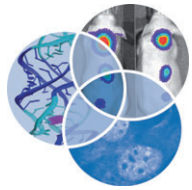
Advanced Review



Insights into RNA structure and dynamics from recent NMR and X-ray studies of the *Neurospora* Varkud satellite ribozyme

Pierre Dagenais,[†] Nicolas Girard,[†] Eric Bonneau[†] and Pascale Legault*

Despite the large number of noncoding RNAs and their importance in several biological processes, our understanding of RNA structure and dynamics at atomic resolution is still very limited. Like many other RNAs, the *Neurospora* Varkud satellite (VS) ribozyme performs its functions through dynamic exchange of multiple conformational states. More specifically, the VS ribozyme recognizes and cleaves its stem-loop substrate via a mechanism that involves several structural transitions within its stem-loop substrate. The recent publications of high-resolution structures of the VS ribozyme, obtained by NMR spectroscopy and X-ray crystallography, offer an opportunity to integrate the data and closely examine the structural and dynamic properties of this model RNA system. Notably, these investigations provide a valuable example of the divide-and-conquer strategy for structural and dynamic characterization of a large RNA, based on NMR structures of several individual subdomains. The success of this divide-and-conquer approach reflects the modularity of RNA architecture and the great care taken in identifying the independently-folding modules. Together with previous biochemical and biophysical characterizations, the recent NMR and X-ray studies provide a coherent picture into how the VS ribozyme recognizes its stem-loop substrate. Such in-depth characterization of this RNA enzyme will serve as a model for future structural and engineering studies of dynamic RNAs and may be particularly useful in planning divide-and-conquer investigations.

How to cite this article:
WIREs RNA 2017, 8:e1421. doi: 10.1002/wrna.1421

INTRODUCTION

Noncoding RNAs form a rich and extremely diverse class of biomolecules found in all domains of life that play critical roles in many biological processes. Although our structural and functional understanding of these RNAs is rapidly growing, there are still very few high-resolution structures of RNA available in comparison with the large number of noncoding RNAs in biology. An important part of what we know at this time comes from the investigations of ribozymes, RNAs that catalyze specific biochemical reactions including phosphoryl transfer and peptide bond formation. Several ribozymes have been shown to function in the absence of proteins and represent ideal model systems for noncoding RNAs due to their relatively small size and easily detectable biological activity. As

[†]These authors contributed equally to this study.

*Correspondence to: pascale.legault@umontreal.ca

Département de Biochimie et Médecine Moléculaire, Université de Montréal, Montréal, Canada

Conflict of interest: The authors have declared no conflicts of interest for this article.

these ribozymes adopt diverse structures and function through distinct mechanisms, previous investigations of ribozymes have given us a broad perspective on the structural diversity and dynamic nature of RNA. The case of the *Neurospora* VS ribozyme is noteworthy because of its unique structure and mode of substrate recognition.

The VS ribozyme was discovered more than 25 years ago as part of the Varkud satellite (VS) RNA, a mitochondrial RNA found in certain isolates of the *Neurospora* filamentous fungi.[1,2] Its self-cleavage and self-ligation activities are essential for the replication cycle of the VS RNA, for which no other specific function has been assigned yet. *In vitro*, a contiguous sequence of approximately 160 nucleotides (nt) containing six helical regions (I–VI) was found sufficient for self-cleavage (Figure 1),[3,5] and extending this sequence with an additional helical domain (VII) was found important for both efficient self-ligation and *trans* cleavage within a dimeric form.[6,7] The cleavage reaction yields products with 2′,3′-cyclic phosphate and 5′-OH termini, as also found for other small nucleolytic ribozymes, like the hairpin, hammerhead, and *glmS* ribozymes.[8–10] A significant number of biochemical, kinetic, and biophysical investigations have allowed extensive functional and mutational characterization of the VS ribozyme[11–15] and have provided low-resolution three-dimensional models (Table 1),[22,27,29,31–33] revealing primary, secondary, and tertiary structures that are unique among known ribozymes. They also support a general acid–base mechanism similar to that of the hairpin ribozyme. Both ribozymes require the intimate association of two internal loops to form the active site where two nucleobases, one from each internal loop, play key roles in the catalytic mechanism.[13,14] A distinctive feature of the VS ribozyme is that one internal loop is located within the catalytic domain (the A_{756} loop also known as the A_{730} loop), whereas the other one (the G_{638} loop) belongs exclusively to its stem-loop substrate, such that the cleavage mechanism can be viewed as substrate-assisted catalysis. In the proposed cleavage mechanism, G_{638} acts as the general base that activates the 2′-oxygen nucleophile, whereas A_{756} acts as the general acid that donates a proton to the 5′-oxygen leaving group.[35] Reported pH-activity profiles of the cleavage reaction fit a double-ionization model with shifted pK_a values of 5.2–5.8 and 8.3–8.4 attributed to A_{756} and G_{638}, respectively.[35–38] Moreover, the metal dependence of these shifted pK_a values point to an indirect role for metal ions in the reaction chemistry.[38] Efficient *trans* cleavage of the stem-loop substrate is made possible through a complex mechanism for substrate recognition (Figure 1), which involves formation of a magnesium-dependent kissing-loop interaction (KLI), an activating conformational change from an unshifted to a shifted state, and intimate association of the G_{638} loop with the A_{756} loop to form the active site.[4,39–41]

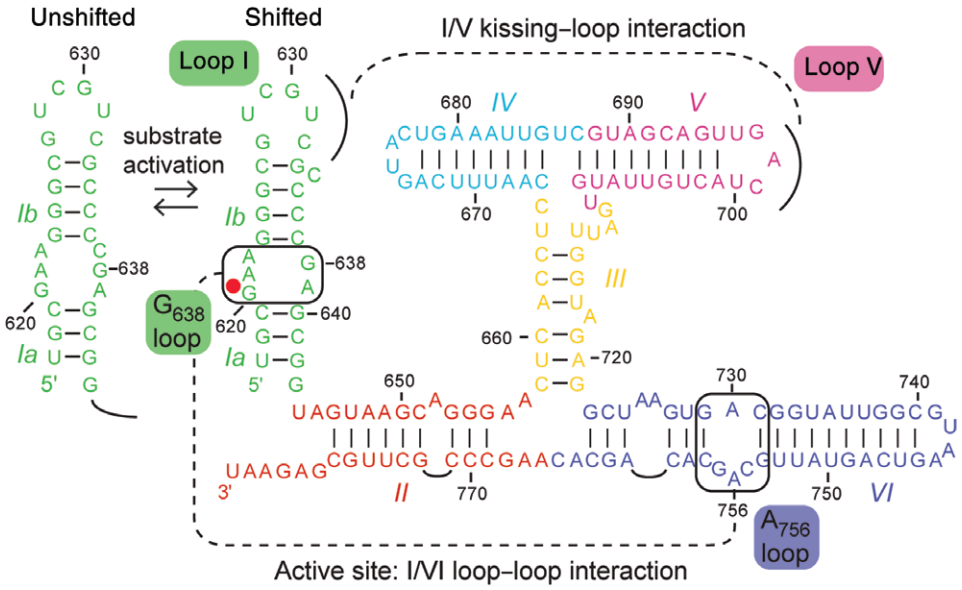


FIGURE 1 | The *Neurospora* VS ribozyme. Primary and secondary structures of the natural *cis*-cleaving VS ribozyme (residues 617–783) determined from mutagenesis and chemical probing studies.[3,4] Recognition of the substrate (stem-loop I or SLI) involves a KLI with stem-loop V (SLV), an activating conformation change (left) and a loop–loop interaction between the G_{638} and A_{756} internal loops to form the active site. Helical domains are color-coded and identified with roman numerals (I–VI). The scissile phosphate is depicted by a red dot.

TABLE 1 | List of Three-Dimensional Models and Structures Reported for VS Ribozyme

Subdomain	Experimental Approach[1]	PDB Code	Divalent Metal Ions in Buffer[2]	Metal Ions in Structure	Reference
SLI unshifted	NMR	1E4P	None		16
	NMR	1HWQ	None		17
SLI shifted	NMR	1OW9	None		18
SLV	NMR	1TBK	None		19
	NMR	1YN1	≥20 mM Mg^{2+}		20
	NMR	1YN2	≥20 mM Mg^{2+} and 20 μM Mn^{2+}	4 $Mn(H_2O)_6^{2+}$	20
	MD		None; 40 mM Mg^{2+}		21
I/V KLI	CP/MI/SDM		0–15 mM Mg^{2+}		22
	NMR	2MI0	5 mM Mg^{2+}		23
A_{756} loop	NMR	1TJZ	None		24
	NMR	2L5Z	5 mM Mg^{2+}		25
	NMR	2MIS	5 mM Mg^{2+} and 10 μM Mn^{2+}; 4.75 mM Mg^{2+} and 0.25 mM Cd^{2+}	3 $Mg(H_2O)_6^{2+}$ 2 $Mg(H_2O)_5^{2+}$	26
III–IV–V junction	FRET/CGE		3 mM Mg^{2+}		27
	NMR	2MTJ	5 mM Mg^{2+}		28
	NMR	2MTK	5 mM Mg^{2+} and 10 μM Mn^{2+}	5 $Mg(H_2O)_6^{2+}$ 1 $Mg(H_2O)_5^{2+}$	28
II–III–VI junction	FRET		2 mM Mg^{2+}		29
	NMR	2N3Q	5 mM Mg^{2+}		30
	NMR	2N3R	5 mM Mg^{2+} and 10 μM Mn^{2+}	8 $Mg(H_2O)_6^{2+}$	30
VS ribozyme (*cis* or *trans* form)	CP/MI/SDM		0–20 mM Mg^{2+}		31
	FRET/CGE		2–3 mM Mg^{2+}		27,32
	smFRET		35 mM Mg^{2+}		32
	SAXS		10 mM Mg^{2+}		33
	NMR		5 mM Mg^{2+}		30
	X-ray ($A_{756}G$)	4R4P	5 mM Mg^{2+}	2 Mg^{2+}	34
	X-ray ($G_{638}A$)	4R4V	5 mM Mg^{2+}	9 Mg^{2+}, 5 K^+	34

[1] The experimental approaches listed here are NMR spectroscopy (NMR), molecular dynamic simulations (MD), Föster resonance energy transfer (FRET), comparative gel electrophoresis (CGE), single-molecule FRET (smFRET), chemical probing (CP), modification interference (MI), site-directed mutagenesis (SDM), small-angle X-ray scattering (SAXS), and X-ray crystallography (X-ray). Only NMR and X-ray studies have yielded high-resolution structures reported in the pdb.
[2] In terms of experimental conditions, only the concentration and type of divalent metal ions used for structure determination is provided (none = none used).

Despite the wealth of knowledge accumulated over the years on the VS ribozyme, it is only very recently that high-resolution three-dimensional structures have become available to provide detailed structural and dynamic insights into its mechanism (Table 1). Both a solution-state NMR model[30] and two crystal structures[34] were reported recently that independently represent significant achievements given the inherently dynamic nature of this ribozyme and its tendency to form dimers and higher-order multimers in solution. The NMR model of a *trans* ribozyme in complex with its stem-loop substrate results from a divide-and-conquer strategy for structural characterization based on NMR structures of several individual subdomains. Two crystal structures are available for *cis* ribozyme variants that each form a dimer in which the substrate of one protomer interacts with the catalytic domain of the other protomer. The NMR model represents an open state of the VS ribozyme prior to formation of the active site, and complementary NMR studies have provided information on important conformational changes associated with substrate recognition and activation. However, the crystal structures capture a closed state in which the active site is formed but cleavage has not taken place, thus providing insights into the catalytic mechanism. Herein, we summarize these recent high-resolution structural studies, highlighting the similarities and differences between the NMR and crystal structures to develop an integrated

understanding of the structural and dynamic properties of this ribozyme.

A DIVIDE-AND-CONQUER NMR STUDY PROVIDES DETAILED STRUCTURAL AND DYNAMIC INSIGHTS

The divide-and-conquer approach represents a simple yet powerful strategy to structurally investigate large and dynamic biomolecular systems. It consists of determining high-resolution structures of smaller and often more amenable subdomains to eventually reconstitute the structure of the complete biomolecule. Its recent application to characterize the complete VS ribozyme structure by solution NMR spectroscopy represents a proof-of-concept that this strategy can be used successfully to investigate the structure and dynamics of large functional RNAs. Prior biochemical and biophysical investigations of the VS ribozyme were instrumental in guiding the efforts. A particular challenge was that the limited number of known natural variants of the VS ribozyme did not provide sufficient sequence diversity to define the secondary structure based on covariation analysis as would normally be done for an RNA of interest. Instead, the secondary structure was delineated from careful biochemical characterization based on site-directed mutagenesis and chemical probing studies.[3] Complementary investigations helped distinguish the subdomains that are functionally required [the KLI between stem-loop I (SLI) and stem-loop V (SLV), the internal loops of SLI and stem-loop VI (SLVI), the three adenosine bulges, and the two three-way junctions] from those that are not (junction-distal ends of stems II, IV, and VI).[42,43] They also helped delineate independent subdomains that when taken out of their natural context would remain meaningful to our structural understanding of the VS ribozyme.[3,5,40]

NMR Studies of Isolated Subdomains Relevant to Substrate Recognition

The KLI between the terminal loops of SLI and SLV is important for substrate recognition by the catalytic domain.[39] Prior investigations revealed that this interaction involves three Watson–Crick (WC) base pairs and is associated with a conformational rearrangement that can be rate-limiting, because it activates the substrate for catalysis (Figure 1).[4,39,40,42,44,45] In its free form, the substrate adopts an unshifted conformation, with a symmetric internal loop of six nt. In the bound form, the substrate adopts a shifted conformation that excludes C_{634} from stem Ib and reconfigures the cleavage site internal loop to an asymmetric loop of five nt.[4,40] Importantly, formation of the KLI and the associated conformation change in SLI can be reproduced when the catalytic domain is replaced by an isolated SLV, validating biophysical investigations using isolated stem-loops.[40]

Unshifted and Shifted States of G_{638} Loop

NMR structures of the SLI substrate were determined to characterize both the unshifted (inactive) and shifted (active) states of the substrate cleavage site internal loop, termed here the G_{638} loop.[16–18] Two individual structures were determined for the unshifted state, one of which is illustrated in Figure 2(a).[16,17] In both studies, the RNA stem-loop used comprises the natural internal loop sequence and the adjacent base pairs from stems Ia and Ib. Both NMR structures were determined at acidic pH (pH = 5.0–6.0) and revealed the same structural elements in the G_{638} loop: two *trans* Hoogsteen/sugar edge (H/SE) G–A base pairs and a *cis* WC/WC A(+)•C base pair involving A_{622}. From pH-dependent NMR studies, pK_a values of 6.1–6.4 were determined for A_{622}, and its protonated state was found critical for stabilizing the three noncanonical base pairs of the unshifted internal loop. At higher pH, the G_{638} loop adopts a more open conformation that may facilitate the displacement of residue C_{637} into stem Ib as part of the conformational change associated with formation of the I/V KLI.[17]

The structure of the shifted conformation of the G_{638} loop was also determined by NMR, as part of an RNA stem-loop with a nonnatural stem Ib and a GNRA tetraloop that stabilize the shifted state (Figure 2(b)).[18] As expected, this RNA was cleaved when incubated with a *trans* VS ribozyme, although at a lower rate than a wild-type sequence given that the I/V KLI could not form due to the terminal loop modification.[18] The NMR structure revealed that residues G_{620} and A_{639} form a *trans* H/SE G–A base pair, as found for the unshifted state (Figure 2(a) and (b)).[16–18] However, residue G_{638} is positioned such that it can participate in a *trans* H/SE G–A base pair with either A_{621} or A_{622} (Figure 2(b)). Importantly, G_{638} is stacked on G_{620} with its WC edge close to the 2′-OH of G_{620}, in agreement with its role as general base in the cleavage reaction. Two Mg^{2+}-binding sites were identified in the internal loop based on chemical shift perturbations (Figure 2(b)). A detailed comparison between the NMR structures of the unshifted and shifted states of the G_{638} loop revealed two distinct features of the shifted state that could be important for catalysis: a Mg^{2+}-binding site near

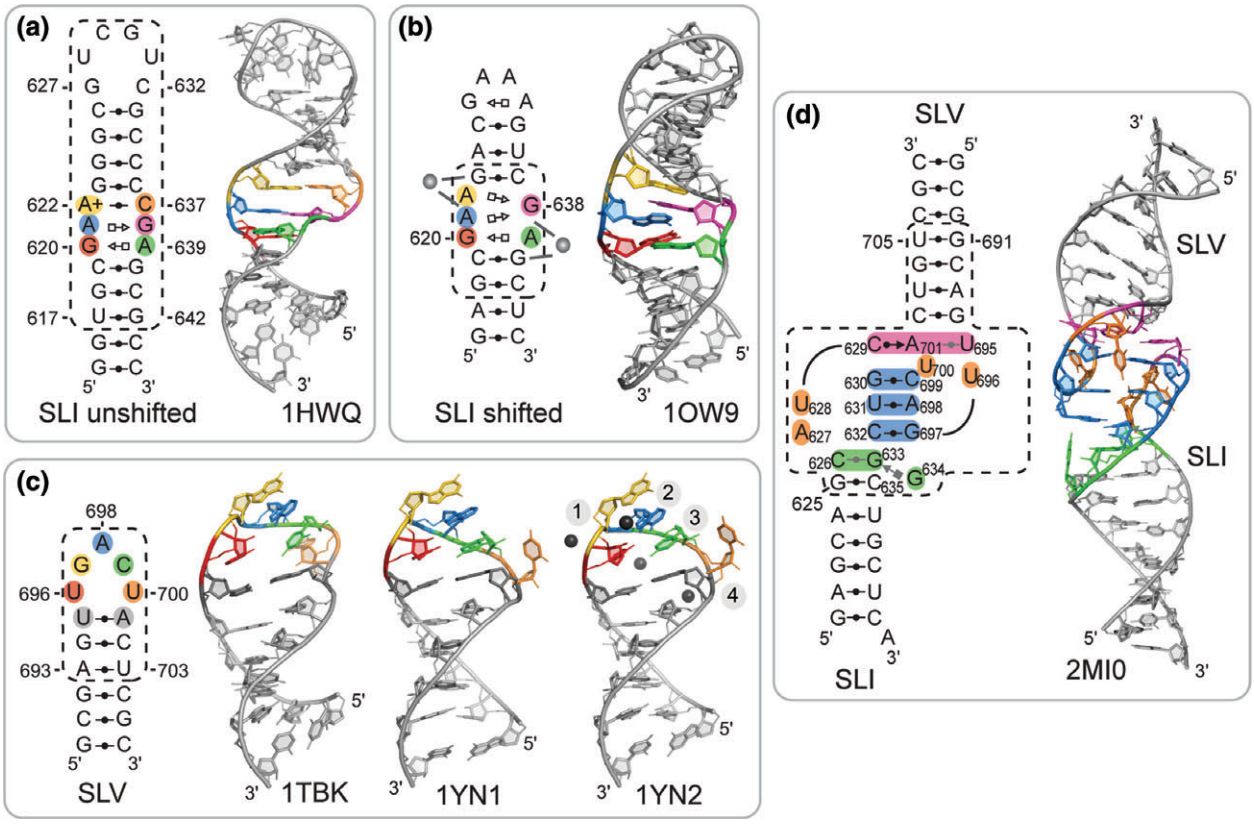


FIGURE 2 | NMR structures of VS ribozyme subdomains determined to characterize substrate recognition and activation. (a) NMR structure of an SLI substrate with the G_{638} loop in the unshifted conformation (PDB code 1HWQ).[17] (b) NMR structure of an SLI substrate with the G_{638} loop in the shifted conformation (PDB code 1OW9).[18] The gray spheres on the structural schematics illustrate Mg^{2+}-binding sites near the 5′-AG-3′ strand of a 5′-AG-3′/5′-CG-3′ motif that are consistent with NMR chemical shift perturbations. (Adapted with permission from Ref 18. Copyright 2003 National Academy of Sciences). (c) NMR structures of the SLV subdomain of the VS ribozyme obtained in the absence of Mg^{2+} ions (PDB code 1TBK)[19] and in the presence of Mg^{2+} ions (PDB code 1YN1 and 1YN2).[20] The NMR structure with the PDB code 1YN2 was determined with RNA-metal restraints to localize hexahydrated Mn^{2+} ions (black spheres). (Adapted with permission. Copyrights 2005 and 2006 American Chemical Society). (d) NMR structure of an SLI/SLV complex (PDB code 2MI0).[23] In panels (a–d), a structural schematic is available alongside a cartoon representation of the NMR structure to illustrate the sequence and base-pairing interactions present in the structure(s): the dashed boxes enclose the sequences derived from the VS ribozyme; specific residues are colored as in the NMR structure representations; and base pairs are depicted using the Leontis–Westhof notation[46,47] in either black (≥2 hydrogen bonds) or gray (1 hydrogen bond).

A_{622} and G_{623} (Figure 2(b)) and strong structural similarity with a tertiary interaction motif found in rRNA.[18] In addition, a pK_a value of 4.0 was estimated for residue A_{622} in this shifted conformation of the G_{638} loop, confirming that the shifted state is stable at neutral pH. Thus, NMR studies established the pH dependence of the equilibrium between the unshifted and shifted states of the SLI substrate and identify structural differences that could be important for its function.

Free Conformations of SLI and SLV Terminal Loops

U-turn folds were initially proposed to form in the terminal loop of both SLI and SLV on the basis of their sequence, mutagenesis studies, and chemical probing experiments.[39] Subsequent NMR structural studies of isolated stem-loops have allowed in-depth characterization of these terminal loops in their free form prior to formation of the I/V KLI.[17,19,20,23] NMR studies of the natural terminal loop of SLI in the absence of Mg^{2+} ions do not provide evidence for the proposed U-turn fold and the adjacent G_{627}–C_{632} base pair, but rather indicate that the structure of the terminal loop is mostly disordered (Figure 2(a), 1HWQ).[17] Addition of Mg^{2+} ions did not bring significant changes in the NMR spectra. Thus, the SLI terminal loop is disordered prior to the formation of the I/V KLI both in the absence and presence of Mg^{2+} ions.

NMR structures of an isolated SLV RNA were determined both in the absence (Figure 2(c), 1TBK) and presence of Mg^{2+} ions (Figure 2(c), 1YN1 and 1YN2).[19,20] In these structures, residues U_{696}, G_{697}, and A_{698} adopt a U-turn fold while residues U_{695} and A_{701} form a canonical base pair that closes the terminal loop. In the presence of Mg^{2+} ions, the terminal loop adopts a more compact structure with more characteristics of the canonical U-turn fold and with U_{700} being completely excluded from the loop fold (Figure 2(c), 1YN1 and 1YN2).[20] Molecular dynamic simulations support the observation that the U-turn of SLV adopts a more compact fold in the presence of Mg^{2+} ions, and indicate that structural characteristics of this compact fold are already present in the structural ensemble obtained in the absence of Mg^{2+} ions.[21] Four divalent metal-binding sites were localized in the SLV terminal loop based on Mn^{2+}-induced paramagnetic relaxation enhancement (PRE) studies (Figure 2(c), 1YN2).[20] Three of them (sites 1–3) correspond to divalent metal-binding sites previously identified in crystal structures of U-turns, and the fourth one (site 4) is associated with the unusual phosphate backbone fold near U_{700}. As a consequence of the U-turn fold, the WC edges of residues after the sharp turn in the phosphate backbone (G_{697}, A_{698}, and C_{699}) are exposed and available for tertiary contact.[19,20]

Highly Stable KLI Between SLI and SLV

To characterize the KLI, several SLI/SLV complexes were investigated by NMR spectroscopy and isothermal titration calorimetry (ITC), but only one was found suitable for NMR structure determination due to its relatively small size (43 nt), the quality of the NMR data and its high thermodynamic stability (K_d = 0.2–0.4 μM).[23,48] This complex was formed with a double-stranded SLI variant that lacks the G_{638} loop, but retains a terminal loop sequence compatible with KLI formation and ribozyme activity ($G_{627}A$ and $C_{634}G$; Figure 2(d)). The three-dimensional structure of this SLI/SLV complex confirms formation of a U-turn fold in each of the two terminal loops and of three canonical *cis* WC/WC base pairs between these loops (Figure 2(d)). In addition, this NMR structure reveals two base triples: a fairly disordered C_{626}–G_{633}–G_{634} triple and a more stable C_{629}–A_{701}–U_{695} triple that contributes a fourth base pair (*cis* WC/SE C_{629}–A_{701}) between the two loops. The structure displays continuous stacking at the KLI that extends to the adjacent stems and likely contributes to the unusually high stability of this interaction.[48] Nevertheless, the interhelical angle between stem I and stem V is quite variable, ranging from 127° to 175° in the ensemble of NMR structures. In addition, there is NMR evidence that several residues of loop I (C_{626}, A_{627}, U_{628}, C_{632}, G_{633}, and G_{634}) and loop V (U_{695} and U_{696}) undergo dynamic exchange on the μs to ms timescale, suggesting that the large range of interhelical angles in the NMR ensemble may not result from the lack of sufficient NMR restraints but rather from the inherent dynamics of the KLI. In the context of the active ribozyme, some degree of flexibility at the kissing-loop junction likely facilitates formation of the active site between the G_{638} and A_{756} loops.

Structural Evidence of Helix Shifting in SLI

NMR studies were conducted to investigate the conformational rearrangement in SLI upon binding to SLV.[23] Briefly, NMR spectra were collected for monitoring the formation of several SLI/SLV complexes in the presence of Mg^{2+} ions using the same SLV, but several SLI variants referred to as shiftable, preshifted, or double-stranded. Comparison of imino proton data between the free and SLV-bound states of these SLI variants provides direct structural evidence for helix shifting in stem Ib upon complex formation with the shiftable SLI variants, but not with the preshifted and the double-stranded SLI variants. The NMR data also reveal that the loop-closing base pair observed in the free SLI is destabilized upon complex formation with both the shiftable and preshifted substrates. Complementary ITC studies confirm that the shiftable substrates have a lower affinity for SLV (12.5 μM ≤ K_d ≤ 65 μM) than the preshifted and double-stranded substrates (0.4 μM ≤ K_d ≤ 0.7 μM), meaning that the helix shift in SLI reduces the stability of the KLI by 2–3 kcal/mol.[48]

These extensive NMR investigations allowed us to define the conformational changes associated with formation of the I/V KLI in the natural VS ribozyme (Figure 3). Prior to binding, the SLV loop adopts a compact U-turn fold and the SLI substrate adopts an unshifted conformation with a disordered terminal loop. Formation of the I/V KLI involves only subtle structural changes in SLV, but a disorder-to-order transition in loop I. The NMR data are consistent with a model in which the initial contact of this disordered loop with the stable loop V destabilizes the closing C_{626}–G_{633} base pair while ordering loop I and thereby initiates the helix shift in SLI. This model is consistent with the ability of U-turns to initiate structural changes in RNA.[49] Thus, the high-energy KLI is likely responsible for triggering the helix shift in SLI and changing the conformation of the G_{638} loop such that it becomes cleavable.

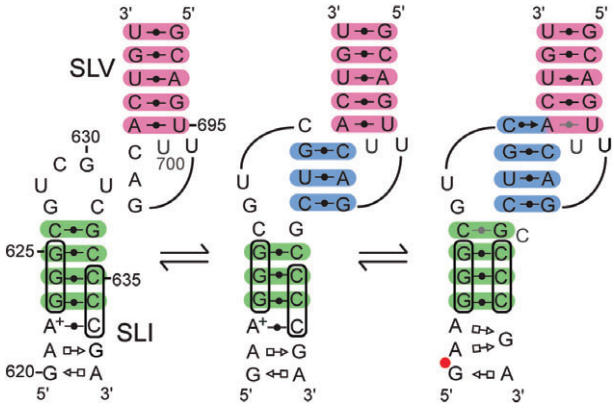


FIGURE 3 | Model for substrate recognition and activation by the VS ribozyme based on NMR data.[23] Structural schematics (as in Figure 2) are used to illustrate the unbound states of SLI and SLV (left), a hypothetical state in which the two loops initiate their contact (middle) and a final bound state (right). Base pairs of stem Ib, the KLI, and stem V are shaded in green, cyan, and pink, respectively.

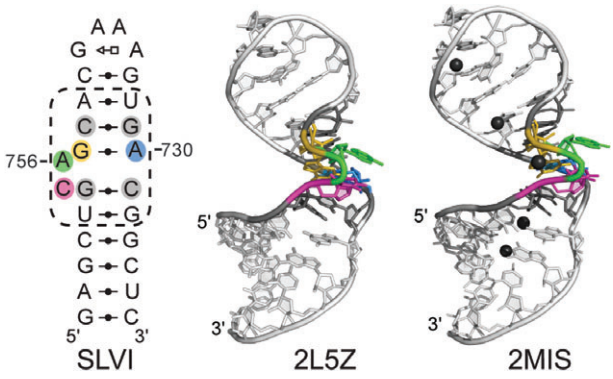


FIGURE 4 | NMR structures of an SLVI subdomain containing the A_{756} loop. Two NMR structures were obtained in the presence of Mg^{2+} ions (PDB code 2L5Z and 2MIS).[25,26] The NMR structure with the PDB code 2MIS was determined with RNA-metal restraints to localize hexahydrated Mn^{2+} ions (black spheres). A structural schematic (as in Figure 2) is available alongside the cartoon representation of the lowest-energy NMR structures to provide the sequence and base-pairing interactions present in both structures. (Adapted from articles that appeared in Oxford University Press[25] and ACS publications[26]).

Free Conformation of A_{756} Loop

For cleavage to occur, the G_{638} loop must interact with the A_{756} loop of SLVI. To investigate the conformation of the A_{756} loop prior to formation of the active site, NMR studies were conducted using a small SLVI subdomain containing the A_{756} loop and its adjacent base pairs (Figure 4).[25] In preliminary NMR studies, it was found that Mg^{2+} ions are required for stable folding of the A_{756} loop and thus Mg^{2+} ions were included in the samples used for structure determination. The NMR structure reveals two main structural elements in the A_{756} loop: a *cis* WC/WC G–A base pair between G_{757} and A_{730} that enlarges the minor groove and an S-turn fold involving residues G_{754}–G_{757}. The S-turn motif generates a distortion of the ribose-phosphate chain that exposes residues C_{755} and A_{756} into the enlarged minor groove with their WC edge facing the solvent. As a result, residue A_{756} adopts an accessible position that is compatible with its role as general acid in the catalytic mechanism.

A dual NMR approach was developed to localize divalent metal ion-binding sites within this SLVI RNA.[26] Mn^{2+}-induced paramagnetic relaxation effect (PRE) studies were used to probe almost all readily detectable ^{1}H, ^{13}C, and ^{15}N atoms in SLVI, whereas Cd^{2+}-induced chemical-shift perturbation (CSP) of RNAs with single phosphorothioate modifications were carried out to specifically probe the contribution of phosphate groups in divalent metal coordination. Using this dual NMR approach, five divalent metal-binding sites were identified (Figure 4), including one at the phosphate cluster formed by the S-turn of the A_{756} loop. Divalent metal ion binding at this site is likely necessary to neutralize the densely-packed electronegative charges at the S-turn and may explain the metal-ion dependence of the A_{756} loop structure. Extrapolating to the potential role of this metal ion at the active site, it may both stabilize the S-turn and play an indirect role in the chemistry of the reaction. Thus, NMR studies of the free conformation of the A_{756} loop provide insights into the formation of the active site and of a potentially important divalent metal ion for catalysis.

NMR Studies of Isolated Subdomains That Define Global Architecture of *trans* Ribozyme

Architecture of III–IV–V Junction

The high-resolution NMR structure of a 47-nt RNA subdomain containing the III–IV–V junction was determined in the presence of Mg^{2+} ions (Figure 5 (a)).[28] It shows a family C topology of three-way junctions, with coaxial stacking of stems III and IV and side-by-side arrangement of stems III and V. A U-turn fold allows backbone reversal between stems III and V and positions the residues after the turn (G_{711} and A_{712}) to form minor groove base triples and a ribose zipper with stem IV.[28] The junction core is stabilized by a continuous stack of four base triples, leading to coaxial stacking of stems III and IV. NMR studies indicate that Mg^{2+} ions are not essential for folding of the junction core. However, the position of six Mg^{2+} ions was determined using Mn^{2+}-induced PRE, revealing two Mg^{2+} ions within the junction core that likely help stabilize the relative orientation of stems III and V. The NMR structure of the III–IV–V junction provided the first evidence,

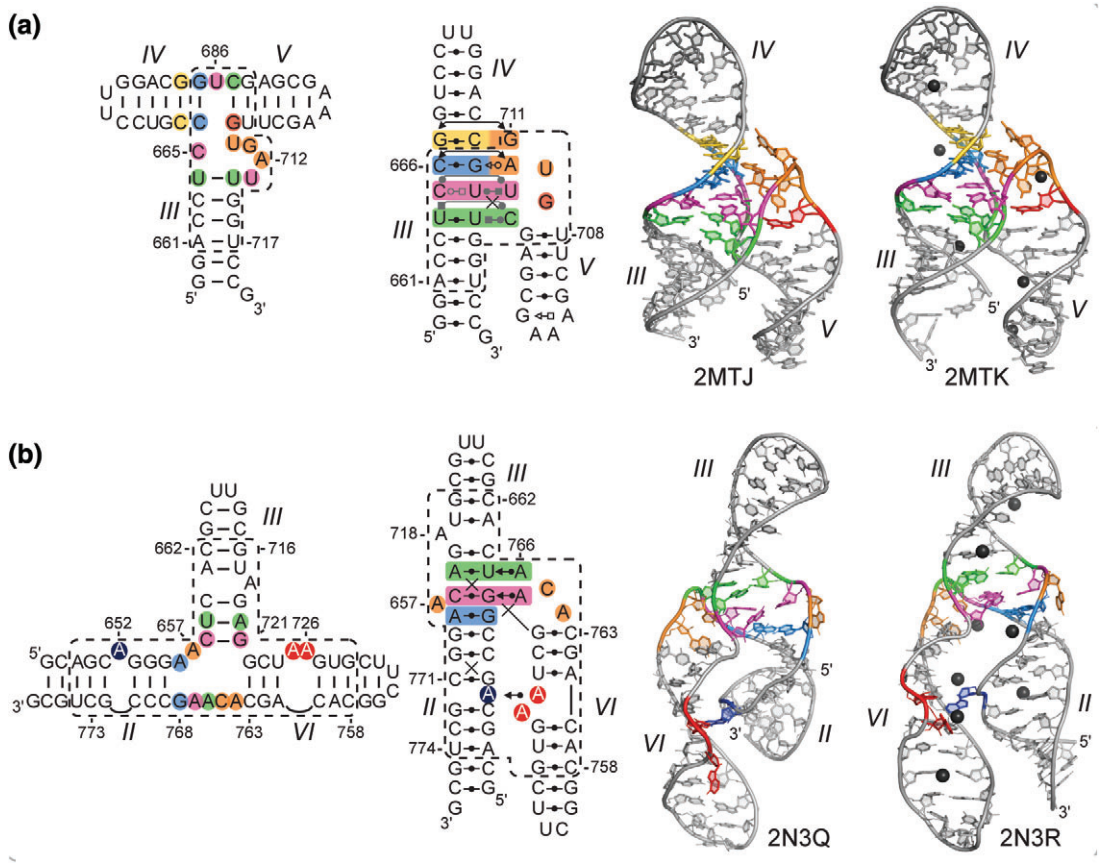


FIGURE 5 | NMR structures of the VS ribozyme (a) III–IV–V(Adapted with permission from Ref 28. Copyrights 2014 American Chemical Society). and (b) II–III–VI[30] three-way junctions. From left to right: proposed secondary structures of the RNA subdomain used for NMR studies; schematic representation (as in Figure 2) of the tertiary structure derived from the NMR structures; cartoon representation of the lowest-energy NMR structure determined in the presence of Mg^{2+} ions; and of the lowest-energy NMR structure determined in the presence of Mg^{2+} ions and using RNA-metal restraints to localize hydrated Mg^{2+} ions (black spheres).

to our knowledge, of a family C three-way junction that does not form a long-range tertiary interaction between the side-by-side helices (stems III and V).[28,50] The extensive stack of base triples allows for formation of a stable junction core in the absence of a distal tertiary interaction. In contrast, less stable three-way junction cores may depend on a distal interaction to stabilize the global folding of the junction, as found for the II–III–VI junction (see below). Given the intrinsic stability of the III–IV–V junction core and lack of a distal tertiary interaction, SLV is free to interact with the SLI substrate to form the I/V KLI. In addition, there is a certain degree of variability in the orientation of SLV with respect to stems III and IV in the NMR ensemble, which is consistent with the III–IV–V junction forming a slightly dynamic hinge. Thus, the III–IV–V junction may assist formation of the active site by allowing the bound SLI substrate to explore the conformational space for optimum docking with the A_{756} loop.[28]

Critical Tertiary Interaction Within II–III–VI Junction

The NMR structure of the II–III–VI junction was determined as part of a 62-nt RNA that incorporates the core residues and adjacent stems with their A-rich bulges, as previous mutational studies had shown the importance of these bulges for the catalytic activity of the VS ribozyme.[27,29,51] The NMR structure determined in the presence of Mg^{2+} ions shows a family C topology of three-way junctions, with coaxial stacking of stems II and III, and side-by-side arrangement of stems II and VI (Figure 5(b)).[30] An ACAA turn (A_{764}–A_{767}) allows backbone reversal between stems II and VI and positions the two residues after the turn (A_{766} and A_{767}) to form base triples (A–U–A and C–G–A) in the minor groove of stem III. The core of the II–III–VI junction is stabilized by these two minor groove base triples and a noncanonical *cis* WC/WC A_{656}–G_{768} base pair, which together form a three-layer stack between stems II and III. The NMR

structure also reveals a long-range tertiary interaction between the bulges of stems II and VI, where A_{652} protrudes from stem II to interact with stem VI, stacking between C_{760} and A_{761} and forming a *cis* SE/WC A_{652}–A_{725} base pair. Interestingly, several Mg^{2+}-binding sites were identified from Mn^{2+}-induced PRE, including two at the junction core and three at the bulge–bulge interaction that are likely needed to counteract the electronegative charges that are brought in proximity of these sites. Complementary NMR studies established that Mg^{2+} ions and the bulge–bulge interaction are both required for the stable folding of the II–III–VI junction.

Model of VS Ribozyme Based on NMR Studies of Isolated Subdomains

Global Architecture of VS Ribozyme Based on NMR Studies: An Open State

NMR studies of VS ribozyme subdomains have allowed the structural characterization of all nonhelical regions of a minimal VS ribozyme competent for cleavage activity (Figure 6(a)). Based on the high-resolution NMR structures of subdomains, a three-dimensional model of a *trans* VS ribozyme bound to a preshifted SLI substrate was built by fragment assembly.[30] This NMR-based model and related structural schematics provided novel insights into the global structure of the VS ribozyme (Figure 6(b)–(d)). Given the specific topology of the two three-way junctions (Figure 5), where stem II stacks on stem III (II–III–VI junction) and stem III stacks on stem IV (III–IV–V junction), the continuous stack of stems II, III, and IV defines the main frame of the VS ribozyme architecture. Stem-loops V and VI are projected alongside this main frame like one arm (SLV) and an opposite leg (SLVI). The SLI substrate associates with the *trans* VS ribozyme via a KLI with the SLV loop, bringing the G_{638} loop of SLI in the vicinity of the A_{756} loop of SLVI. The NMR-based model captures an open-state conformation in which these two internal loops are apart, and this is compatible with chemical probing and biophysical data.[32,53] In all likelihood, such an open state is populated by a dynamic ensemble of ground-state structures. The remodeling of the G_{638} and A_{756} loops as well as the relative motion between the supporting helical domains would be necessary to allow formation of the active site and subsequent cleavage.

Validation of Divide-and-Conquer NMR Approach

The divide-and-conquer approach was undertaken with a good understanding of the structural elements important for structure and function of the VS ribozyme, and this facilitated the identification of RNA subdomains that can fold autonomously. To define suitable RNA subdomains for structural studies, we routinely prepared several different RNA sequences for each subdomain and performed preliminary biochemical, biophysical, and structural characterization. It was important to define an optimal RNA sequence, buffer composition, and folding conditions for which each RNA subdomain adopts a homogeneous and stable fold, and this was verified initially using native gel electrophoresis and NMR spectroscopy, and later using small-angle X-ray scattering (SAXS; Dagenais P and Legault, unpublished data). Imino NMR spectroscopy was particularly valuable to quickly assess conformational homogeneity and define the secondary structure of the isolated RNA subdomains. In addition, the imino NMR spectra of isolated RNA subdomains and that of the *trans* ribozyme were compared to validate structure determination of a given RNA subdomain (Bonneau E, Dagenais P, and Legault P, unpublished data). In some cases, NMR studies of RNA subdomain variants contributed to defining the specific residues required for folding of a given subdomain. For example, NMR studies of two variants of the isolated II–III–VI RNA subdomain revealed the importance of a bulge–bulge interaction for stable folding of this junction.[30] Thus, the II–III–VI RNA subdomain used for NMR structure determination contained the bulges of stems II and VI that form this critical bulge–bulge interaction. In summary, great care was taken to make sure that RNA subdomains were well circumscribed and represented independent folding modules.

Once the high-resolution NMR structure of an RNA subdomain was obtained, it became essential to exhaustively compare the structure with previous biochemical and biophysical data. This comparison with previous studies was particularly important to validate our NMR structure of the SLVI RNA subdomain containing the A_{756} loop, because no structural or functional validation of this subdomain was available prior to structure determination.[25] In contrast, a previous NMR study of the A_{756} loop in a different structural context revealed an alternate conformation that may be relevant to the folding pathway, but that did not agree with biochemical data.[24] Otherwise, the NMR structural studies of isolated subdomains are essentially consistent with a large body of biochemical and biophysical data, thereby supporting the modularity of the VS ribozyme architecture.

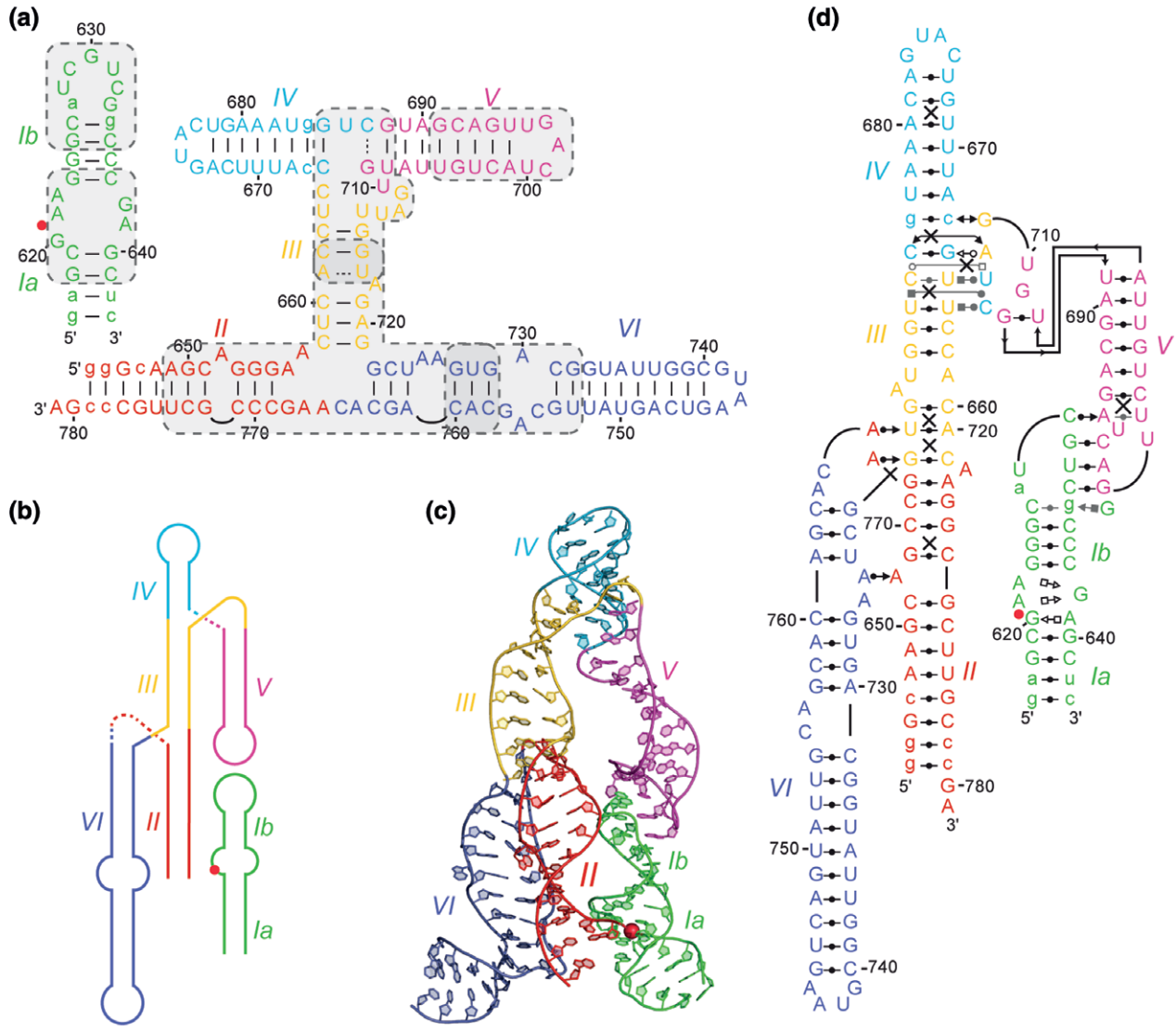


FIGURE 6 | Structural overview of the open state of the VS ribozyme captured by the NMR-based model of a *trans* ribozyme in complex with a preshifted stem-loop substrate.[30] (a) Proposed secondary structure of the complex,[3,4] which was previously characterized kinetically[52] and used to identify important subdomains for NMR structural determination (shaded gray boxes). (b) Simplified global architecture reflecting the relative orientation of helical domains. (c) Cartoon representation of the three-dimensional NMR model. (d) Tertiary structure schematics (as in Figure 2) based on high-resolution NMR structures of isolated subdomains. Other legend details: In panels (a) and (d), nonnatural nucleotides are in lowercase. Also, base pairs of the proposed secondary structure are shown as a solid line when compatible with the structure shown in panel (c) and as a dashed line when not. A uniform color scheme is used for helical domains, and the scissile phosphate is depicted by a red dot or sphere.

CRYSTAL STRUCTURES PROVIDE A DETAILED UNDERSTANDING OF THE VS RIBOZYME STRUCTURE

Overall Crystal Structure and Comparison with NMR Structures of Isolated Subdomains

Trans Recognition of a Stem-Loop Substrate Within a Symmetric Dimer: A Closed State

Crystal structures were determined at 3.1-Å resolution for two active-site variants ($G_{638}A$ and $A_{756}G$) of an extended *cis* VS ribozyme containing helical domain VII in addition to domains I–VI (Figure 7 (a)).[34] The otherwise identical RNA sequence of these two variants was optimized to improve conformational homogeneity and reduce aggregation as needed for structural studies; it incorporates a preshifted SLI substrate and a substitution in loop IV by an ACAAA pentaloop, which serves as epitope for a synthetic antibody fragment.[54] Although such antibody fragment–RNA loop interaction can facilitate crystallization,[54] the crystal lattice did not

incorporate the antibody fragment, but was rather stabilized by crystal contacts mediated by the loop IV pentaloop itself.

The two crystal structures are very similar and both show that the VS ribozyme forms a symmetric dimer, in which the substrate of each protomer interacts with the catalytic domain of the other protomer (Figure 7(b)). This dimeric form may be relevant to efficient *in vivo* processing of the transcribed concatemeric form of the VS RNA, in which the upstream ribozyme catalytic domain could perform cleavage of the downstream stem-loop substrate.[55]

The two central helical axes of the dimer are formed by continuous stacking of helices I–II–III–IV of one protomer on helix V of the partner protomer via the I–V KLI. The III–IV–V and II–III–VI three-way junctions impart a parallel orientation to helices V and VI with respect to the central helices. In

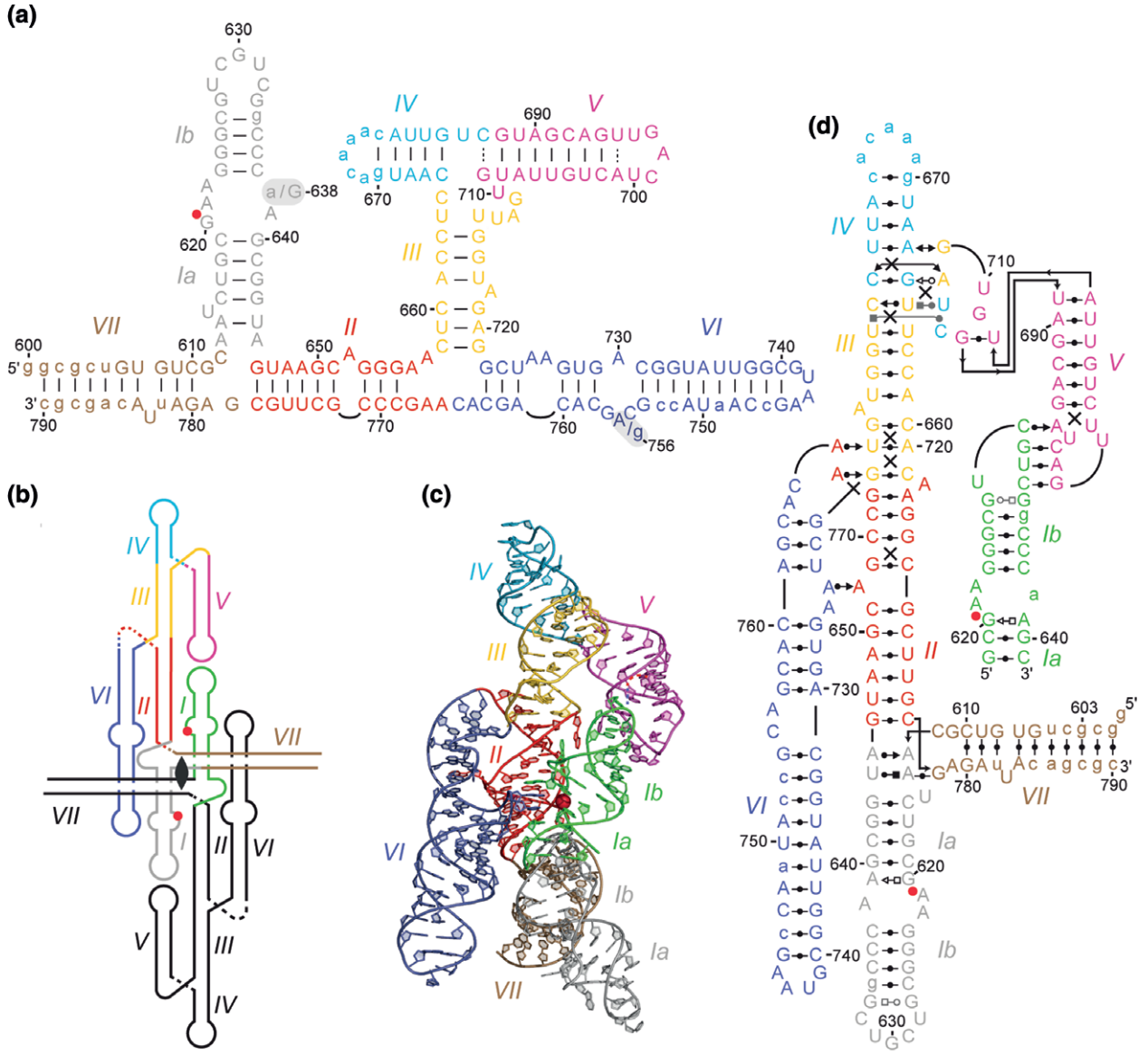


FIGURE 7 | Structural overview of the closed state of the VS ribozyme captured by the crystal structure of a dimeric form.[34] (a) Proposed secondary structure (as in Figure 6) of the two *cis* ribozyme sequences used for crystal structure determination, a $G_{638}A$ and an $A_{756}G$ variant. Of note, the secondary structure was defined for a smaller RNA sequence where the I–II–VII junction and stem VII were not present.[3,4,34] (b) Simplified global architecture of the dimeric form reflecting the relative orientation of helical domains and showing the center of two-fold rotational symmetry. (c) Cartoon representation of the crystal structure of the $G_{638}A$ variant (PDB code 4R4V) shown in a simplified *trans* form with one entire protomer and the SLI substrate of the partner protomer. (d) Tertiary structure schematics (as in Figure 2) of the simplified *trans* form. Due to constraints inherent to this type of representation, the proximity of the G_{638} and A_{756} loops is not depicted here. (Adapted with permission from Ref 34. Copyright Macmillan Publishers Ltd: Nature Chemical Biology)

contrast, the I–II–VII junction directs helix VII almost perpendicular to the flat-shaped structure formed by the side-by-side packing of the other helical segments. The dimeric structure can be simplified by showing only the substrate of one protomer bound to the ribozyme of the other protomer (Figure 7(c) and (d)). Interestingly, this structure captures a closed state, in which the substrate is bound via both the I/V KLI and close association of its G_{638} loop with the A_{756} loop to form a precatalytic active site.

Important Structural Features Within Crystal Structure

The global architecture of the VS ribozyme is governed by local structural features that direct its folding. With the exception of the I–II–VII junction, important structural features were previously structurally characterized by NMR. Thus, it is interesting to compare these structural features present in the crystal structures with the NMR structures of isolated subdomains.

I/V KLI

The I/V KLI within the crystal structure is very similar to the one found in the NMR structure of this subdomain, with the U-turns of loop I and loop V directing the formation of three WC/WC base pairs and one SE/WC base pair between the two loops (Figure 8(a)). Loop V displays a compact U-turn structure with the U_{700} residue bulged out and the closing A_{701}–U_{695} base pair destabilized, as previously found in the NMR structure of the SLI/SLV complex (Figure 2(d)). For loop I, there is some discrepancy between the available structures (Figures 6 (d) and 7(d)). In the NMR structure, there is a disordered WC/WC C_{626}–G_{633} base pair that, in some models, forms a base triple with the Hoogsteen edge of the bulged out G_{634} (Figure 2(d)). In the X-ray structure, it is G_{634} instead of G_{633} that forms a WC/WC base pair with C_{626}, and G_{633} is involved in the H/SE G_{633}–G_{627} base pair. This structural discrepancy likely results from the fact that different SLI substrates were used for X-ray (C_{634}G variant) and NMR (C_{634}G/G_{627}A variant) studies, in agreement with the notion that both the sequence and structure can be varied in this region without significantly affecting cleavage activity.[4,22]

III–IV–V Junction

The III–IV–V junction within the crystal structure is very similar to the one found in the NMR structure of this subdomain, both adopting a family C topology with coaxial stacking of stems III and IV and close packing of stems III and V (Figure 8(b)). A U-turn fold defined by residues U_{710}, G_{711}, and A_{712} allows backbone reversal between stems III and V as observed in the NMR ensemble, with the two typical hydrogen bonds, the sharp turn in the phosphate backbone, and continuous stacking before and after the turn. As in the NMR structure, the splayed conformation between C_{687} and G_{688} allows backbone redirection between stems IV and V. The core of the junction is stabilized by a stack of four base triples and a ribose zipper that are very similar between the crystal and NMR structures; however, there are small variations in hydrogen bond interactions that leads to different definitions of base-pairing interactions (Figures 6 and 7).

II–III–VI Junction

The II–III–VI junction within the crystal structure is also very similar to the one found in the NMR structure of this subdomain, both adopting a family C topology with coaxial stacking of stems II and III as well as close packing of stems II and VI stabilized by a bulge–bulge interaction (Figure 8(c)). An ACAA turn defined by residues A_{764}, C_{765}, A_{766}, and A_{767} allows backbone reversal between stems II and VI and base stacking of residues before and after the turn. Although this ACAA turn is very similar in the NMR and X-ray structures, the NMR structure shows a slightly more compact fold with a hydrogen bond (C_{765} 2′-OH-A_{766} N7) not present in the X-ray structures. As observed in the NMR structure, the splayed conformation between G_{721} and G_{722} allows backbone redirection between stems III and VI, and the core of the junction is stabilized by the stack of two base triples and a WC/WC G–A base pairs. Nevertheless, there are small variations in the junction core between the crystal and NMR structures in terms of base planarity and hydrogen bond interactions (Figures 6 and 7). As in the NMR structure, a bulge–bulge interaction between stems II and VI involves A_{652} (stem II) stacking between C_{760} and A_{761} (stem VI) and forming a SE/WC A–A base pair with A_{725} (stem VI).

Despite the remarkable similarity between the structures of the II–III–VI junction core and bulge–bulge interaction, a notable difference is observed in stem III. Whereas A_{718} stacks on G_{719} within the helix in the NMR structure, it is completely bulged out in the crystal structure (Figure 8(c)). Thus, the transition of the VS ribozyme from an open state to a closed state may involve A_{718} shifting from an intrahelical position to a bulged-out state. Such conformational change, known to affect the rise and twist of flanking base pairs,[56] would change the relative orientation of helices connected to the III–IV–V and

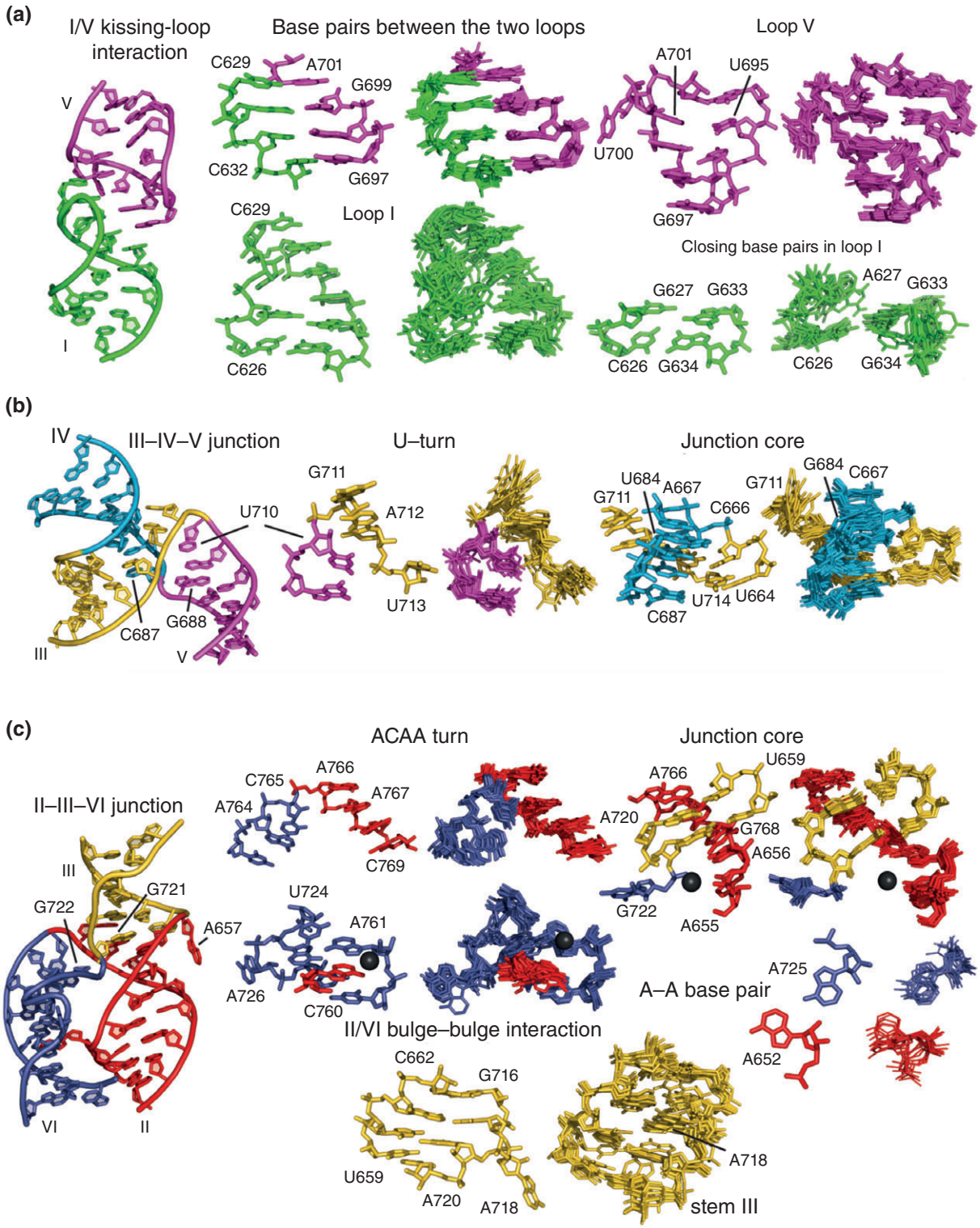


FIGURE 8 | Comparative analysis of the crystal and NMR structures in terms of the structural features that direct the folding of the VS ribozyme. A general view of the crystal structure (PDB code 4R4V; left) and specific characteristic features (right) are provided for (a) the I/V kissing-loop interaction (residues 623–637 and 692–704), (b) the III–IV–V junction (residues 661–670, 681–691, and 705–717), and (c) the II–III–VI junction (residues 649–662, 716–729, and 758–774). Each specific feature is displayed as a side-by-side comparison between the X-ray crystallography structure (PDB code 4R4V) on the left and a superposition of the 10 lowest-energy NMR structures [PDB codes 2MI0 (a), 2MTJ (b), and 2N3R (c)] on the right. Divalent metal ions that are common to both the NMR and X-ray structures are shown as black spheres.

II–III–VI junctions and may help bring the A_{638} loop in proximity of the A_{756} loop.

I–II–VII Junction

The I–II–VII junction is only represented in the crystal structures, where it adopts a family C topology with coaxial stacking of stems I and II and bending of stem VII toward stem I (Figure 7). This coaxial stacking of stems I and II helps orient the substrate helix outward from its own catalytic domain to interact with the catalytic domain of the partner protomer. A GCAA turn, involving G_{611}, C_{612}, A_{613}, and A_{614}, allows backbone reversal between C_{612} and A_{613} as well as base stacking of residues before and after the turn. In the core of the junction, a SE/WC A–A base pair piles up on a G–A–U base triple via cross-stand purine stacking. Residues A_{614}, U_{615}, and C_{616} adopt an S-turn fold that directs stacking of the U_{615} nucleobase on G_{778}. This stacking is likely critical to stabilize the splayed conformation of G_{778} and A_{779}, which respectively stack on stems II and VII. A superposition of the two crystal structures (not shown) reveal that residues G_{611}, C_{612}, U_{615}, and A_{779} are less well defined, suggesting that they are relatively more dynamic than other residues of the junction. This may be relevant to the possibility of the I–II–VII junction to adopt two conformations, one that favors cleavage and one that favors ligation, as previously suggested.[6]

Interactions with Divalent Metal Ions

NMR studies of VS ribozyme subdomains have established that Mg^{2+} ions are required for folding of the A_{756} loop and the II–III–VI junction and help stabilize the compact fold of free SLV and the III–IV–V junction. In addition, several divalent metal-binding sites have been identified in RNA subdomains using NMR methods (Figures 2, 4, and 5). Likewise, several metal ions have been localized from the two available crystal structures, 4R4P ($2Mg^{2+}$) and 4R4V ($9Mg^{2+}$ and $5K^{+}$). However, only two metal-binding sites have been identified by both NMR spectroscopy and X-ray crystallography methods. Interestingly, these are associated with the II–III–VI junction and are likely required for the folding of this junction. One is localized at the II–VI bulge–bulge interaction near the A_{652}–A_{761} stack, such that it can counteract the electronegative repulsion from nearby phosphates (C_{651}, A_{652}, A_{761}, and G_{762}; Figure 8(c)). The other is localized in the core of the II–III–VI junction where it may stabilize a phosphate cluster (G_{721}, A_{722}, G_{655}, and A_{656}), the splayed conformation of G_{721} and G_{722} and base interactions at the junction core. Surprisingly, the Mg^{2+} ion localized in the S-turn of the A_{756} loop by NMR methods (Figure 4) and important for folding of the free A_{756} loop has no equivalent in crystal structures of the VS ribozyme, even though the S-turn displays a similar phosphate cluster in these structures.

VS Ribozyme Active Site

Mechanistic investigations of the VS ribozyme have revealed two key nucleobases, G_{638} and A_{756}, responsible for general acid–base catalysis and highlighted similarities with the hairpin ribozyme, where a guanine (G_8) and an adenine (A_{38}) of two distinct internal loops also respectively serve as proton acceptor and proton donor in the cleavage reaction.[13] NMR studies of individual subdomains encompassing the G_{638} and A_{756} loops of the VS ribozyme have revealed important structural features in these loops and similarities with the hairpin ribozyme active site.[18,25,26] However, the two crystal structures of the VS ribozyme provide a detailed view of the active site architecture. They also clearly establish that the VS and hairpin ribozymes display comparable structural features at the active site despite low similarity in their sequences and tertiary structures, possibly as a result of convergent evolution.[13,14,34]

Remodeling of G_{638} and A_{756} Loops for Formation of Active Site

The NMR structures of the G_{638} and A_{756} loops were superposed to those of the crystal structures to provide insights into the local structural changes associated with formation of the active site (Figure 9 (a)). For the A_{756} loop, the main structural characteristics of the free loop are found in the active site, namely the WC/WC G–A base pair and the S-turn. However, formation of the active site leads to a more pronounced S-turn and brings the bulged-out residues (C_{755} and A_{756}) from their minor groove position to a completely exposed position, away from the main body of the helix. Similarly, the largest structural change in the G_{638} loop upon formation of the active site is the protrusion of two adenines (A_{621} and A_{622}) that were both positioned in the free G_{638} loop to form a G–A base pair with G_{638}. This structural change is associated with formation of an S-turn centered at A_{621} that brings the G_{620} 2′-OH closer to the general base G_{638} (from ~6.1 to 5.4 Å). Thus, the A_{756} and G_{638} internal loops are prefolded in their free form, and relatively minor conformational changes allow their less constrained residues to become completely protruded to form the active site. This is in drastic contrast with the more significant

structural rearrangements that occur when the two interacting loops of the hairpin ribozyme associate to form the active site. The NMR structures of the isolated internal loops of the hairpin ribozyme[57,58] were previously compared with the crystal structure of a hairpin ribozyme–noncleavable substrate complex in which the two loops intimately associate,[59] revealing important changes in base-pairing interactions and backbone geometry upon formation of the active site.

Active Site Architecture

As originally envisioned, the G_{638} loop is nested in a cleft formed by the helical domains II and VI,[41] and this involves extensive interloop stacking between the G_{638} and A_{756} loops as well as interactions of these loops with the minor groove of stem II (Figure 9(b)). Starting with the protruded adenines of the G_{638} loop, A_{622} stacks with the bulged out A_{657} from the II–III–VI junction and A_{621} forms an A-minor interaction with helix II and stacks with A_{756} of stem VI. The protruded residues of the A_{756} loop, C_{755} and A_{756}, are stacked on one another and sandwiched between A_{621} and A_{639} from SLI. These stacking interactions are consistent with the observation of a UV-inducible crosslink between A_{756} and a 4-thiouridine at position 621.[53] As observed in the hairpin ribozyme, the residues flanking the scissile phosphate [N – 1 (G_{620}) and N + 1 (A_{621})] adopt a splayed conformation that allow stacking of the key catalytic nucleobases (G_{638} and A_{756}) and position their reactive groups near the scissile phosphate (Figure 9(c)). The similarity with the hairpin ribozyme extends further as the general acid (A_{756}) is protruded from the internal loop by an S-turn structural element and the N – 1 residue forms a SE/H base-pair with the residue immediately 3′ of the general base (A_{639}). In the proposed cleavage mechanism, a proton transfer occurs between the G_{620} 2′-OH group and the deprotonated N1 group of the G_{638} nucleobase, followed by a nucleophile attack of

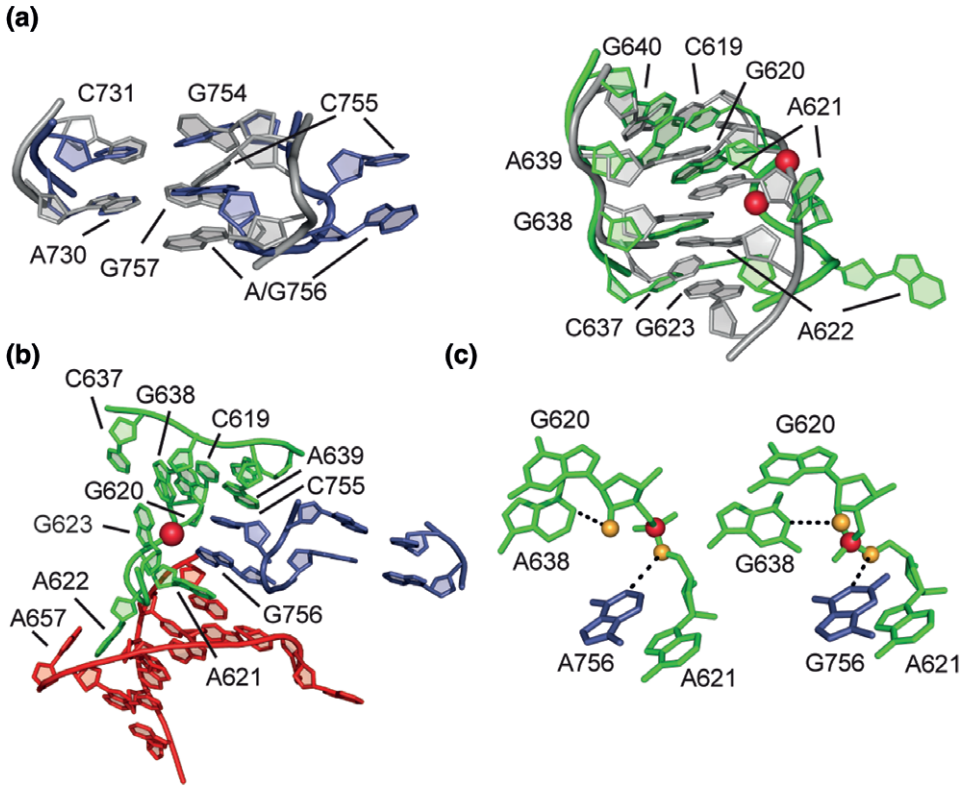


FIGURE 9 | Formation of the active site in the VS ribozyme. (a) Remodeling of the A_{756} loop (left) and G_{638} loop (right) upon formation of the active site as seen from the superposition of the lowest-energy NMR structures of the free loops [in gray; PDB codes 2L5Z (left) and 1OW9 (right)] onto those forming the active site in the crystal structure (in color; PDB code 4R4P). (b) Local environment at the active site (PDB code 4R4P). (c) Key residues at the active site of the G_{638}A variant (left; PDB code 4R4V) and the A_{756}G variant (right; PDB code 4R4P). The representation highlights the atoms that are expected to adopt an inline geometry in the transition state, namely the 2′-oxygen (gold sphere), scissile P (red), and 5′-oxygen (gold sphere) that rather form angles of 97° (left) and 129° (right).[34] The N1 atoms of G_{638} and A_{638} are respectively 5.7 Å (left) and 4.9 Å (right) away from the 2′-oxygen nucleophile (dashed line), whereas the N1 atoms of A_{756} and G_{756} are respectively 4.2 Å (left) and 4.0 Å (right) away from the 5′-oxygen leaving group (dashed line).[34] (Adapted with permission from Ref 34. Macmillan Publishers Ltd: Nature Chemical Biology)

the 2′-oxygen on the scissile phosphate to form a trigonal bipyramidal transition state, which is resolved by protonation of the 5′-oxygen leaving group by the $N1^{+}$-H group of the A_{756} nucleobase.[35] In many nucleolytic ribozymes, the splayed conformation of the cleavage site nucleotides is associated with the in-line geometry of reacting groups (2′-oxygen, P, and 5′-oxygen) and is required in the transition state; however, such in-line geometry was not observed in the available crystal structures of the VS ribozyme (Figure 9(c)). Also, the key nucleobases are not within hydrogen bond distances of their associated reacting groups (Figure 9(c)). Nevertheless, the crystal structures are in general agreement with the respective roles of G_{638} and A_{756} as the general base and acid in the cleavage reaction since only minor structural rearrangement may be required to bring the active site in a fully functional state.

Shifted pK_as of Two Key Nucleobases in Light of Available Structures

In light of the available NMR and X-ray structures of the VS ribozyme, it is interesting to consider the shifted pK_as of 5.2–5.8 and 8.3–8.4 previously assigned to A_{756} and G_{638}, respectively, based on the pH dependence of the VS ribozyme cleavage reaction.[35–38] From NMR studies of the isolated SLVI subdomain, a pK_a value of 4.4 was attributed to A_{756}, in agreement with its nucleobase being exposed in the open state conformation of the VS ribozyme.[25] The increase of its pK_a value to 5.2–5.8 could result from its proximity to the scissile phosphate in the active site. In contrast, the lowering of the pK_a value of G_{638} by approximately one log unit compared to 5′-GMP is more intriguing and suggests the presence of a metal ion at the active site. Modulation of the pK_a values of the two key nucleobases upon changes in the ionic environment also suggests the presence of a metal ion at the active site.[38] Although no metal ion is observed at the active site in the available X-ray structures, NMR studies have revealed divalent metal ion binding sites in both the A_{756} and G_{638} loops that may remain bound in the active site, namely the metal ion at the S-turn of the A_{756} loop (Figure 4) and the one associated with the 5′-A_{639}–G_{640}-3′ sequence of the G_{638} loop (Figure 2(b)). Alternatively, the stabilization of an S-turn in the G_{638} loop upon formation of the active site (Figure 9(a); near A_{621}) may create a favorable environment for binding of a metal ion that could stabilize the transition state. Higher resolution crystal structures are clearly needed to provide a more detailed characterization of Mg^{2+} ion binding sites in the VS ribozyme.

CONCLUSION

Several high-resolution NMR and X-ray studies reported in the past few years have provided novel insights into the dynamic structure of the *Neurospora* VS ribozyme. More specifically, they have allowed us to better understand: (1) the global architecture of the VS ribozyme, (2) the specific structural features that direct folding, (3) important features of the active site that direct catalysis, (4) the role of magnesium ions, and (5) the conformational changes associated with substrate recognition, activation, and docking. Several hotspots for dynamics have been identified that are likely important for the cleavage and ligation activities of this ribozyme. Follow-up investigations are needed to capture a fully active ribozyme with an in-line conformation at the active site and better understand the role of metal ions and dynamics in the reaction. Together with available structural data, such information will support current efforts aimed at engineering the VS ribozyme to cleave alternate substrates.[52,60]

ACKNOWLEDGMENTS

The authors thank Dr. Richard A. Collins for critical reading of the manuscript. This work was supported by a grant from the Canadian Institutes of Health Research (CIHR 148801 to P.L.).

REFERENCES

1. Saville BJ, Collins RA. A site-specific self-cleavage reaction performed by a novel RNA in *Neurospora* mitochondria. *Cell* 1990, 61:685–696.
2. Saville BL, Collins RA. RNA-mediated ligation of self-cleavage products of a *Neurospora* mitochondrial plasmid transcript. *Proc Natl Acad Sci USA* 1991, 88:8826–8830.
3. Beattie TL, Olive JE, Collins RA. A secondary-structure model for the self-cleaving region of *Neurospora* VS RNA. *Proc Natl Acad Sci USA* 1995, 92:4686–4690.

4. Andersen A, Collins RA. Rearrangement of a stable RNA secondary structure during VS ribozyme catalysis. *Mol Cell* 2000, 5:469–478.
5. Guo HCT, De Abreu DM, Tillier ERM, Saville BJ, Olive JE, Collins RA. Nucleotide sequence requirements for self-cleavage of *Neurospora* VS RNA. *J Mol Biol* 1993, 232:351–361.
6. Jones FD, Ryder SP, Strobel SA. An efficient ligation reaction promoted by a Varkud satellite ribozyme with extended 5′- an 3′-termini. *Nucleic Acids Res* 2001, 29:5115–5120.
7. Ouellet J, Byrne M, Lilley DM. Formation of an active site in trans by interaction of two complete Varkud Satellite ribozymes. *RNA* 2009, 15:1822–1826.
8. Lilley DMJ. The hairpin and Varkud satellite ribozymes. In: Lilley DMJ, Eckstein F, eds. *Ribozymes and RNA Catalysis*. Cambridge: Royal Society of Chemistry; 2008, 66–91.
9. Cochrane JC, Strobel SA. Catalytic strategies of self-cleaving ribozymes. *Acc Chem Res* 2008, 41:1027–1035.
10. Serganov A, Patel DJ. Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nat Rev Genet* 2007, 8:776–790.
11. Collins RA. The *Neurospora* Varkud satellite ribozyme. *Biochem Soc Trans Rev* 2002, 30:1122–1126.
12. Lilley DM. The Varkud satellite ribozyme. *RNA* 2004, 10:151–158.
13. Lilley DM. Catalysis by the nucleolytic ribozymes. *Biochem Soc Trans* 2011, 39:641–646.
14. Wilson TJ, Lilley DM. 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, 17:213–221.
15. Lilley DM. Mechanisms of RNA catalysis. *Philos Trans R Soc Lond B Biol Sci* 2011, 366:2910–2917.
16. Michiels PJ, Schouten CHJ, Hilbers CW, Heus HA. Structure of the ribozyme substrate hairpin of *Neurospora* VS RNA: a close look at the cleavage site. *RNA* 2000, 6:1821–1832.
17. Flinders J, Dieckmann T. A pH controlled conformational switch in the cleavage site of the VS ribozyme substrate RNA. *J Mol Biol* 2001, 308:665–679.
18. Hoffmann B, Mitchell GT, Gendron P, Major F, Andersen AA, Collins RA, Legault P. NMR structure of the active conformation of the Varkud satellite ribozyme cleavage site. *Proc Natl Acad Sci USA* 2003, 100:7003–7008.
19. Campbell DO, Legault P. NMR structure of the Varkud satellite ribozyme stem-loop V RNA and magnesium-ion binding from chemical-shift mapping. *Biochemistry* 2005, 44:4157–4170.
20. Campbell DO, Bouchard P, Desjardins G, Legault P. NMR structure of Varkud satellite ribozyme stem-loop V in the presence of magnesium ions and localization of metal-binding sites. *Biochemistry* 2006, 45:10591–10605.
21. Bergonzo C, Hall KB, Cheatham TE 3rd.. Stem-loop V of Varkud satellite RNA exhibits characteristics of the Mg(2+) bound structure in the presence of monovalent ions. *J Phys Chem B* 2015, 119:12355–12364.
22. Bouchard P, Lacroix-Labonté J, Desjardins G, Lampron P, Lisi V, Lemieux S, Major F, Legault P. Role of SLV in SLI substrate recognition by the *Neurospora* VS ribozyme. *RNA* 2008, 14:736–748.
23. Bouchard P, Legault P. Structural insights into substrate recognition by the *Neurospora* Varkud satellite ribozyme: importance of U-turns at the kissing-loop junction. *Biochemistry* 2014, 53:258–269.
24. Flinders J, Dieckmann T. The solution structure of the VS ribozyme active site loop reveals a dynamic "hot-spot". *J Mol Biol* 2004, 341:935–949.
25. Desjardins G, Bonneau E, Girard N, Boisbouvier J, Legault P. NMR structure of the A730 loop of the *Neurospora* VS ribozyme: insights into the formation of the active site. *Nucleic Acids Res* 2011, 39:4427–4437.
26. Bonneau E, Legault P. NMR localization of divalent cations at the active site of the *Neurospora* VS ribozyme provides insights into RNA-metal-ion interactions. *Biochemistry* 2014, 53:579–590.
27. Lafontaine DA, Norman DG, Lilley DM. The global structure of the VS ribozyme. *EMBO J* 2002, 21:2461–2471.
28. Bonneau E, Legault P. 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, 53:6264–6275.
29. Lafontaine DA, Norman DG, Lilley DM. Structure, folding and activity of the VS ribozyme: importance of the 2-3-6 helical junction. *EMBO J* 2001, 20:1415–1424.
30. Bonneau E, Girard N, Lemieux S, Legault P. The NMR structure of the II-III-VI three-way junction from the *Neurospora* VS ribozyme reveals a critical tertiary interaction and provides new insights into the global ribozyme structure. *RNA* 2015, 21:1621–1632.
31. Hiley SL, Collins RA. Rapid formation of a solvent-inaccessible core in the *Neurospora* Varkud satellite ribozyme. *EMBO J* 2001, 20:5461–5469.
32. Pereira MJ, Nikolova EN, Hiley SL, Jaikaran D, Collins RA, Walter NG. Single VS ribozyme molecules reveal dynamic and hierarchical folding toward catalysis. *J Mol Biol* 2008, 382:496–509.
33. Lipfert J, Ouellet J, Norman DG, Doniach S, Lilley DM. The complete VS ribozyme in solution studied by small-angle X-ray scattering. *Structure* 2008, 16:1357–1367.

34. Suslov NB, DasGupta S, Huang H, Fuller JR, Lilley DM, Rice PA, Piccirilli JA. Crystal structure of the Varkud satellite ribozyme. *Nat Chem Biol* 2015, 11:840–846.
35. Wilson TJ, Li NS, Lu J, Frederiksen JK, Piccirilli JA, Lilley DM. Nucleobase-mediated general acid–base catalysis in the Varkud satellite ribozyme. *Proc Natl Acad Sci USA* 2010, 107:11751–11756.
36. Smith MD, Collins RA. Evidence for proton transfer in the rate-limiting step of a fast-cleaving Varkud satellite ribozyme. *Proc Natl Acad Sci USA* 2007, 104:5818–5823.
37. Wilson TJ, McLeod AC, Lilley DM. A guanine nucleobase important for catalysis by the VS ribozyme. *EMBO J* 2007, 26:2489–2500.
38. Smith MD, Mehdizadeh R, Olive JE, Collins RA. The ionic environment determines ribozyme cleavage rate by modulation of nucleobase pK_a. *RNA* 2008, 14:1942–1949.
39. Rastogi T, Beattie TL, Olive JE, Collins RA. A long-range pseudoknot is required for activity of the *Neurospora* VS ribozyme. *EMBO J* 1996, 15:2820–2825.
40. Andersen AA, Collins RA. Intramolecular secondary structure rearrangement by the kissing interaction of the *Neurospora* VS ribozyme. *Proc Natl Acad Sci USA* 2001, 98:7730–7735.
41. Sood VD, Collins RA. Identification of the catalytic subdomain of the VS ribozyme and evidence for remarkable sequence tolerance in the active site loop. *J Mol Biol* 2002, 320:443–454.
42. Rastogi T, Collins RA. Smaller, faster ribozymes reveal the catalytic core of *Neurospora* VS RNA. *J Mol Biol* 1998, 277:215–224.
43. Beattie TL, Collins RA. Identification of functional domains in the self-cleaving *Neurospora* VS ribozyme using damage selection. *J Mol Biol* 1997, 267:830–840.
44. Zamel R, Collins RA. Rearrangement of substrate secondary structure facilitates binding to the *Neurospora* VS ribozyme. *J Mol Biol* 2002, 324:903–915.
45. Zamel R, Poon A, Jaikaran D, Andersen A, Olive J, De Abreu D, Collins RA. Exceptionally fast self-cleavage by a *Neurospora* Varkud satellite ribozyme. *Proc Natl Acad Sci USA* 2004, 101:1467–1472.
46. Leontis NB, Westhof E. Geometric nomenclature and classification of RNA base pairs. *RNA* 2001, 7:499–512.
47. Abu Almakarem AS, Petrov AI, Stombaugh J, Zirbel CL, Leontis NB. Comprehensive survey and geometric classification of base triples in RNA structures. *Nucleic Acids Res* 2012, 40:1407–1423.
48. Bouchard P, Legault P. A remarkably stable kissing-loop interaction defines substrate recognition by the *Neurospora* VS Ribozyme. *RNA* 2014, 20:1451–1464.
49. Franch T, Petersen M, Wagner EGH, Jacobsen JP, Gerdes K. Antisense RNA regulation in prokaryotes: rapid RNA/RNA interaction facilitated by a general U-turn loop structure. *J Mol Biol* 1999, 294:1115–1125.
50. de la Pena M, Dufour D, Gallego J. Three-way RNA junctions with remote tertiary contacts: a recurrent and highly versatile fold. *RNA* 2009, 15:1949–1964.
51. Lafontaine DA, Wilson TJ, Norman DG, Lilley DM. The A730 loop is an important component of the active site of the VS ribozyme. *J Mol Biol* 2001, 312:663–674.
52. Lacroix-Labonté J, Girard N, Lemieux S, Legault P. Helix-length compensation studies reveal the adaptability of the VS ribozyme architecture. *Nucleic Acids Res* 2012, 40:2284–2293.
53. Hiley SL, Sood VD, Fan J, Collins RA. 4-Thio-U cross-linking identifies the active site of the VS ribozyme. *EMBO J* 2002, 21:4691–4698.
54. Koldobskaya Y, Duguid EM, Shechner DM, Suslov NB, Ye J, Sidhu SS, Bartel DP, Koide S, Kossiakoff AA, Piccirilli JA. A portable RNA sequence whose recognition by a synthetic antibody facilitates structural determination. *Nat Struct Mol Biol* 2011, 18:100–106.
55. Poon AH, Olive JE, McLaren M, Collins RA. Identification of separate structural features that affect rate and cation concentration dependence of self-cleavage by the *Neurospora* VS ribozyme. *Biochemistry* 2006, 45:13394–13400.
56. Barthel A, Zacharias M. Conformational transitions in RNA single uridine and adenosine bulge structures: a molecular dynamics free energy simulation study. *Biophys J* 2006, 90:2450–2462.
57. Cai Z, Tinoco IJ. Solution structure of loop A from the hairpin ribozyme from Tobacco ringspot virus satellite. *Biochemistry* 1996, 35:6026–6036.
58. Butcher SE, Allain FH, Feigon J. Solution structure of the loop B domain from the hairpin ribozyme. *Nat Struct Biol* 1999, 6:212–216.
59. Rupert PB, Ferré-D'Amaré AR. Crystal structure of a haripin ribozyme-inhibitor complex with implications for catalysis. *Nature* 2001, 410:780–786.
60. Lacroix-Labonté J, Girard N, Dagenais P, Legault P. Rational engineering of the *Neurospora* VS ribozyme to allow substrate recognition via different kissing-loop interactions. *Nucleic Acids Res* 2016, 44:6924–6934.

