenant de ces modèles à basse résolution est limitée puisqu'ils ne révèlent pas les interactions formées dans le ribozyme VS d'un point de vue atomique. Il serait donc intéressant de produire un modèle tridimensionnel du ribozyme VS dans son état pré-catalytique (Figure 6.2). Afin de produire ce modèle, il serait possible d'utiliser une approche bioinformatique de modélisation hybride, qui permet de produire un modèle par l'assemblage de fragments d'ARN (Section 6.2). En terme de fragments, il serait possible d'utiliser la structure RMN de la boucle interne du substrat dans sa conformation *shifted* [253], la structure RMN de l'interaction *kissing-loop* I/V [227], la structure RMN de la boucle A730 (Chapitres 2 et 3) et les structures RMN des jonctions III-IV-V et II-III-VI (Chapitres 4 et 5; Figure 6.2). Afin de produire un modèle complet cohérent, il serait même possible de modéliser l'état de transition en faisant subir à la boucle interne de SLI un léger changement conformationnel lui permettant d'adopter la configuration étendue caractéristique des sites actifs des petits ribozymes nucléolytiques [289]. Pour ce faire, la modélisation du site actif présenté dans cette thèse pourrait être incorporée dans le modèle (Chapitre 2).

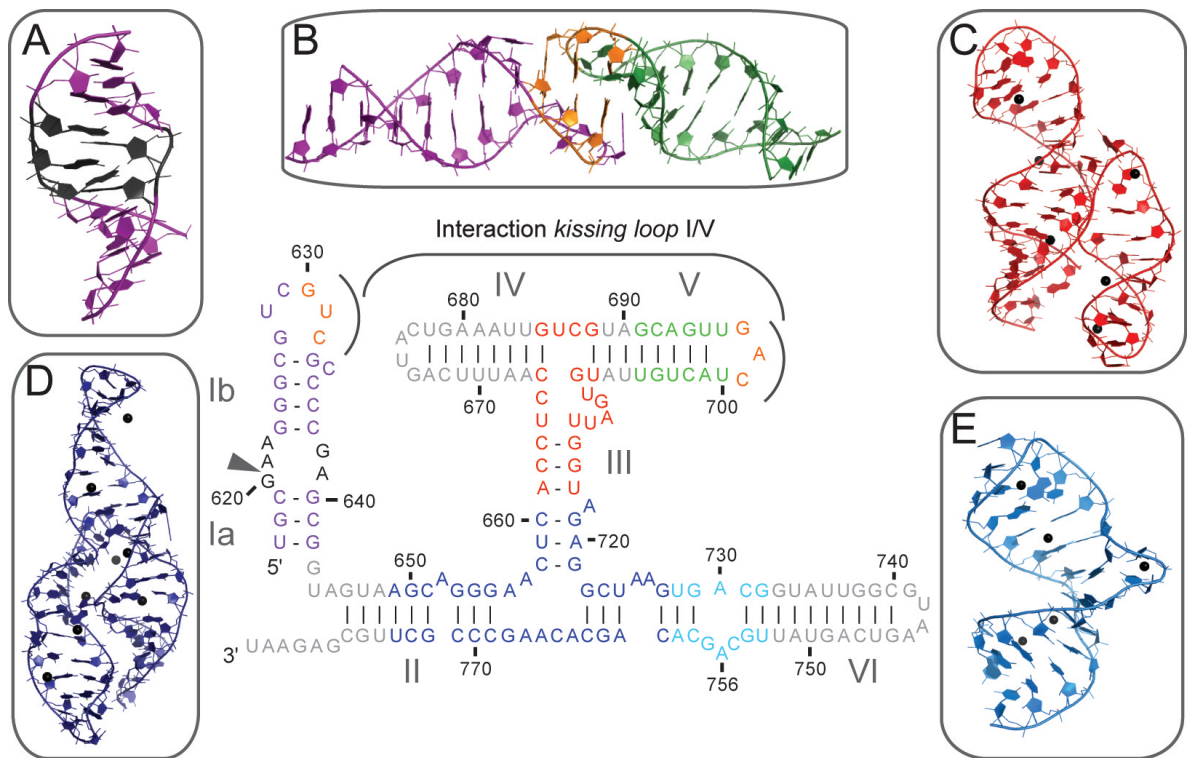


Figure 6.2. Structures RMN pouvant être utilisées pour modéliser la structure complète à haute résolution du ribozyme VS dans une conformation pré-catalytique. Structures RMN (A) de la conformation active de la boucle interne du substrat (PDB : 1OW9; [253]), (B) de l'interaction *kissing-loop* I/V (PDB : 2MI0; [227]), (C) de la jonction III-IV-V en présence d'ions Mg^{2+} (PDB : 2MTK; [313]), (D) de la jonction II-III-VI en présence d'ions Mg^{2+} (Chapitre 5) et (E) de la boucle A730 en présence d'ions Mg^{2+} (PDB : 2MIS; [314]). Les couleurs utilisées pour représenter les structures RMN correspondent aux sous-domaines présentés dans la structure secondaire, au centre.

Il serait aussi avantageux d'obtenir des contraintes expérimentales supplémentaires sur un ribozyme VS complet pour augmenter la qualité d'un tel modèle. En combinant des données supplémentaires aux structures RMN des sous-domaines isolés, il serait possible d'obtenir une structure RMN à haute résolution du ribozyme VS complet. Des données biochimiques pourraient être utilisées, par exemple en identifiant le rôle du A_{657} qui est présent dans la jonction II-III-VI et qui est susceptible de former une interaction tertiaire avec un autre sous-domaine du ribozyme VS (Chapitre 5). Pour ce faire, il serait possible de remplacer le A_{657} par un nucléotide photo-activable afin de créer un *cross-link* et d'identifier un partenaire d'interaction [248]. L'identification d'une telle interaction tertiaire supplémentaire pourrait ensuite être incorporée pour raffiner le modèle du ribozyme VS. Aussi, des contraintes supplémentaires pourraient provenir d'expériences de

RMN réalisées sur un ribozyme VS complet. J'ai d'ailleurs déjà initiées des expériences dans le laboratoire avec un ribozyme complet et il serait possible d'en dériver des contraintes de type RDC, qui sont des contraintes à longue distance provenant de données RMN et qui permettent d'orienter les sous-domaines les uns par rapport aux autres. Il serait aussi possible de réaliser des expériences de SAXS qui permettraient d'obtenir des enveloppes moléculaires pour raffiner la structure. L'importance d'une telle structure est considérable puisqu'elle permettrait de mieux comprendre l'organisation et le fonctionnement du ribozyme VS et constituerait une base rationnelle sur laquelle pourrait s'appuyer plusieurs autres études.

De plus, il serait possible de poursuivre les études structurales du ribozyme VS complet en utilisant une approche de dynamique moléculaire. En effet, cette approche bioinformatique a l'avantage de permettre de faire évoluer un système dans un champ de forces donné tout en observant l'évolution de régions potentiellement dynamiques. Par exemple, nous avons démontré que la jonction III-IV-V est une charnière dynamique qui favorise la formation de l'interaction *kissing-loop* I/V ainsi que la formation du site actif (Chapitre 4). Il avait aussi été précédemment démontré que les tiges de SLI et SLV peuvent adopter une gamme assez large d'angles interhélicaux lorsqu'ils forment l'interaction *kissing-loop* [227] et que la jonction II-III-VI pourrait adopter une conformation relâchée et une conformation repliée [250]. Il serait donc intéressant d'évaluer, par dynamique moléculaire, l'amplitude de mouvements potentiels de ces domaines ainsi que la participation de ces régions flexibles dans la formation de l'interaction *kissing-loop* I/V et la formation du site actif. Cette approche de dynamique moléculaire permettrait ainsi d'élucider les principaux mouvements potentiels associés à l'activité du ribozyme VS de *Neurospora*.

6.3.4. Obtention de données expérimentales sur le site actif

Grâce aux structures RMN de la boucle interne de SLI et de la boucle A730, nous avons pu généré un modèle du site actif du ribozyme VS (Chapitre 2). Bien que ce modèle ait été produit par homologie avec le site actif dans l'état de transition du ribozyme *hairpin*, il reste qu'il ne s'agit que d'un modèle. Il serait donc intéressant de pouvoir étudier, d'un point de vue expérimental, la formation du site actif du ribozyme VS. En effet, aucune

donnée n'est présentement disponible sur la réorganisation nécessaire au niveau de la boucle interne de SLI pour obtenir la conformation étendue (*splayed*) dans le site actif du ribozyme VS. Cette conformation, facilitant l'adoption d'une configuration *in-line* qui favorise la réaction enzymatique, est proposée, mais n'est pas confirmée expérimentalement. Il serait donc nécessaire de caractériser cet état, autant d'un point de vue biochimique que structural. Pour ce faire, il serait nécessaire de piéger la conformation étendue. Ceci pourrait être envisageable par le biais d'expériences de *cross-linking* sur le ribozyme VS pour essayer de capturer cet état. En effet, de telles expériences de *cross-linking* ont précédemment été utilisées pour identifier le site actif du ribozyme VS. La présence d'un *cross-link* pourrait permettre de stabiliser le site actif afin de pouvoir favoriser sa caractérisation [248]. De plus, notre meilleure compréhension de la structure du ribozyme VS pourrait permettre de stabiliser le ribozyme VS (par exemple en favorisant l'état replié de la jonction II-III-VI) afin de faire des essais cristallographiques sur le ribozyme VS. Si de tels essais étaient concluants, il serait possible d'insérer un groupement vanadate au site de clivage, ce qui permettrait de mimer l'état de transition du ribozyme VS pour ainsi capturer l'état étendu. En effet, un groupement vanadate a déjà été utilisé sur le ribozyme *hairpin* et il serait possible d'utiliser une approche similaire pour le ribozyme VS [203]. De telles données biochimiques et structurales de l'état étendu du site de clivage permettraient de raffiner le modèle existant du site actif ainsi que le modèle complet du ribozyme VS tout en permettant de mieux comprendre les détails mécanistiques menant à la catalyse par le ribozyme VS.

6.3.5. Localisation des ions Mg^{2+} dans le ribozyme VS

Les résultats présentés dans cette thèse démontrent que les ions Mg^{2+} jouent un rôle crucial dans l'adoption de la structure de plusieurs sous-domaines du ribozyme VS (Chapitres 2 à 5). La localisation de cations divalents avait été précédemment accomplie pour la boucle terminale de SLV tandis que les ions Mg^{2+} ont été positionnés dans les structures RMN de la boucle A730 et des jonctions III-IV-V et II-III-VI (Chapitres 3 à 5). Cependant, les ions Mg^{2+} n'ont pas été localisés dans la structure de l'interaction *kissing-loop* I/V ni dans celle de la boucle interne de SLI et il serait intéressant de faire cette localisation afin de pouvoir raffiner un modèle complet du ribozyme VS. De plus, la

localisation des métaux divalents dans l'interaction *kissing-loop* I/V est importante puisque des expériences d'interférence aux phosphorothioates avaient démontré que cette interaction nécessite la présence d'ions Mg^{2+} pour se former et que certains groupements phosphates pourraient agir comme des ligands potentiels [233]. Il avait aussi été proposé que la boucle interne de SLI sous sa forme active puisse former un site de liaison de métaux [253]. Cependant, le rôle et la localisation de ces métaux lors de la formation du site actif ne sont pas connus. Sachant que le pK_a des nucléobases participant au mécanisme enzymatique varie selon les métaux en solution [254] et sachant que les ions métalliques peuvent jouer un rôle dans la stabilisation des intermédiaires réactionnels, il serait intéressant de pouvoir localiser les ions Mg^{2+} dans le site actif du ribozyme VS. Évidemment, les approches de RMN de localisation de métaux utilisées dans le cadre de cette thèse (Section 6.1.4) pourraient être utilisées pour localiser les ions Mg^{2+} dans ces structures.

6.4. Ingénierie du ribozyme VS

L'obtention d'un modèle complet à haute résolution du ribozyme VS devrait pouvoir faciliter l'ingénierie rationnelle du ribozyme VS. En effet, la disponibilité des structures des différents sous-domaines ou d'un modèle du ribozyme VS complet permet de plus facilement prédire l'effet de la substitution d'un résidu ou d'une partie d'une séquence par une autre dans un sous-domaine donné. Plusieurs autres ribozymes nucléolytiques ont d'ailleurs été la cible d'études d'ingénierie rationnelle fructueuses [315-317]. Une des caractéristiques particulières du ribozyme VS est qu'il reconnaît un substrat qui adopte une forme repliée en tige-boucle. Comme la majorité des ARN adoptent également des formes repliées dans les cellules, cette caractéristique confère un attrait intéressant au ribozyme VS pour des études d'ingénierie en permettant de cibler des ARN spécifiques qui adoptent des formes structurées. Au niveau du ribozyme VS, il n'y a eu à ce jour qu'une seule étude d'ingénierie d'accomplie [235]. Cette étude a permis de démontrer qu'une compensation entre la longueur des tiges Ib et V est compatible avec l'activité. Le nombre restreint d'études d'ingénierie du ribozyme VS provient possiblement de l'absence de données structurales sur le domaine catalytique et les études présentées dans cette thèse devraient permettre de faciliter de telles études d'ingénierie.

Des études d'ingénierie pourraient être initiées afin de modifier les éléments structuraux retrouvés à travers le ribozyme. Ce genre d'études serait intéressant pour mieux comprendre la relation structure-fonction de l'ARN. Par exemple, il serait donc possible de modifier le motif *U-turn* retrouvé dans la jonction III-IV-V par un autre motif ou par une séquence aléatoire afin de voir comment ces changements affectent l'activité du ribozyme.

Dans l'optique où des études d'ingénierie du ribozyme VS viseraient à lui faire reconnaître et cliver un ARN différent de son substrat naturel, tel qu'une cible thérapeutique comme un ARN viral ou un ARNmi, il est important de pouvoir évaluer le potentiel de formation des interactions tertiaires clés du ribozyme VS en trois dimensions. Par exemple, l'ARN TAR du VIH-1 forme une interaction *kissing-loop* qui permet l'initiation de la dimérisation du génome viral [109]. Il serait donc intéressant de modifier le ribozyme VS pour qu'il reconnaisse l'ARN TAR par une telle interaction *kissing-loop*, et que par la suite, d'autres études d'ingénierie permettent de modifier le site actif du ribozyme VS afin de supporter le clivage de l'ARN TAR. Par exemple, l'ARN TAR contient un renflement qui favorise son interaction avec la protéine Tat et ce renflement pourrait être ciblé pour le clivage [318]. Évidemment, cet ARN représente un exemple mais une stratégie similaire pourrait être considérée sur d'autres ARN cibles. Ainsi, les structures RMN des domaines isolés contribuent à faciliter l'ingénierie rationnelle du ribozyme VS.

Une autre approche d'ingénierie est l'utilisation de la sélection *in vitro* (*systematic evolution of ligands by exponential enrichment*) ou SELEX. Cette approche est pratique puisqu'elle permet de modifier aléatoirement certaines séquences d'un système [319]. Cette approche pourrait être utilisée sur le ribozyme VS afin de rendre le ribozyme actif dans des conditions sous-optimales (changement de pH ou faible concentration d'ions divalents) ou encore pour lui permettre de cliver un ARN cible différent du substrat naturel. La connaissance des éléments structuraux importants pour la structure du ribozyme peut faciliter la décision de modifier ou non certaines régions du ribozyme. La caractérisation des ribozymes obtenues par SELEX pourrait aussi mener à une caractérisation biochimique et structurale de ces nouveaux ribozymes qui nécessiterait la comparaison avec les structures des sous-domaines naturels du ribozyme VS afin de comprendre l'activité des nouvelles séquences. Ainsi, la disponibilité des structures RMN de ces domaines est essentielle pour l'établissement de bases comparatives cohérentes en ingénierie moléculaire.

7. Références

1. Giege, R., et al., *Structure of transfer RNAs: similarity and variability*. Wiley Interdiscip. Rev. RNA, 2012. **3**(1): p. 37-61.
2. Noller, H.F., *RNA structure: reading the ribosome*. Science, 2005. **309**(5740): p. 1508-14.
3. Lerner, M.R., et al., *Two novel classes of small ribonucleoproteins detected by antibodies associated with lupus erythematosus*. Science, 1981. **211**(4480): p. 400-2.
4. Feng, J., et al., *The RNA component of human telomerase*. Science, 1995. **269**(5228): p. 1236-41.
5. Winter, J., et al., *Many roads to maturity: microRNA biogenesis pathways and their regulation*. Nat. Cell. Biol., 2009. **11**(3): p. 228-34.
6. Agrawal, N., et al., *RNA interference: biology, mechanism, and applications*. Microbiol. Mol. Biol. Rev., 2003. **67**(4): p. 657-85.
7. Mani, S.R. and C.E. Juliano, *Untangling the web: the diverse functions of the PIWI/piRNA pathway*. Mol. Reprod. Dev., 2013. **80**(8): p. 632-64.
8. Sorek, R., V. Kunin, and P. Hugenholtz, *CRISPR--a widespread system that provides acquired resistance against phages in bacteria and archaea*. Nat. Rev. Microbiol., 2008. **6**(3): p. 181-6.
9. Winkler, W.C., *Riboswitches and the role of noncoding RNAs in bacterial metabolic control*. Curr. Opin. Chem. Biol., 2005. **9**(6): p. 594-602.
10. Maxwell, E.S. and M.J. Fournier, *The small nucleolar RNAs*. Annu. Rev. Biochem., 1995. **64**: p. 897-934.
11. Valadkhan, S., *snRNAs as the catalysts of pre-mRNA splicing*. Curr. Opin. Chem. Biol., 2005. **9**(6): p. 603-8.
12. Mulhbachter, J., P. St-Pierre, and D.A. Lafontaine, *Therapeutic applications of ribozymes and riboswitches*. Curr. Opin. Pharmacol., 2010. **10**(5): p. 551-6.
13. Han, J.Q., et al., *A phylogenetically conserved RNA structure in the poliovirus open reading frame inhibits the antiviral endoribonuclease RNase L*. J. Virol., 2007. **81**(11): p. 5561-72.

14. Lin, K.C., H.L. Chang, and R.Y. Chang, *Accumulation of a 3'-terminal genome fragment in Japanese encephalitis virus-infected mammalian and mosquito cells*. J. Virol., 2004. **78**(10): p. 5133-8.
15. Kruger, K., et al., *Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena*. Cell, 1982. **31**: p. 147-157.
16. Guerrier-Takada, C., et al., *The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme*. Cell, 1983. **35**: p. 849-857.
17. Crick, F.H., *The origin of the genetic code*. J. Mol. Biol., 1968. **38**(3): p. 367-79.
18. Orgel, L.E., *Evolution of the genetic apparatus*. J. Mol. Biol., 1968. **38**(3): p. 381-93.
19. Saenger, W., *Principles of Nucleic Acid Structure*. Springer Advanced Texts in Chemistry, ed. C.R. Cantor. 1984, New York: Springer-Verlag. 556.
20. Batey, R.T., R.P. Rambo, and J.A. Doudna, *Tertiary Motifs in RNA Structure and Folding*. Angew. Chem. Int. Ed. Engl., 1999. **38**(16): p. 2326-2343.
21. Leontis, N.B., J. Stombaugh, and E. Westhof, *The non-Watson-Crick base pairs and their associated isostericity matrices*. Nucleic Acids Res., 2002. **30**(16): p. 3497-531.
22. Leontis, N.B. and E. Westhof, *Geometric nomenclature and classification of RNA base pairs*. RNA, 2001. **7**(4): p. 499-512.
23. Rosenberg, J.M., et al., *Double helix at atomic resolution*. Nature, 1973. **243**(5403): p. 150-4.
24. Day, R.O., et al., *A crystalline fragment of the double helix: the structure of the dinucleoside phosphate guanylyl-3',5'-cytidine*. Proc. Natl. Acad. Sci. U. S. A., 1973. **70**(3): p. 849-53.
25. Heus, H.A. and A. Pardi, *Structural features that give rise to the unusual stability of RNA hairpins containing GNRA tetraloops*. Science, 1991. **253**(5016): p. 191-194.
26. Jucker, F.M., et al., *A network of heterogeneous hydrogen bonds in GNRA tetraloops*. J. Mol. Biol., 1996. **264**(5): p. 968-80.
27. Schroeder, S.J., M.E. Burkard, and D.H. Turner, *The energetics of small internal loops in RNA*. Biopolymers, 1999. **52**(4): p. 157-67.
28. Elgavish, T., et al., *AAAG@helix.ends: A:A and A:G base-pairs at the ends of 16 S and 23 S rRNA helices*. J. Mol. Biol., 2001. **310**(4): p. 735-53.

29. Moore, P.B., *Structural motifs in RNA*. Annu. Rev. Biochem., 1999. **68**: p. 287-300.
30. Woese, C.R., S. Winker, and R.R. Gutell, *Architecture of ribosomal RNA: constraints on the sequence of "tetra-loops"*. Proc. Natl. Acad. Sci. U. S. A., 1990. **87**(21): p. 8467-71.
31. Antao, V.P., S.Y. Lai, and I. Tinoco, Jr., *A thermodynamic study of unusually stable RNA and DNA hairpins*. Nucleic Acids Res., 1991. **19**(21): p. 5901-5.
32. Quigley, G.J. and A. Rich, *Structural domains of transfer RNA molecules*. Science, 1976. **194**: p. 796-806.
33. von Ahsen, U., et al., *Identification of 2'-hydroxyl groups required for interaction of a tRNA anticodon stem-loop region with the ribosome*. RNA, 1997. **3**(1): p. 49-56.
34. Lambert, D., J.E. Heckman, and J.M. Burke, *Three conserved guanosines approach the reaction site in native and minimal hammerhead ribozymes*. Biochemistry, 2006. **45**(23): p. 7140-7.
35. Silvian, L.F., J. Wang, and T.A. Steitz, *Insights into editing from an ile-tRNA synthetase structure with tRNA^{ile} and mupirocin*. Science, 1999. **292**: p. 897-902.
36. Nozinovic, S., et al., *High-resolution NMR structure of an RNA model system: the 14-mer cUUCGg tetraloop hairpin RNA*. Nucleic Acids Res., 2010. **38**(2): p. 683-94.
37. Jucker, F.M. and A. Pardi, *Solution structure of the CUUG hairpin loop: a novel RNA tetraloop motif*. Biochemistry, 1995. **34**(44): p. 14416-27.
38. Jucker, F.M. and A. Pardi, *GNRA tetraloops make a U-turn*. RNA, 1995. **1**(2): p. 219-22.
39. Tuerk, C., et al., *CUUCGG hairpins: extraordinarily stable RNA secondary structures associated with various biochemical processes*. Proc. Natl. Acad. Sci. U. S. A., 1988. **85**(5): p. 1364-8.
40. Varani, G., C. Cheong, and I. Tinoco, Jr., *Structure of an unusually stable RNA hairpin*. Biochemistry, 1991. **30**(13): p. 3280-9.
41. Allain, F.H. and G. Varani, *Structure of the P1 helix from group I self-splicing introns*. J. Mol. Biol., 1995. **250**(3): p. 333-53.
42. Shen, L.X., Z. Cai, and I. Tinoco, Jr., *RNA structure at high resolution*. FASEB J., 1995. **9**(11): p. 1023-33.

43. Cate, J.H., et al., *RNA tertiary structure mediation by adenosine platforms*. *Science*, 1996. **273**(5282): p. 1696-9.
44. Lilley, D.M., *The K-turn motif in riboswitches and other RNA species*. *Biochim. Biophys. Acta*, 2014.
45. Puglisi, J.D., et al., *Conformation of the TAR RNA-arginine complex by NMR spectroscopy*. *Science*, 1992. **257**(5066): p. 76-80.
46. GuhaThakurta, D. and D.E. Draper, *Contributions of basic residues to ribosomal protein L11 recognition of RNA*. *J. Mol. Biol.*, 2000. **295**(3): p. 569-80.
47. Wimberly, B., G. Varani, and I. Tinoco, Jr., *The conformation of loop E of eukaryotic 5S ribosomal RNA*. *Biochemistry*, 1993. **32**(4): p. 1078-87.
48. Correll, C.C., et al., *Crystal structure of the ribosomal RNA domain essential for binding elongation factors*. *Proc. Natl. Acad. Sci. U. S. A.*, 1998. **95**(23): p. 13436-41.
49. Butcher, S.E., F.H. Allain, and J. Feigon, *Solution structure of the loop B domain from the hairpin ribozyme*. *Nat. Struct. Biol.*, 1999. **6**(3): p. 212-6.
50. Alam, S., et al., *Conformational heterogeneity at position U37 of an all-RNA hairpin ribozyme with implications for metal binding and the catalytic structure of the S-turn*. *Biochemistry*, 2005. **44**(44): p. 14396-408.
51. Szewczak, A.A., et al., *The conformation of the sarcin/ricin loop from 28S ribosomal RNA*. *Proc. Natl. Acad. Sci. U. S. A.*, 1993. **90**(20): p. 9581-5.
52. Leontis, N.B. and E. Westhof, *A common motif organizes the structure of multi-helix loops in 16 S and 23 S ribosomal RNAs*. *J. Mol. Biol.*, 1998. **283**(3): p. 571-83.
53. Zimmermann, G.R., et al., *Molecular interactions and metal binding in the theophylline-binding core of an RNA aptamer*. *RNA*, 2000. **6**: p. 659-667.
54. Allain, F.H., et al., *Solution structure of the two N-terminal RNA-binding domains of nucleolin and NMR study of the interaction with its RNA target*. *J. Mol. Biol.*, 2000. **303**(2): p. 227-41.
55. Tereshko, V., E. Skripkin, and D.J. Patel, *Encapsulating streptomycin within a small 40-mer RNA*. *Chem. Biol.*, 2003. **10**(2): p. 175-87.
56. Lilley, D.M., et al., *A nomenclature of junctions and branchpoints in nucleic acids*. *Nucleic Acids Res.*, 1995. **23**(17): p. 3363-4.

57. Leontis, N.B., W. Kwok, and J.S. Newman, *Stability and structure of three-way DNA junctions containing unpaired nucleotides*. *Nucleic Acids Res.*, 1991. **19**(4): p. 759-66.
58. Woodson, S.A. and N.B. Leontis, *Structure and dynamics of ribosomal RNA*. *Curr. Opin. Struct. Biol.*, 1998. **8**(3): p. 294-300.
59. Batey, R.T., S.D. Gilbert, and R.K. Montange, *Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine*. *Nature*, 2004. **432**(7015): p. 411-5.
60. Serganov, A., et al., *Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs*. *Chem. Biol.*, 2004. **11**(12): p. 1729-41.
61. Serganov, A., et al., *Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch*. *Nature*, 2006. **441**(7097): p. 1167-71.
62. Thore, S., M. Leibundgut, and N. Ban, *Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand*. *Science*, 2006. **312**(5777): p. 1208-11.
63. Tanner, M.A. and T.R. Cech, *Joining the two domains of a group I ribozyme to form the catalytic core*. *Science*, 1997. **275**(5301): p. 847-9.
64. Krasilnikov, A.S., et al., *Crystal structure of the specificity domain of ribonuclease P*. *Nature*, 2003. **421**(6924): p. 760-4.
65. De la Pena, M., S. Gago, and R. Flores, *Peripheral regions of natural hammerhead ribozymes greatly increase their self-cleavage activity*. *EMBO J.*, 2003. **22**(20): p. 5561-70.
66. Wilson, T.J., et al., *Folding and catalysis of the hairpin ribozyme*. *Biochem. Soc. Trans.*, 2005. **33**(Pt 3): p. 461-5.
67. Klein, D.J. and A.R. Ferre-D'Amare, *Structural basis of glmS ribozyme activation by glucosamine-6-phosphate*. *Science*, 2006. **313**(5794): p. 1752-6.
68. Nehdi, A., et al., *A novel structural rearrangement of hepatitis delta virus antigenomic ribozyme*. *Nucleic Acids Res.*, 2007. **35**(20): p. 6820-31.
69. Toor, N., et al., *Crystal structure of a self-spliced group II intron*. *Science*, 2008. **320**(5872): p. 77-82.

70. Roth, A., et al., *A widespread self-cleaving ribozyme class is revealed by bioinformatics*. Nat. Chem. Biol., 2014. **10**(1): p. 56-60.
71. Conn, G.L., et al., *Crystal structure of a conserved ribosomal protein-RNA complex*. Science, 1999. **284**(5417): p. 1171-4.
72. Wimberly, B.T., et al., *A detailed view of a ribosomal active site: the structure of the L11-RNA complex*. Cell, 1999. **97**(4): p. 491-502.
73. Hainzl, T., S. Huang, and A.E. Sauer-Eriksson, *Structure of the SRP19 RNA complex and implications for signal recognition particle assembly*. Nature, 2002. **417**(6890): p. 767-71.
74. Laing, C., et al., *Predicting coaxial helical stacking in RNA junctions*. Nucleic Acids Res., 2012. **40**(2): p. 487-98.
75. Lescoute, A. and E. Westhof, *Topology of three-way junctions in folded RNAs*. RNA, 2006. **12**(1): p. 83-93.
76. de la Pena, M., D. Dufour, and J. Gallego, *Three-way RNA junctions with remote tertiary contacts: a recurrent and highly versatile fold*. RNA, 2009. **15**(11): p. 1949-64.
77. Kim, S.H., et al., *Three-dimensional tertiary structure of yeast phenylalanine transfer RNA*. Science, 1974. **185**(4149): p. 435-40.
78. Robertus, J.D., et al., *Structure of yeast phenylalanine tRNA at 3 Å resolution*. Nature, 1974. **250**(467): p. 546-51.
79. Jack, A., J.E. Ladner, and A. Klug, *Crystallographic refinement of yeast phenylalanine transfer RNA at 2-5Å resolution*. J. Mol. Biol., 1976. **108**(4): p. 619-49.
80. Pley, H.W., K.M. Flaherty, and D.B. McKay, *Three-dimensional structure of a hammerhead ribozyme*. Nature, 1994. **372**: p. 68-74.
81. Cate, J.H., et al., *Crystal structure of a group I ribozyme domain: principles of RNA packing*. Science, 1996. **273**: p. 1678-1685.
82. Ferre-D'Amare, A.R., K. Zhou, and J.A. Doudna, *Crystal structure of a hepatitis delta virus ribozyme*. Nature, 1998. **395**(6702): p. 567-74.
83. Jaeger, L., F. Michel, and E. Westhof, *Involvement of a GNRA tetraloop in long-range RNA tertiary interactions*. J. Mol. Biol., 1994. **236**(5): p. 1271-6.

84. Murphy, F.L. and T.R. Cech, *GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain*. J. Mol. Biol., 1994. **236**(1): p. 49-63.
85. Costa, M. and F. Michel, *Frequent use of the same tertiary motif by self-folding RNAs*. EMBO J., 1995. **14**(6): p. 1276-85.
86. Oubridge, C., et al., *Crystal structure of SRP19 in complex with the S domain of SRP RNA and its implication for the assembly of the signal recognition particle*. Mol. Cell., 2002. **9**(6): p. 1251-61.
87. Kuglstatter, A., C. Oubridge, and K. Nagai, *Induced structural changes of 7SL RNA during the assembly of human signal recognition particle*. Nat. Struct. Biol., 2002. **9**(10): p. 740-4.
88. Krasilnikov, A.S. and A. Mondragon, *On the occurrence of the T-loop RNA folding motif in large RNA molecules*. RNA, 2003. **9**(6): p. 640-643.
89. Klein, D.J., P.B. Moore, and T.A. Steitz, *The roles of ribosomal proteins in the structure assembly, and evolution of the large ribosomal subunit*. J. Mol. Biol., 2004. **340**(1): p. 141-77.
90. Martick, M. and W.G. Scott, *Tertiary contacts distant from the active site prime a ribozyme for catalysis*. Cell, 2006. **126**(2): p. 309-20.
91. Chi, Y.I., et al., *Capturing hammerhead ribozyme structures in action by modulating general base catalysis*. PLoS Biol., 2008. **6**(9): p. e234.
92. Blake, R.D., J. Massoulié, and J.R. Fresco, *Polynucleotides. 8. A spectral approach to the equilibria between polyriboadenylate and polyribouridylylate and their complexes*. J. Mol. Biol., 1967. **30**(2): p. 291-308.
93. Conrad, N.K., *The emerging role of triple helices in RNA biology*. Wiley Interdiscip. Rev. RNA, 2014. **5**(1): p. 15-29.
94. Brown, J.A., et al., *Structural insights into the stabilization of MALAT1 noncoding RNA by a bipartite triple helix*. Nat. Struct. Mol. Biol., 2014. **21**(7): p. 633-40.
95. Nissen, P., et al., *RNA tertiary interactions in the large ribosomal subunit: the A-minor motif*. Proc. Natl. Acad. Sci. U. S. A., 2001. **98**(9): p. 4899-903.
96. Tamura, M. and S.R. Holbrook, *Sequence and structural conservation in RNA ribose zippers*. J. Mol. Biol., 2002. **320**: p. 455-474.

97. Pley, H.W., K.M. Flaherty, and D.B. McKay, *Model for an RNA tertiary interaction from the structure of an intermolecular complex between a GAAA tetraloop and an RNA helix*. *Nature*, 1994. **372**(6501): p. 111-3.
98. Kolk, M.H., et al., *Structure of the 3'-hairpin of the TYMV pseudoknot: preformation in RNA folding*. *EMBO J.*, 1998. **17**(24): p. 7498-504.
99. Kolk, M.H., et al., *NMR structure of a classical pseudoknot: interplay of single- and double-stranded RNA*. *Science*, 1998. **280**(5362): p. 434-8.
100. Brierley, I., S. Pennell, and R.J. Gilbert, *Viral RNA pseudoknots: versatile motifs in gene expression and replication*. *Nat. Rev. Microbiol.*, 2007. **5**(8): p. 598-610.
101. Chapman, E.G., et al., *The structural basis of pathogenic subgenomic flavivirus RNA (sfRNA) production*. *Science*, 2014. **344**(6181): p. 307-10.
102. Lemay, J.F., et al., *Folding of the adenine riboswitch*. *Chem. Biol.*, 2006. **13**(8): p. 857-68.
103. Blouin, S. and D.A. Lafontaine, *A loop loop interaction and a K-turn motif located in the lysine aptamer domain are important for the riboswitch gene regulation control*. *RNA*, 2007. **13**(8): p. 1256-67.
104. Garst, A.D., et al., *Crystal structure of the lysine riboswitch regulatory mRNA element*. *J. Biol. Chem.*, 2008. **283**(33): p. 22347-51.
105. Serganov, A., L. Huang, and D.J. Patel, *Structural insights into amino acid binding and gene control by a lysine riboswitch*. *Nature*, 2008. **455**(7217): p. 1263-7.
106. Khvorova, A., et al., *Sequence elements outside the hammerhead ribozyme catalytic core enable intracellular activity*. *Nat. Struct. Biol.*, 2003. **10**(9): p. 708-12.
107. Dufour, D., et al., *Structure-function analysis of the ribozymes of chrysanthemum chlorotic mottle viroid: a loop-loop interaction motif conserved in most natural hammerheads*. *Nucleic Acids Res.*, 2009. **37**(2): p. 368-81.
108. Rastogi, T., et al., *A long-range pseudoknot is required for activity of the Neurospora VS ribozyme*. *EMBO J.*, 1996. **15**(11): p. 2820-2825.
109. Skripkin, E., et al., *Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro*. *Proc. Natl. Acad. Sci. U. S. A.*, 1994. **91**(11): p. 4945-9.
110. Fedor, M.J., *Structure and function of the hairpin ribozyme*. *J. Mol. Biol.*, 2000. **297**(2): p. 269-91.

111. Cochrane, J.C. and S.A. Strobel, *Catalytic strategies of self-cleaving ribozymes*. *Acc. Chem. Res.*, 2008. **41**(8): p. 1027-35.
112. Sumita, M., et al., *Intermolecular domain docking in the hairpin ribozyme: metal dependence, binding kinetics and catalysis*. *RNA Biol.*, 2013. **10**(3): p. 425-35.
113. Ennifar, E., et al., *Crystal structures of coaxially stacked kissing complexes of the HIV-1 RNA dimerization initiation site*. *Nat. Struct. Biol.*, 2001. **8**(12): p. 1064-8.
114. Feig, A.L. and O.C. Uhlenbeck, *The role of metal ions in RNA biochemistry*, In *The RNA world (Gesteland, R.F., Cech, T. R., and Atkins, J. F., Eds.) 2nd ed.* Cold Spring Harbor Laboratory Press, New York. 1999.
115. Hsiao, C., et al., *Complexes of nucleic acids with group I and II cations in Nucleic Acid-Metal Ion Interactions 2009*, RCS Publishing, Cambridge, UK.
116. Misra, V.K. and D.E. Draper, *On the role of magnesium ions in RNA stability*. *Biopolymers*, 1998. **48**(2-3): p. 113-35.
117. Lilley, D.M.J. and F. Eckstein, eds. *Ribozymes and RNA catalysis*. 2008, Royal Society of Chemistry: Cambridge.
118. Wu, Q., L. Huang, and Y. Zhang, *The structure and function of catalytic RNAs*. *Sci. China C. Life Sci.*, 2009. **52**(3): p. 232-44.
119. Weng, D.E., et al., *A phase I clinical trial of a ribozyme-based angiogenesis inhibitor targeting vascular endothelial growth factor receptor-1 for patients with refractory solid tumors*. *Mol. Cancer Ther.*, 2005. **4**(6): p. 948-55.
120. Li, M.J., et al., *Long-term inhibition of HIV-1 infection in primary hematopoietic cells by lentiviral vector delivery of a triple combination of anti-HIV shRNA, anti-CCR5 ribozyme, and a nucleolar-localizing TAR decoy*. *Mol. Ther.*, 2005. **12**(5): p. 900-9.
121. Belousoff, M.J., et al., *Ancient machinery embedded in the contemporary ribosome*. *Biochem. Soc. Trans.*, 2010. **38**(2): p. 422-7.
122. Kazantsev, A.V. and N.R. Pace, *Bacterial RNase P: a new view of an ancient enzyme*. *Nat. Rev. Microbiol.*, 2006. **4**(10): p. 729-40.
123. Mondragon, A., *Structural studies of RNase P*. *Annu. Rev. Biophys.*, 2013. **42**: p. 537-57.
124. Cech, T.R., *Self-splicing of group I introns*. *Annu. Rev. Biochem.*, 1990. **59**: p. 543-68.

125. Lilley, D.M., *Structure, folding and mechanisms of ribozymes*. Curr. Opin. Struct. Biol., 2005. **15**(3): p. 313-23.
126. Fedorova, O. and N. Zingler, *Group II introns: structure, folding and splicing mechanism*. Biol. Chem., 2007. **388**(7): p. 665-78.
127. Johansen, S. and V.M. Vogt, *An intron in the nuclear ribosomal DNA of Didymium iridis codes for a group I ribozyme and a novel ribozyme that cooperate in self-splicing*. Cell, 1994. **76**(4): p. 725-34.
128. Fica, S.M., et al., *RNA catalyses nuclear pre-mRNA splicing*. Nature, 2013. **503**(7475): p. 229-34.
129. Salehi-Ashtiani, K., et al., *A genomewide search for ribozymes reveals an HDV-like sequence in the human CPEB3 gene*. Science, 2006. **313**(5794): p. 1788-92.
130. Teixeira, A., et al., *Autocatalytic RNA cleavage in the human beta-globin pre-mRNA promotes transcription termination*. Nature, 2004. **432**(7016): p. 526-30.
131. Martick, M., et al., *A discontinuous hammerhead ribozyme embedded in a mammalian messenger RNA*. Nature, 2008. **454**(7206): p. 899-902.
132. Robart, A.R., et al., *Crystal structure of a eukaryotic group II intron lariat*. Nature, 2014. **514**(7521): p. 193-7.
133. Rupert, P.B. and A.R. Ferré-D'Amaré, *Crystal structure of a haripin ribozyme-inhibitor complex with implications for catalysis*. Nature, 2001. **410**: p. 780-786.
134. Saville, B.J. and R.A. Collins, *A site-specific self-cleavage reaction performed by a novel RNA in Neurospora mitochondria*. Cell, 1990. **61**: p. 685-696.
135. Lilley, D.M., *The Varkud satellite ribozyme*. RNA, 2004. **10**: p. 151-158.
136. Dorner, S., et al., *Molecular aspects of the ribosomal peptidyl transferase*. Biochem. Soc. Trans., 2002. **30**(Pt 6): p. 1131-6.
137. Bashan, A., et al., *Structural basis of the ribosomal machinery for peptide bond formation, translocation, and nascent chain progression*. Mol. Cell, 2003. **11**(1): p. 91-102.
138. Agmon, I., et al., *Symmetry at the active site of the ribosome: structural and functional implications*. Biol. Chem., 2005. **386**(9): p. 833-44.
139. Schmeing, T.M., et al., *Structural insights into the roles of water and the 2' hydroxyl of the P site tRNA in the peptidyl transferase reaction*. Mol. Cell, 2005. **20**(3): p. 437-48.

140. Davidovich, C., et al., *The evolving ribosome: from non-coded peptide bond formation to sophisticated translation machinery*. Res. Microbiol., 2009. **160**(7): p. 487-92.
141. Yonath, A., *Large facilities and the evolving ribosome, the cellular machine for genetic-code translation*. J. R. Soc. Interface, 2009. **6 Suppl 5**: p. S575-85.
142. Robertson, H.D., S. Altman, and J.D. Smith, *Purification and properties of a specific Escherichia coli ribonuclease which cleaves a tyrosine transfer ribonucleic acid precursor*. J. Biol. Chem., 1972. **247**(16): p. 5243-51.
143. Stark, B.C., et al., *Ribonuclease P: an enzyme with an essential RNA component*. Proc. Natl. Acad. Sci. U. S. A., 1978. **75**(8): p. 3717-21.
144. Altman, S., *Ribonuclease P*. Philos. Trans. R. Soc. Lond. B. Biol. Sci., 2011. **366**(1580): p. 2936-41.
145. Haugen, P., D.M. Simon, and D. Bhattacharya, *The natural history of group I introns*. Trends Genet., 2005. **21**(2): p. 111-9.
146. Matera, A.G. and Z. Wang, *A day in the life of the spliceosome*. Nat. Rev. Mol. Cell Biol., 2014. **15**(2): p. 108-21.
147. van der Veen, R., et al., *Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro*. Cell, 1986. **44**(2): p. 225-34.
148. Cech, T.R., *The chemistry of self-splicing RNA and RNA enzymes*. Science, 1987. **236**(4808): p. 1532-9.
149. Steitz, T.A. and J.A. Steitz, *A general two-metal-ion mechanism for catalytic RNA*. Proc. Natl. Acad. Sci. U. S. A., 1993. **90**(14): p. 6498-502.
150. Fedor, M.J., *The role of metal ions in RNA catalysis*. Current Opin. in Struct. Biol., 2002. **12**: p. 289-295.
151. Valadkhan, S., *The spliceosome: a ribozyme at heart?* Biol. Chem., 2007. **388**(7): p. 693-7.
152. Bagheri, S. and M. Kashani-Sabet, *Ribozymes in the age of molecular therapeutics*. Curr. Mol. Med., 2004. **4**(5): p. 489-506.
153. Johnson, I.M., *RNA as a drug target: recent patents on the catalytic activity of trans-splicing ribozymes derived from group I intron RNA*. Recent. Pat. DNA Gene Seq., 2010. **4**(1): p. 17-28.

154. Fiskaa, T. and A.B. Birgisdottir, *RNA reprogramming and repair based on trans-splicing group I ribozymes*. N. Biotechnol., 2010. **27**(3): p. 194-203.
155. Liu, Y., et al., *Crystal structure and mechanistic investigation of the twister ribozyme*. Nat. Chem. Biol., 2014. **10**(9): p. 739-44.
156. Eiler, D., J. Wang, and T.A. Steitz, *Structural basis for the fast self-cleavage reaction catalyzed by the twister ribozyme*. Proc. Natl. Acad. Sci. U. S. A., 2014. **111**(36): p. 13028-33.
157. Ren, A., et al., *In-line alignment and Mg(2+) coordination at the cleavage site of the env22 twister ribozyme*. Nat. Commun., 2014. **5**: p. 5534.
158. Serganov, A. and D.J. Patel, *Ribozymes, riboswitches and beyond: regulation of gene expression without proteins*. Nat. Rev. Genet., 2007. **8**(10): p. 776-90.
159. Fedor, M.J. and J.R. Williamson, *The catalytic diversity of RNAs*. Nat. Rev. Mol. Cell. Biol., 2005. **6**(5): p. 399-412.
160. de la Pena, M. and I. Garcia-Robles, *Ubiquitous presence of the hammerhead ribozyme motif along the tree of life*. RNA, 2010. **16**(10): p. 1943-50.
161. Lilley, D.M., *Mechanisms of RNA catalysis*. Philos. Trans. R. Soc. Lond. B. Biol. Sci., 2011. **366**(1580): p. 2910-7.
162. Findlay, D., et al., *The active site and mechanism of action of bovine pancreatic ribonuclease*. Nature, 1961. **190**: p. 781-84.
163. Moody, E.M., J.T. Lecomte, and P.C. Bevilacqua, *Linkage between proton binding and folding in RNA: a thermodynamic framework and its experimental application for investigating pKa shifting*. RNA, 2005. **11**(2): p. 157-72.
164. Shih, I.H. and M.D. Been, *Ribozyme cleavage of a 2,5-phosphodiester linkage: mechanism and a restricted divalent metal-ion requirement*. RNA, 1999. **5**(9): p. 1140-8.
165. Han, J. and J.M. Burke, *Model for general acid-base catalysis by the hammerhead ribozyme: pH-activity relationships of G8 and G12 variants at the putative active site*. Biochemistry, 2005. **44**(21): p. 7864-70.
166. van Tol, H., et al., *Two autolytic processing reactions of a satellite RNA proceed with inversion of configuration*. Nucleic Acids Res., 1990. **18**(8): p. 1971-5.

167. Slim, G. and M.J. Gait, *Configurationaly defined phosphorothioate-containing oligoribonucleotides in the study of the mechanism of cleavage of hammerhead ribozymes*. Nucleic Acids Res., 1991. **19**(6): p. 1183-8.
168. Milewski, S., *Glucosamine-6-phosphate synthase--the multi-facets enzyme*. Biochim. Biophys. Acta, 2002. **1597**(2): p. 173-92.
169. Winkler, W.C., et al., *Control of gene expression by a natural metabolite-responsive ribozyme*. Nature, 2004. **428**(6980): p. 281-6.
170. Barrick, J.E., et al., *New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control*. Proc. Natl. Acad. Sci. U. S. A., 2004. **101**(17): p. 6421-6.
171. McCarthy, T.J., et al., *Ligand requirements for glmS ribozyme self-cleavage*. Chem. Biol., 2005. **12**(11): p. 1221-6.
172. Roth, A., et al., *Characteristics of the glmS ribozyme suggest only structural roles for divalent metal ions*. RNA, 2006. **12**(4): p. 607-19.
173. Wilkinson, S.R. and M.D. Been, *A pseudoknot in the 3' non-core region of the glmS ribozyme enhances self-cleavage activity*. RNA, 2005. **11**(12): p. 1788-94.
174. Cochrane, J.C., S.V. Lipchok, and S.A. Strobel, *Structural investigation of the GlmS ribozyme bound to Its catalytic cofactor*. Chem. Biol., 2007. **14**(1): p. 97-105.
175. Blount, K., et al., *Development and application of a high-throughput assay for glmS riboswitch activators*. RNA Biol., 2006. **3**(2): p. 77-81.
176. Mayer, G. and M. Famulok, *High-throughput-compatible assay for glmS riboswitch metabolite dependence*. Chembiochem, 2006. **7**(4): p. 602-4.
177. Lai, M.M., *The molecular biology of hepatitis delta virus*. Annu. Rev. Biochem., 1995. **64**: p. 259-86.
178. Flores, R., et al., *Rolling-circle replication of viroids, viroid-like satellite RNAs and hepatitis delta virus: variations on a theme*. RNA Biol., 2011. **8**(2): p. 200-6.
179. Filipovska, J. and M.M. Konarska, *Specific HDV RNA-templated transcription by pol II in vitro*. RNA, 2000. **6**(1): p. 41-54.
180. Perrotta, A.T. and M.D. Been, *The self-cleaving domain from the genomic RNA of hepatitis delta virus: sequence requirements and the effects of denaturant*. Nucleic Acids Res., 1990. **18**(23): p. 6821-7.
181. Perrotta, A.T. and M.D. Been, *A pseudoknot-like structure required for efficient self-cleavage of hepatitis delta virus RNA*. Nature, 1991. **350**(6317): p. 434-6.

182. Chadalavada, D.M., et al., *A role for upstream RNA structure in facilitating the catalytic fold of the genomic hepatitis delta virus ribozyme*. J. Mol. Biol., 2000. **301**(2): p. 349-67.
183. Been, M.D. and G.S. Wickham, *Self-cleaving ribozymes of hepatitis delta virus RNA*. Eur. J. Biochem., 1997. **247**(3): p. 741-53.
184. Wadkins, T.S., et al., *A nested double pseudoknot is required for self-cleavage activity of both the genomic and antigenomic hepatitis delta virus ribozymes*. RNA, 1999. **5**(6): p. 720-7.
185. Thaplyal, P., et al., *Thio effects and an unconventional metal ion rescue in the genomic hepatitis delta virus ribozyme*. Biochemistry, 2013. **52**(37): p. 6499-514.
186. Bergeron, L.J., C. Reymond, and J.P. Perreault, *Functional characterization of the SOFA delta ribozyme*. RNA, 2005. **11**(12): p. 1858-68.
187. Prody, G.A., et al., *Autolytic processing of dimeric plant virus satellite RNA*. Science, 1986. **231**(4745): p. 1577-80.
188. Hutchins, C.J., et al., *Self-cleavage of plus and minus RNA transcripts of avocado sunblotch viroid*. Nucleic Acids Res., 1986. **14**(9): p. 3627-40.
189. Forster, A.C. and R.H. Symons, *Self-cleavage of virusoid RNA is performed by the proposed 55-nucleotide active site*. Cell, 1987. **50**(1): p. 9-16.
190. Uhlenbeck, O.C., *A small catalytic oligoribonucleotide*. Nature, 1987. **328**(6131): p. 596-600.
191. Haseloff, J. and W.L. Gerlach, *Sequences required for self-catalysed cleavage of the satellite RNA of tobacco ringspot virus*. Gene, 1989. **82**(1): p. 43-52.
192. Ferbeyre, G., J.M. Smith, and R. Cedergren, *Schistosome satellite DNA encodes active hammerhead ribozymes*. Mol. Cell. Biol., 1998. **18**(7): p. 3880-8.
193. Bourdeau, V., et al., *The distribution of RNA motifs in natural sequences*. Nucleic Acids Res., 1999. **27**(22): p. 4457-67.
194. Hammann, C., et al., *The ubiquitous hammerhead ribozyme*. RNA, 2012. **18**(5): p. 871-85.
195. Perreault, J., et al., *Identification of hammerhead ribozymes in all domains of life reveals novel structural variations*. PLoS Comput. Biol., 2011. **7**(5): p. e1002031.
196. Blount, K.F. and O.C. Uhlenbeck, *The structure-function dilemma of the hammerhead ribozyme*. Annu. Rev. Biophys. Biomol. Struct., 2005. **34**: p. 415-40.

197. Nelson, J.A. and O.C. Uhlenbeck, *Minimal and extended hammerheads utilize a similar dynamic reaction mechanism for catalysis*. RNA, 2008. **14**(1): p. 43-54.
198. Garcia-Robles, I., J. Sanchez-Navarro, and M. de la Pena, *Intronic hammerhead ribozymes in mRNA biogenesis*. Biol. Chem., 2012. **393**(11): p. 1317-26.
199. Weinberg, Z., et al., *Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes*. Genome Biol., 2010. **11**(3): p. R31.
200. Murchie, A.I., et al., *Folding of the hairpin ribozyme in its natural conformation achieves close physical proximity of the loops*. Mol. Cell, 1998. **1**(6): p. 873-81.
201. Walter, N.G., J.M. Burke, and D.P. Millar, *Stability of hairpin ribozyme tertiary structure is governed by the interdomain junction*. Nat. Struct. Biol., 1999. **6**(6): p. 544-9.
202. Wilson, T.J., et al., *Importance of specific nucleotides in the folding of the natural form of the hairpin ribozyme*. Biochemistry, 2001. **40**(7): p. 2291-302.
203. Rupert, P.B., et al., *Transition state stabilization by a catalytic RNA*. Science, 2002. **298**(5597): p. 1421-4.
204. Kuzmin, Y.I., et al., *Role of an active site adenine in hairpin ribozyme catalysis*. J. Mol. Biol., 2005. **349**(5): p. 989-1010.
205. Lilley, D.M.J., *The hairpin and Varkud satellite ribozymes*, in *Ribozymes and RNA catalysis*, D.M.J. Lilley and F. Eckstein, Editors. 2008, Royal Society of Chemistry: Cambridge. p. 66-91.
206. Wilson, T.J. and D.M. Lilley, *Do the hairpin and VS ribozymes share a common catalytic mechanism based on general acid-base catalysis? A critical assessment of available experimental data*. RNA, 2011. **17**(2): p. 213-21.
207. Kennell, J.C., et al., *The VS catalytic RNA replicates by reverse transcription as a satellite of a retroplasmid*. Genes & Devel., 1995. **9**: p. 294-303.
208. D'Souza, A.D., S. Sultana, and R. Maheshwari, *Characterization and prevalence of a circular mitochondrial plasmid in senescence-prone isolates of Neurospora intermedia*. Curr. Genet., 2005. **47**(3): p. 182-93.
209. Saville, B.L. and R.A. Collins, *RNA-mediated ligation of self-cleavage products of a Neurospora mitochondrial plasmid transcript*. Proc. Natl. Acad. Sci. U. S. A., 1991. **88**: p. 8826-8830.

210. Keeping, A.J. and R.A. Collins, *Evidence for the phenotypic neutrality of the Neurospora Varkud (V) and Varkud satellite (VS) plasmids*. Fungal Genet. Biol., 2011. **48**(7): p. 741-6.
211. Guo, H.C.T., et al., *Nucleotide sequence requirements for self-cleavage of Neurospora VS RNA*. J. Mol. Biol., 1993. **232**: p. 351-361.
212. Collins, R.A. and J.E. Olive, *Reaction conditions and kinetics of self-cleavage of a ribozyme derived from Neurospora VS RNA*. Biochemistry, 1993. **32**: p. 2795-2799.
213. Murray, J.B., et al., *The hammerhead, hairpin and VS ribozymes are catalytically proficient in monovalent cations alone*. Chem. Biol., 1998. **5**: p. 587-595.
214. Guo, H.C.T. and R.A. Collins, *Efficient trans-cleavage of a stem-loop RNA substrate by a ribozyme derived from Neurospora VS RNA*. EMBO J., 1995. **14**(12): p. 368-376.
215. Maguire, J.L. and R.A. Collins, *Effects of cobalt hexammine on folding and self-cleavage of the Neurospora VS ribozyme*. J. Mol. Biol., 2001. **309**: p. 45-56.
216. Olive, J.E. and R.A. Collins, *Spermine switches a Neurospora VS ribozyme from slow cis cleavage to fast trans cleavage*. Biochemistry, 1998. **37**(18): p. 6476-6484.
217. Ouellet, J., M. Byrne, and D.M. Lilley, *Formation of an active site in trans by interaction of two complete Varkud Satellite ribozymes*. RNA, 2009. **15**(10): p. 1822-6.
218. Beattie, T.L., J.E. Olive, and R.A. Collins, *A secondary-structure model for the self-cleaving region of Neurospora VS RNA*. Proc. Natl. Acad. Sci. U. S. A., 1995. **92**: p. 4686-4690.
219. Collins, R.A., *The Neurospora Varkud satellite ribozyme*. Bioch. Soc. Trans. Rev., 2002. **30**: p. 1122-1126.
220. Lafontaine, D.A., et al., *The A730 loop is an important component of the active site of the VS ribozyme*. J. Mol. Biol., 2001. **312**: p. 663-674.
221. Rastogi, T. and R.A. Collins, *Smaller, faster ribozymes reveal the catalytic core of Neurospora VS RNA*. J. Mol. Biol., 1998. **277**: p. 215-224.
222. Beattie, T.L. and R.A. Collins, *Identification of functional domains in the self-cleaving Neurospora VS ribozyme using damage selection*. J. Mol. Biol., 1997. **267**: p. 830-840.

223. Sood, V.D. and R.A. Collins, *Identification of the catalytic subdomain of the VS ribozyme and evidence for remarkable sequence tolerance in the active site loop*. J. Mol. Biol., 2002. **320**: p. 443-454.
224. Hiley, S.L. and R.A. Collins, *Rapid formation of a solvent-inaccessible core in the Neurospora Varkud satellite ribozyme*. EMBO J., 2001. **20**(19): p. 5461-5469.
225. Andersen, A.A. and R.A. Collins, *Intramolecular secondary structure rearrangement by the kissing interaction of the Neurospora VS ribozyme*. Proc. Natl. Acad. Sci U. S. A., 2001. **98**(14): p. 7730-7735.
226. Andersen, A. and R.A. Collins, *Rearrangement of a stable RNA secondary structure during VS ribozyme catalysis*. Mol. Cell, 2000. **5**: p. 469-478.
227. Bouchard, P. and P. Legault, *Structural insights into substrate recognition by the Neurospora Varkud satellite ribozyme: importance of U-turns at the kissing-loop junction*. Biochemistry, 2014. **53**(1): p. 258-69.
228. Flinders, J. and T. Dieckmann, *A pH controlled conformational switch in the cleavage site of the VS ribozyme substrate RNA*. J. Mol. Biol., 2001. **308**: p. 665-679.
229. Bouchard, P. and P. Legault, *A remarkably stable kissing-loop interaction defines substrate recognition by the Neurospora VS Ribozyme*. RNA, 2014. **20**(9): p. 1451-64.
230. Campbell, D.O. and P. Legault, *NMR structure of the Varkud satellite ribozyme stem-loop V RNA and magnesium-ion binding from chemical-shift mapping*. Biochemistry, 2005. **44**: p. 4157-4170.
231. Campbell, D.O., et al., *NMR structure of Varkud satellite ribozyme stem-loop V in the presence of magnesium ions and localization of metal-binding sites*. Biochemistry, 2006. **45**(35): p. 10591-605.
232. Bouchard, P. and P. Legault, *Structural insights into substrate recognition by the Neurospora Varkud satellite ribozyme: importance of U-turns at the kissing-loop junction*. Biochemistry, 2014. **53**(1): p. 258-69.
233. Sood, V.D., T.L. Beattie, and R.A. Collins, *Identification of phosphate groups involved in metal binding and tertiary interactions in the core of the Neurospora VS ribozyme*. J. Mol. Biol., 1998. **282**: p. 741-750.

234. DeAbreu, D.M., J.E. Olive, and R.A. Collins, *Additional roles of a peripheral loop-loop interaction in the Neurospora VS ribozyme*. Nucleic Acids Res., 2011. **39**(14): p. 6223-8.
235. Lacroix-Labonte, J., et al., *Helix-length compensation studies reveal the adaptability of the VS ribozyme architecture*. Nucleic Acids Res., 2012. **40**(5): p. 2284-93.
236. Zamel, R. and R.A. Collins, *Rearrangement of substrate secondary structure facilitates binding to the Neurospora VS ribozyme*. J. Mol. Biol., 2002. **324**: p. 903-915.
237. Wilson, T.J., A.C. McLeod, and D.M. Lilley, *A guanine nucleobase important for catalysis by the VS ribozyme*. EMBO J., 2007. **26**(10): p. 2489-500.
238. Jaikaran, D., et al., *An important role of G638 in the cis-cleavage reaction of the Neurospora VS ribozyme revealed by a novel nucleotide analog incorporation method*. RNA, 2008. **14**(5): p. 938-49.
239. Wilson, T.J., et al., *Nucleobase-mediated general acid-base catalysis in the Varkud satellite ribozyme*. Proc. Natl. Acad. Sci. U. S. A., 2010. **107**(26): p. 11751-6.
240. Tzokov, S.B., I.A. Murray, and J.A. Grasby, *The role of magnesium ions and 2'-hydroxyl groups in the VS ribozyme-substrate interaction*. J. Mol. Biol., 2002. **324**: p. 215-226.
241. Kovacheva, Y.S., et al., *The role of phosphate groups in the VS ribozyme-substrate interaction*. Nucleic Acids Res., 2004. **32**(21): p. 6240-50.
242. Sood, V.D. and R.A. Collins, *Functional equivalence of the uridine turn and the hairpin as building blocks of tertiary structure in the Neurospora VS ribozyme*. J. Mol. Biol., 2001. **313**: p. 1013-1019.
243. Lafontaine, D.A., D.G. Norman, and D.M. Lilley, *The global structure of the VS ribozyme*. EMBO J., 2002. **21**(10): p. 2461-2471.
244. Zhao, Z.Y., et al., *Nucleobase participation in ribozyme catalysis*. J. Am. Chem. Soc., 2005. **127**(14): p. 5026-7.
245. Lafontaine, D.A., et al., *Functional group requirements in the probable active site of the VS ribozyme*. J. Mol. Biol., 2002. **323**: p. 23-34.

246. Jones, F.D. and S.A. Strobel, *Ionization of a critical adenosine residue in the Neurospora Varkud satellite ribozyme active site*. *Biochemistry*, 2003. **42**: p. 4265-4276.
247. Suydam, I.T. and S.A. Strobel, *Fluorine substituted adenosines as probes of nucleobase protonation in functional RNAs*. *J. Am. Chem. Soc.*, 2008. **130**(41): p. 13639-48.
248. Hiley, S.L., et al., *4-thio-U cross-linking identifies the active site of the VS ribozyme*. *EMBO J.*, 2002. **21**(17): p. 4691-4698.
249. Lafontaine, D.A., D.G. Norman, and D.M. Lilley, *Structure, folding and activity of the VS ribozyme: importance of the 2-3-6 helical junction*. *EMBO J.*, 2001. **20**(6): p. 1415-1424.
250. Pereira, M.J., et al., *Single VS ribozyme molecules reveal dynamic and hierarchical folding toward catalysis*. *J. Mol. Biol.*, 2008. **382**(2): p. 496-509.
251. Smith, M.D. and R.A. Collins, *Evidence for proton transfer in the rate-limiting step of a fast-cleaving Varkud satellite ribozyme*. *Proc. Natl. Acad. Sci. U. S. A.*, 2007. **104**(14): p. 5818-23.
252. Wilcox, J.L., A.K. Ahluwalia, and P.C. Bevilacqua, *Charged nucleobases and their potential for RNA catalysis*. *Acc. Chem. Res.*, 2011. **44**(12): p. 1270-9.
253. Hoffmann, B., et al., *NMR structure of the active conformation of the Varkud satellite ribozyme cleavage site*. *Proc. Natl. Acad. Sci. U. S. A.*, 2003. **100**(12): p. 7003-7008.
254. Smith, M.D., et al., *The ionic environment determines ribozyme cleavage rate by modulation of nucleobase pKa*. *RNA*, 2008. **14**(9): p. 1942-9.
255. Smith, M.D. and R.A. Collins, *Use of ribozyme cleavage kinetics to measure salt-induced changes in solution pH*. *Anal. Biochem.*, 2011. **415**(1): p. 12-20.
256. Duss, O., P.J. Lukavsky, and F.H. Allain, *Isotope labeling and segmental labeling of larger RNAs for NMR structural studies*. *Adv. Exp. Med. Biol.*, 2012. **992**: p. 121-44.
257. Michiels, P.J., et al., *Structure of the ribozyme substrate hairpin of Neurospora VS RNA: A close look at the cleavage site*. *RNA*, 2000. **6**: p. 1821-1832.
258. Flinders, J. and T. Dieckmann, *The solution structure of the VS ribozyme active site loop reveals a dynamic "hot-spot"*. *J. Mol. Biol.*, 2004. **341**: p. 935-949.

259. Lafontaine, D.A., D.G. Norman, and D.M.J. Lilley, *Folding and catalysis by the VS ribozyme*. *Biochimie*, 2002. **84**: p. 889-896.
260. Lipfert, J., et al., *The complete VS ribozyme in solution studied by small-angle X-ray scattering*. *Structure*, 2008. **16**(9): p. 1357-67.
261. Friebolin, H., *Basic one- and two-dimensional NMR spectroscopy*. John Wiley & Sons. 406, 2005.
262. Furtig, B., et al., *NMR spectroscopy of RNA*. *Chembiochem*, 2003. **4**(10): p. 936-62.
263. Vuister, G.W. and A. Bax, *Resolution enhancement and spectral editing of uniformly ¹³C-enriched proteins by homonuclear broadband ¹³C decoupling*. *J. Magn. Reson.*, 1992. **98**: p. 428-435.
264. Santoro, J. and G.C. King, *A constant-time 2D overboderhausen experiment for inverse correlation of isotopically enriched species*. *J. Magn. Reson.*, 1992. **97**: p. 202-207.
265. Ikura, M., et al., *Three-dimensional NOESY-HMQC spectroscopy of a ¹³C-labeled protein*. *J. Magn. Reson.*, 1990. **86**: p. 204-209.
266. Pardi, A. and E.P. Nikonowicz, *Simple procedure for resonance assignment of the sugar protons in ¹³C-labeled RNAs*. *J. Am. Chem. Soc.*, 1992. **114**: p. 9202-9203.
267. Wüthrich, K., *NMR of proteins and acids nucleic*. Wiley, 1986.
268. Zhang, O., et al., *Backbone ¹H and ¹⁵N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques*. *J. Biomol. NMR*, 1994. **4**(6): p. 845-58.
269. Pascal, S.M., et al., *Simultaneous acquisition of ¹⁵N and ¹³C-edited NOE spectra of proteins dissolved in H₂O*. *J. Magn. Res. Series B*, 1994. **103**: p. 197-201.
270. Dieckmann, T., et al., *Biomolecular NMR Spectroscopy of Ribonucleic Acids*. eLS. John Wiley & Sons, Ltd Chichester, 2014.
271. Lipsitz, R.S. and N. Tjandra, *Residual dipolar couplings in NMR structure analysis*. *Annu. Rev. Biophys. Biomol. Struct.*, 2004. **33**: p. 387-413.
272. Schwieters, C.D., et al., *The Xplor-NIH NMR molecular structure determination package*. *J. Magn. Reson.*, 2003. **160**: p. 66-74.
273. Bruschiweiler, R., *New approaches to the dynamic interpretation and prediction of NMR relaxation data from proteins*. *Curr. Opin. Struct. Biol.*, 2003. **13**(2): p. 175-83.

274. Cromsigt, J., et al., *Preparation of partially 2H/13C-labelled RNA for NMR studies. Stereo-specific deuteration of the H5' in nucleotides*. Nucleic Acids Res., 2002. **30**(7): p. 1639-45.
275. Dagenais, P. and P. Legault, *Preparative separation of ribonucleoside monophosphates by ion-pair reverse-phase HPLC*. Methods Mol. Biol., 2012. **941**: p. 247-56.
276. D'Souza, V., et al., *NMR structure of the 101-nucleotide core encapsidation signal of the Moloney murine leukemia virus*. J. Mol. Biol., 2004. **337**(2): p. 427-42.
277. Kim, I., P.J. Lukavski, and J.D. Puglisi, *NMR study of 100 kDa HCV IRES RNA using segmental isotope labeling*. J. Am. Chem. Soc., 2002. **124**: p. 9338-9339.
278. Sinha, K., L. Jen-Jacobson, and G.S. Rule, *Divide and conquer is always best: sensitivity of methyl correlation experiments*. J. Biomol. NMR, 2013. **56**(4): p. 331-5.
279. Frueh, D.P., et al., *NMR methods for structural studies of large monomeric and multimeric proteins*. Curr. Opin. Struct. Biol., 2013. **23**(5): p. 734-9.
280. Ziegeler, M., et al., *NMR studies of HAR1 RNA secondary structures reveal conformational dynamics in the human RNA*. Chembiochem, 2012. **13**(14): p. 2100-12.
281. Lukavsky, P.J., et al., *Structure of HCV IRES domain II determined by NMR*. Nat. Struct. Biol., 2003. **10**(12): p. 1033-8.
282. Bothe, J.R., et al., *Characterizing RNA dynamics at atomic resolution using solution-state NMR spectroscopy*. Nat. Methods, 2011. **8**(11): p. 919-31.
283. Legault, P. and A. Pardi, *Unusual dynamics and pKa shifts at the active site of a lead-dependent ribozyme*. J. Am. Chem. Soc., 1997. **119**: p. 6621-6628.
284. Gonzalez, R.L., Jr. and I. Tinoco, Jr., *Identification and characterization of metal ion binding sites in RNA*. Methods Enzymol., 2001. **338**: p. 421-443.
285. Davis, J.H., et al., *Role of metal ions in the tetraloop-receptor complex as analyzed by NMR*. RNA, 2007. **13**(1): p. 76-86.
286. Hammann, C., D.G. Norman, and D.M. Lilley, *Dissection of the ion-induced folding of the hammerhead ribozyme using 19F NMR*. Proc. Natl. Acad. Sci. U. S. A., 2001. **98**(10): p. 5503-8.

287. Maderia, M., L.M. Hunsicker, and V.J. DeRose, *Metal-phosphate interactions in the hammerhead ribozyme observed by ^{31}P NMR and phosphorothioate substitutions*. *Biochemistry*, 2000. **39**(40): p. 12113-20.
288. Maderia, M., T.E. Horton, and V.J. DeRose, *Metal interactions with a GAAA RNA tetraloop characterized by ^{31}P NMR and phosphorothioate substitution*. *Biochemistry*, 2000. **39**: p. 8193-8200.
289. Lilley, D.M., *Catalysis by the nucleolytic ribozymes*. *Biochem. Soc. Trans.*, 2011. **39**(2): p. 641-6.
290. McLeod, A.C. and D.M. Lilley, *Efficient, pH-dependent RNA ligation by the VS ribozyme in trans*. *Biochemistry*, 2004. **43**(4): p. 1118-25.
291. Sood, V.D., S. Yekta, and R.A. Collins, *The contribution of 2'-hydroxyls to the cleavage activity of the Neurospora VS ribozyme*. *Nucleic Acids Res.*, 2002. **30**(5): p. 1132-1138.
292. Keel, A.Y., et al., *A general strategy to solve the phase problem in RNA crystallography*. *Structure*, 2007. **15**(7): p. 761-72.
293. Rieder, R., et al., *Ligand-induced folding of the adenosine deaminase A-riboswitch and implications on riboswitch translational control*. *Chembiochem*, 2007. **8**(8): p. 896-902.
294. Kinoshita, Y., K. Nishigaki, and Y. Husimi, *Fluorescence-, isotope- or biotin-labeling of the 5'-end of single-stranded DNA/RNA using T4 RNA ligase*. *Nucleic Acids Res.*, 1997. **25**(18): p. 3747-8.
295. Duss, O., et al., *A fast, efficient and sequence-independent method for flexible multiple segmental isotope labeling of RNA using ribozyme and RNase H cleavage*. *Nucleic Acids Res.*, 2010. **38**(20): p. e188.
296. Hobartner, C., et al., *Syntheses of RNAs with up to 100 nucleotides containing site-specific 2'-methylseleno labels for use in X-ray crystallography*. *J. Am. Chem. Soc.*, 2005. **127**(34): p. 12035-45.
297. Serganov, A., et al., *Structural basis for Diels-Alder ribozyme-catalyzed carbon-carbon bond formation*. *Nat. Struct. Mol. Biol.*, 2005. **12**(3): p. 218-24.
298. Dieckmann, T. and J. Feigon, *Assignment methodology for larger RNA oligonucleotides: application to an ATP-binding RNA aptamer*. *J. Biomol. NMR*, 1997. **9**(3): p. 259-72.

299. Anfinsen, C.B., *Principles that govern the folding of protein chains*. Science, 1973. **181**(4096): p. 223-30.
300. Sela, M., F.H. White, Jr., and C.B. Anfinsen, *Reductive cleavage of disulfide bridges in ribonuclease*. Science, 1957. **125**(3250): p. 691-2.
301. Rother, K., et al., *RNA and protein 3D structure modeling: similarities and differences*. J. Mol. Model., 2011. **17**(9): p. 2325-36.
302. McDowell, S.E., et al., *Molecular dynamics simulations of RNA: an in silico single molecule approach*. Biopolymers, 2007. **85**(2): p. 169-84.
303. Chothia, C. and A.M. Lesk, *The relation between the divergence of sequence and structure in proteins*. EMBO J., 1986. **5**(4): p. 823-6.
304. Dror, O., R. Nussinov, and H. Wolfson, *ARTS: alignment of RNA tertiary structures*. Bioinformatics, 2005. **21 Suppl 2**: p. ii47-53.
305. Magnus, M., et al., *Computational modeling of RNA 3D structures, with the aid of experimental restraints*. RNA Biol., 2014. **11**(5): p. 522-36.
306. Das, R. and D. Baker, *Automated de novo prediction of native-like RNA tertiary structures*. Proc. Natl. Acad. Sci. U. S. A., 2007. **104**(37): p. 14664-9.
307. Das, R., J. Karanicolas, and D. Baker, *Atomic accuracy in predicting and designing noncanonical RNA structure*. Nat. Methods, 2010. **7**(4): p. 291-4.
308. Parisien, M. and F. Major, *The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data*. Nature, 2008. **452**(7183): p. 51-5.
309. Chojnowski, G., T. Walen, and J.M. Bujnicki, *RNA Bricks--a database of RNA 3D motifs and their interactions*. Nucleic Acids Res., 2014. **42**(Database issue): p. D123-31.
310. Tus, A., et al., *BioMe: biologically relevant metals*. Nucleic Acids Res., 2012. **40**(Web Server issue): p. W352-7.
311. Schnabl, J., P. Suter, and R.K. Sigel, *MINAS--a database of Metal Ions in Nucleic AcidS*. Nucleic Acids Res., 2012. **40**(Database issue): p. D434-8.
312. Petrov, A.I., C.L. Zirbel, and N.B. Leontis, *WebFR3D--a server for finding, aligning and analyzing recurrent RNA 3D motifs*. Nucleic Acids Res., 2011. **39**(Web Server issue): p. W50-5.
313. Bonneau, E. and P. Legault, *Nuclear magnetic resonance structure of the III-IV-V three-way junction from the Varkud Satellite ribozyme and identification of*

- magnesium-binding sites using paramagnetic relaxation enhancement.* Biochemistry, 2014. **53**(39): p. 6264-75.
314. Bonneau, E. and P. Legault, *NMR localization of divalent cations at the active site of the Neurospora VS ribozyme provides insights into RNA-metal-ion interactions.* Biochemistry, 2014. **53**(3): p. 579-90.
315. Laine, S., et al., *In vitro and in vivo cleavage of HIV-1 RNA by new SOFA-HDV ribozymes and their potential to inhibit viral replication.* RNA Biol., 2011. **8**(2): p. 343-53.
316. Klauser, B. and J.S. Hartig, *An engineered small RNA-mediated genetic switch based on a ribozyme expression platform.* Nucleic Acids Res., 2013. **41**(10): p. 5542-52.
317. Levesque, D., F.P. Briere, and J.P. Perreault, *A modern mode of activation for nucleic acid enzymes.* PLoS One, 2007. **2**(7): p. e673.
318. Davidson, A., et al., *Essential structural requirements for specific recognition of HIV TAR RNA by peptide mimetics of Tat protein.* Nucleic Acids Res., 2011. **39**(1): p. 248-56.
319. Gold, L., et al., *Diversity of oligonucleotide functions.* Annu. Rev. Biochem., 1995. **64**: p. 763-97.