

UC San Diego

UC San Diego Previously Published Works

Title

Identification of a GUAAY Pentaloop Sequence Involved in a Novel RNA Loop-Helix Interaction

Permalink

<https://escholarship.org/uc/item/4jh803vq>

Journal

Journal of Molecular Biology, 428(24)

ISSN

0022-2836

Authors

Chan, Russell T
Keating, Kevin S
Go, Michaela C
[et al.](#)

Publication Date

2016-12-01

DOI

10.1016/j.jmb.2016.10.015

Peer reviewed



HHS Public Access

Author manuscript

J Mol Biol. Author manuscript; available in PMC 2017 December 04.

Published in final edited form as:

J Mol Biol. 2016 December 04; 428(24 Pt B): 4882–4889. doi:10.1016/j.jmb.2016.10.015.

Identification of a GUAAY pentaloop sequence involved in a novel RNA loop-helix interaction

Russell T. Chan¹, Kevin S. Keating², Michaela C. Go¹, and Navtej Toor¹

¹Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA

²Schrödinger, LLC, New York, New York

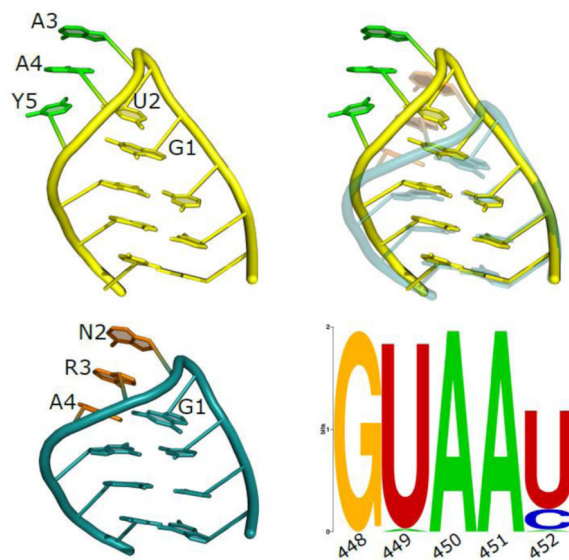
Abstract

Large RNAs often utilize GNRA tetraloops as structural elements to stabilize the overall tertiary fold. These tetraloop-receptor (TR) interactions have a conserved geometry in which the tetraloop docks into the receptor at an angle of ~15 degrees from the helix containing the receptor. Here we show that the conserved GUAAY pentaloop found in domain III of group IIB1 introns participates in a novel class of RNA tertiary interaction with a geometry and mode of binding that is significantly different from that found in GNRA TR interactions. This pentaloop is highly conserved within the IIB1 class and interacts with the minor groove of the catalytic DV. The base planes of the loop and receptor nucleotides are not coplanar and greatly deviate from standard A-minor motifs. The helical axis of the GUAAY stem loop diverges ~70° from the angle of insertion found in a typical GNRA TR interaction. Therefore, the loop architecture and insertion orientation is distinctive, with *in vitro* splicing data indicating that a GNRA tetraloop is incompatible at this position. The GUAAY pentaloop-receptor motif is also found in the structure of the eukaryotic thiamine pyrophosphate riboswitch in the context of a hexanucleotide loop sequence. We therefore propose, based on phylogenetic, structural, and biochemical data, that the GUAAY pentaloop-receptor interaction represents a novel structural motif that is present in multiple structured RNAs.

Graphical abstract

Corresponding should be addressed to N.T. (ntoor@ucsd.edu).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Introduction

One of the most ubiquitous RNA structures is the tetraloop [1, 2], a four nucleotide sequence that caps an A-form double helix. While numerous tetraloop motifs have been identified, three are the most prevalent: CUUG [3], UNCG [4], and GNRA [5], with the latter demonstrating a robust ability to form RNA tertiary structures. GNRA tetraloops can engage in tetraloop-receptor interactions, where the loop nucleotides insert into the minor groove of an RNA receptor helix. In addition, different GNRA tetraloop sequences have varying affinities for receptor sequences, allowing this tertiary structure motif to be sequence specific [6, 7]. GNRA tetraloop-receptor interactions have a defined geometry between the two interacting motifs, with the tetraloop docking into the minor groove of the receptor helix at an angle of $\sim 15^\circ$. Furthermore, the nucleobases of the tetraloop and the receptor nucleotides are coplanar. Despite many large RNA structures being determined over the past two decades, there are only two identified classes of tetraloop-receptor interactions (along with the GANC tetraloop [8, 9]).

Previously, we determined the crystal structure of a eukaryotic group II intron from the brown alga *Pylaiella littoralis* (*P.li.LSUI2*) in which we described the conformation of the μ - μ' interaction (PDB 4R0D) [10]. This interaction was first biochemically identified through nucleotide analogue interference mapping (NAIM) and involves a GUAAY pentaloop docking into the base of the catalytic domain V [11] (Figure 1a). This μ - μ' interaction occurs directly adjacent to the κ - κ' interaction in the group IIB intron crystal structure [10] and results in a quintuple adenosine base stack inserting into the minor groove (Figure 1b). Here we present phylogenetic, structural, and biochemical evidence that distinguishes the GUAAY pentaloop from the ubiquitous GNRA tetraloop and reveals a new mode of interaction between an RNA loop and a receptor helix.

Results

The consensus secondary structure of the group IIB1 phylogenetic subclass reveals that nucleotides 1 to 4 of the pentaloop sequence are highly conserved as GUAA [12] (Table 1). At the fifth position, a pyrimidine residue is always observed, with a U and C representing 71% and 19% of the sequences, thus forming the GUAAY consensus sequence. Additionally, the stem containing the loop contains a closing conserved U-G wobble pair (85%), but there are no observed occurrences of the G-U wobble pair. This differs from GNRA tetraloops, which tend to have a preference for C-G base pairs at this position, but can also contain a closing G-C pair [13, 14]. In addition to the sequence conservation, the loop is attached to the domain IIIa helix, which has a highly conserved length of seven base-pairs in group IIB1 introns (Figure 1a).

The first nucleotide of the pentaloop (G448) stacks directly on top of the capping U-G wobble pair (Figure 2a). Interestingly, instead of stacking on the 5' adjacent base U447, G448 instead stacks almost entirely on G453, which is paired with U447. This may explain the preference for a U-G over G-U wobble pair, as the latter would orient the loop further from its receptor interaction. The second pentaloop nucleotide (U449) continues the base stack above the first nucleotide. The Watson-Crick edges of these two nucleotides are oriented in a similar direction, which keeps the phosphate backbone in an A-form helix-like conformation. U449 stacks on G448 and forms a base-phosphate interaction with C452 using its Watson-Crick edge. In this orientation, U449 (N3) hydrogen bonds with the non-bridging oxygen of the phosphate of C452. This type of base-phosphate interaction is typical of a U-turn, which was first observed in tRNA anti-codon loops [15]. Substitution of cytosine for U449 could only be maintained if N3 were protonated, while a purine substitution would result in a steric clash or force the backbone into a sub-optimal configuration. Altogether, the configuration of the second nucleotide in context within the overall loop structure provides a rationale for its conservation. The third loop nucleotide (A450) initiates the U-turn [15] and causes its Watson-Crick edge to deviate from an A-form helix rotation. This contrasts with the GNRA tetraloop, where the U-turn is initiated between nucleotides one and two (GNRA) [16]. The fourth pentaloop nucleotide (A451) stacks beneath the previous nucleotide and also inserts its nucleobase into the minor groove of DV. Finally, the last nucleotide of the pentaloop (C542) finishes the base stack below A451. It makes a single hydrogen bond between O2 and N2 of G555 from the receptor. Because O2 is present in both uracil and cytosine, this provides a rationale for the conservation of a pyrimidine at this position.

There is some commonality between the GNRA tetraloop and the GUAAY pentaloop. Both loop structures allow the last three nucleobases to splay out in a manner that allows them to interact with a receptor (Figure 2b). However, the base planes of the loops represent a significant difference between the two systems. In the GNRA tetraloop, the first and last nucleotide are on the same plane and form a trans-Watson-Crick/Hoogsteen pair between the guanosine sugar edge and the Hoogsteen edge of the adenosine. In contrast, the guanosine of the pentaloop is the only base in its plane. Rather, the second nucleotide (U449) is located on the same base plane as the final nucleotide (GUAAY). Furthermore, the second and last nucleobase are not interacting with each other. Instead, the second nucleobase interacts with

the phosphate backbone and Watson-Crick edge of the last nucleotide is oriented in the opposite direction. Because the first and second nucleotide stack with one another, this allows the pentaloop to extend one base plane higher, with respect to the GNRA tetraloop (Figure 2a).

The last three nucleotides of the GNRA and GUAAY loop are positioned so that their nucleobases are ordered and oriented outward from the helix. This allows the nucleobases to insert into their respective receptors, but the orientation of insertion differs greatly. The inserting nucleotides of the GNRA tetraloop are approximately planar with respect to the receptor base planes (Figure 2b). In contrast, the GUAAY nucleotides insert into the minor groove of DV $\sim 70^\circ$ relative to the receptor nucleotides. While this orientation precludes the formation of canonical A-minor motifs, it allows the possibility of more nucleotides to stack above and below without clashing with the phosphate backbone.

A structural motif search was done to determine if similar loop structures existed in the PDB that were not yet annotated. Using the AMIGOS II program [17], we initiated an (η, θ) search, with the pentaloop values corresponding to (182.1, 218.3), (175.2, 256.4), (18.7, 231.2), (162.9, 223.6), and (167.8, 273.0). Based on this search, no pentaloop match was identified. However, the RNA 3D Motif Atlas [18, 19] revealed a hexanucleotide loop containing the sequence GUAAU from the eukaryotic thiamine pyrophosphate riboswitch (PDB 3D2V and 2CKY) that docks into a receptor helix in a similar manner as seen in the group IIB1 intron (Figure 2c). In this case, there is no analogous κ - κ' interaction found near the inserting base stack of the GUAAY sequence, therefore suggesting that this novel loop-receptor combination can exist as an independent functional unit. In addition, the presence of the GUAAY sequence in a hexaloop indicates that it can function in a variety of sequence contexts in a manner similar to that seen for GNRA tetraloops [20].

The initial discovery of μ - μ' indicated a strict requirement for the loop to be a pentaloop [11]. Due to the lack of structural information at the time, only two tetraloop substitutions were conducted: mutating the entire loop to UUCG (μ), and deletion of the last loop nucleotide to a GUAA sequence. With the information obtained from the crystal structure of the group IIB1 intron [10], the latter loop mutation actually results in a significant change of sequence register of the loop structure. Specifically, shortening the loop from five to four nucleotides results in the substitution of a uracil at a position that usually involves an adenosine inserting into the receptor. In the same study, a point mutation at that position from an adenosine to uracil resulted in defects in splicing activity. Utilizing the structural data as a guide, we aimed to characterize the sequence and spatial requirements that would promote the formation of μ - μ' through mutagenesis (Figure 3). While the *in vitro* splicing assay is an indirect measure for the formation of μ - μ' , the only known function of this tertiary interaction is to stabilize the active conformation of the active site. We are therefore using splicing as an indication for the formation of the μ - μ' interaction. Magnesium is an essential cation for ribozymes, both as a structural element [21-23] to stabilize the negative phosphate backbone as well as for catalytic function [24, 25]. In the case of the group II intron, two catalytic metals are bound in the active site [26, 27]. One serves to activate the nucleophile for attack at the scissile phosphate, while the other stabilizes the transition state of the leaving group. In standard splicing conditions for *P.li.LSUI2* [28], 10 mM Mg^{2+} is

used to stimulate splicing *in vitro*. However, at this concentration we observed no significant difference in splicing activity for any of the D3a mutants (Figure 4). This led us to believe that this magnesium concentration is capable of overcoming structural defects caused by mutagenesis, as seen in Figure 4 (orange bars). This is especially apparent in the μ mutant, which should be incapable of forming the loop-receptor interaction (Figure 4b). To address this, the Mg^{2+} concentration was lowered to 2.5mM. However, it is still useful to test the splicing at 10 mM Mg^{2+} in order to validate the retention of activity in the mutants assayed in this study.

The high level of conservation of the GUAAY pentaloop is remarkable for group IIB1 introns, with the first four positions being completely conserved across all the sequences. Despite this conservation, mutants with single base substitutions seemed to retain splicing activity comparable to wild type (Figure 4a). By comparison, mutating the loop entirely to adenosines leads to ~3-fold difference. This suggests that there is some level of redundancy with related sequences having a capacity to form a similar three-dimensional structure as the wild-type loop.

While the previous study deleted the terminal pyrimidine to test the effects of a GUAA-tetraloop sequence, a GAAA sequence was also tested. The latter allows for a canonical GNRA-tetraloop fold to still maintain an adenosine at the very tip of the loop to potentially interact with the DV minor groove. Both of these mutants result in ~3-fold reduction in splicing activity (Figure 4b), consistent with the findings of Fedorova *et al.* (2005). This still does not rule out the possibility that a tetraloop is incompatible at this position, because a tetraloop substitution would effectively lower the loop by one base plane (*vide supra*). Different helical lengths were tested (-2 to +2) to favor docking of this tetraloop and affect splicing activity. Increasing the helix by 1 base pair had a modest reduction on the effect of splicing, however the other helical changes (-2, -1, +2) resulted in significant defects in splicing (Figure 4c). Therefore, the addition of a GAAA-loop with varying helical lengths (-2 to +3) does not rescue splicing activity, rather it exacerbated the defect further to levels comparable to μ (Figure 4d).

Discussion

The GUAAY pentaloop-receptor interaction found in group IIB1 introns and the eukaryotic thiamine pyrophosphate riboswitch represents a new class of loop-receptor interactions in addition to the known GNRA and GANC [8, 9] tetraloop-receptors. Consistent with previous studies, the GUAAY loop cannot be replaced with a GNRA tetraloop. In addition, the defect caused by a tetraloop substitution cannot be rescued by changing the length of the attached helix. The helix attached to the μ -loop emanates from a highly conserved four-way junction, with two other tertiary interactions in close proximity to this junction. It could be possible to adapt a GNRA-tetraloop at this position, but would likely require the remodeling of the aforementioned junction. It is probable, however, that a GNRA tetraloop is completely incompatible in this position. Due to the close proximity of κ - κ' , the receptor must be able to accommodate five adenosines in the minor groove. Rather than inserting planar with respect to the receptor nucleotides, the loop nucleotides insert at an unusual angle that allows the base stacking vector to be parallel to the direction of the minor groove. This

orientation allows five, and potentially more, nucleotides to base stack effectively into the minor groove without clashing with the surrounding phosphate backbone. Additionally, there may be other stable variants of the pentaloop, similar to the GUAAY, but unable to accommodate splicing activity in the specific context of the μ - μ' interaction. It should be added that this new class of loop-helix interaction seems to tolerate insertions as seen in the hexanucleotide sequence of the loop from the thiamine pyrophosphate riboswitch. GNRA tetraloops behave in a similar manner with additional inserted nucleotides in the form GNR[X_n]A [7, 29].

Materials and Methods

IIB1 intron sequences were obtained from the Group II Intron Database [30]. Alignment was done using Bioedit [31]. Sequence logos were generated using the WebLogo [32]. AMIGOS II [17] motif search was done on the following data set: PDBs accessed on September 2015, with a minimum of 30 RNA nucleotides, and a resolution greater than 3.5Å.

The constructs used for the *in vitro* self-splicing assays contained wild-type *P.lj.LSUI2* sequence with a 250-nt 5' exon and 75-nt 3' exon. The domain IV sequence in this construct was deleted and replaced with a UUCG stem loop. This construct was inserted into the pUC57 plasmid at the EcoRV site. Overlapping PCR was used to generate the mutant constructs and inserted into pUC57 with EcoRV and T4 DNA ligase. Plasmid was linearized using HindIII and used for *in vitro* transcription with T7 RNA polymerase. Radiolabeled transcripts were prepared using 10 μ Ci [α -³²P] UTP (3000 Ci/mmol), 0.5 mM UTP, 1 mM other NTPs, and 2.5mM MgCl₂. Transcripts were gel purified on a 4% polyacrylamide (19:1)/8 M urea gel, RNA was recovered by diffusion into 300 mM NaCl, 0.01% SDS, 1 mM EDTA. Self-splicing experiments were performed for 30 min at 45°C in a splicing buffer containing, 1 M NH₄Cl, 40 mM Tris-HCl (pH 7.5), and 0.02% SDS, and either 2.5 mM or 10 mM MgCl₂. Reactions were stopped by addition of 2.5 volumes of 95% EtOH and 0.5 volume of 3M sodium acetate pH 5.2. This mixture was then incubated at -80°C for 30 mins to precipitate the RNA followed centrifugation. Supernatant was removed and pellets resuspended in 5 μ l of 0.5 \times TE, 50% formamide. Splicing products were resolved using a denaturing 4% polyacrylamide (19:1)/8 M urea gels. The Bio-Rad PMI Imager was used to image and quantitate gels. All splicing assays were done in triplicate.

Acknowledgements

We thank Timothy Wiryaman for comments on the manuscript. R.T.C. was supported by the Cell, Molecular, and Genetics (CMG) Training Program funded by NIH predoctoral training grant 5T32GM007240. This work was supported by NIH grant 5R01GM102216 awarded to N.T.

References

- [1]. Cate JH, Gooding AR, Podell E, Zhou K, Golden BL, Kundrot CE, et al. Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science*. 1996; 273:1678–85. [PubMed: 8781224]
- [2]. Woese CR, Winker S, Gutell RR. Architecture of ribosomal RNA: constraints on the sequence of “tetra-loops”. *Proc Natl Acad Sci U S A*. 1990; 87:8467–71. [PubMed: 2236056]
- [3]. Jucker FM, Pardi A. Solution structure of the CUUG hairpin loop: a novel RNA tetraloop motif. *Biochemistry*. 1995; 34:14416–27. [PubMed: 7578046]

- [4]. Molinaro M, Tinoco I. Use of ultra stable UNCG tetraloop hairpins to fold RNA structures: thermodynamic and spectroscopic applications. *Nucleic Acids Res.* 1995; 23:3056–63. [PubMed: 7544890]
- [5]. Jaeger L, Michel F, Westhof E. Involvement of a GNRA tetraloop in long-range RNA tertiary interactions. *J Mol Biol.* 1994; 236:1271–6. [PubMed: 7510342]
- [6]. Fiore JL, Nesbitt DJ. An RNA folding motif: GNRA tetraloop-receptor interactions. *Q Rev Biophys.* 2013; 46:223–64. [PubMed: 23915736]
- [7]. Wu L, Chai D, Fraser ME, Zimmerly S. Structural variation and uniformity among tetraloop-receptor interactions and other loop-helix interactions in RNA crystal structures. *PLoS One.* 2012; 7:e49225. [PubMed: 23152878]
- [8]. Keating KS, Toor N, Pyle AM. The GANC tetraloop: a novel motif in the group IIC intron structure. *J Mol Biol. England.* 2008:475–81.
- [9]. Ishikawa J, Furuta H, Ikawa Y. An in vitro-selected RNA receptor for the GAAC loop: modular receptor for non-GNRA-type tetraloop. *Nucleic Acids Res.* 2013; 41:3748–59. [PubMed: 23382175]
- [10]. Robart AR, Chan RT, Peters JK, Rajashankar KR, Toor N. Crystal structure of a eukaryotic group II intron lariat. *Nature.* 2014; 514:193–7. [PubMed: 25252982]
- [11]. Fedorova O, Pyle AM. Linking the group II intron catalytic domains: tertiary contacts and structural features of domain 3. *EMBO J.* 2005; 24:3906–16. [PubMed: 16252007]
- [12]. Toor N, Hausner G, Zimmerly S. Coevolution of group II intron RNA structures with their intron-encoded reverse transcriptases. *RNA.* 2001; 7:1142–52. [PubMed: 11497432]
- [13]. Moody EM, Feerrar JC, Bevilacqua PC. Evidence that folding of an RNA tetraloop hairpin is less cooperative than its DNA counterpart. *Biochemistry.* 2004; 43:7992–8. [PubMed: 15209494]
- [14]. Antao VP, Lai SY, Tinoco I. A thermodynamic study of unusually stable RNA and DNA hairpins. *Nucleic Acids Res.* 1991; 19:5901–5. [PubMed: 1719483]
- [15]. Quigley GJ, Rich A. Structural domains of transfer RNA molecules. *Science.* 1976; 194:796–806. [PubMed: 790568]
- [16]. Jucker FM, Pardi A. GNRA tetraloops make a U-turn. *RNA.* 1995; 1:219–22. [PubMed: 7585251]
- [17]. Wadley LM, Keating KS, Duarte CM, Pyle AM. Evaluating and learning from RNA pseudotorsional space: quantitative validation of a reduced representation for RNA structure. *J Mol Biol.* 2007; 372:942–57. [PubMed: 17707400]
- [18]. Parlea LG, Sweeney BA, Hosseini-Asanjan M, Zirbel CL, Leontis NB. The RNA 3D Motif Atlas: Computational methods for extraction, organization and evaluation of RNA motifs. *Methods.* 2016; 103:99–119. [PubMed: 27125735]
- [19]. Petrov AI, Zirbel CL, Leontis NB. Automated classification of RNA 3D motifs and the RNA 3D Motif Atlas. *RNA.* 2013; 19:1327–40. [PubMed: 23970545]
- [20]. Abramovitz DL, Pyle AM. Remarkable morphological variability of a common RNA folding motif: the GNRA tetraloop-receptor interaction. *J Mol Biol.* 1997; 266:493–506. [PubMed: 9067606]
- [21]. Cate JH, Hanna RL, Doudna JA. A magnesium ion core at the heart of a ribozyme domain. *Nat Struct Biol.* 1997; 4:553–8. [PubMed: 9228948]
- [22]. Misra VK, Draper DE. On the role of magnesium ions in RNA stability. *Biopolymers.* 1998; 48:113–35. [PubMed: 10333741]
- [23]. Draper DE. A guide to ions and RNA structure. *RNA.* 2004; 10:335–43. [PubMed: 14970378]
- [24]. Steitz TA, Steitz JA. A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci U S A.* 1993; 90:6498–502. [PubMed: 8341661]
- [25]. Bowman JC, Lenz TK, Hud NV, Williams LD. Cations in charge: magnesium ions in RNA folding and catalysis. *Curr Opin Struct Biol.* 2012; 22:262–72. [PubMed: 22595008]
- [26]. Toor N, Keating KS, Taylor SD, Pyle AM. Crystal structure of a self-spliced group II intron. *Science. United States.* 2008:77–82.
- [27]. Toor N, Rajashankar K, Keating KS, Pyle AM. Structural basis for exon recognition by a group II intron. *Nat Struct Mol Biol. United States.* 2008:1221–2.

- [28]. Costa M, Fontaine JM, Loiseaux-de Goër S, Michel F. A group II self-splicing intron from the brown alga *Pylaiella littoralis* is active at unusually low magnesium concentrations and forms populations of molecules with a uniform conformation. *J Mol Biol.* 1997; 274:353–64. [PubMed: 9405145]
- [29]. Lemieux S, Major F. Automated extraction and classification of RNA tertiary structure cyclic motifs. *Nucleic Acids Res.* 2006; 34:2340–6. [PubMed: 16679452]
- [30]. Dai L, Toor N, Olson R, Keeping A, Zimmerly S. Database for mobile group II introns. *Nucleic Acids Res.* 2003; 31:424–6. [PubMed: 12520040]
- [31]. Tippmann HF. Analysis for free: comparing programs for sequence analysis. *Brief Bioinform.* 2004; 5:82–7. [PubMed: 15153308]
- [32]. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. *Genome Res.* 2004; 14:1188–90. [PubMed: 15173120]

Research Highlights

- We have discovered a novel class of interacting loops present in a group IIB intron and a previous structure of the thiamine pyrophosphate riboswitch
- This pentaloop sequence interacts with a receptor helix at a 70° angle deviation from the more common GNRA tetraloop-receptor interaction
- This work highlights the fact that there may be additional classes of RNA tertiary interactions to be discovered in the future

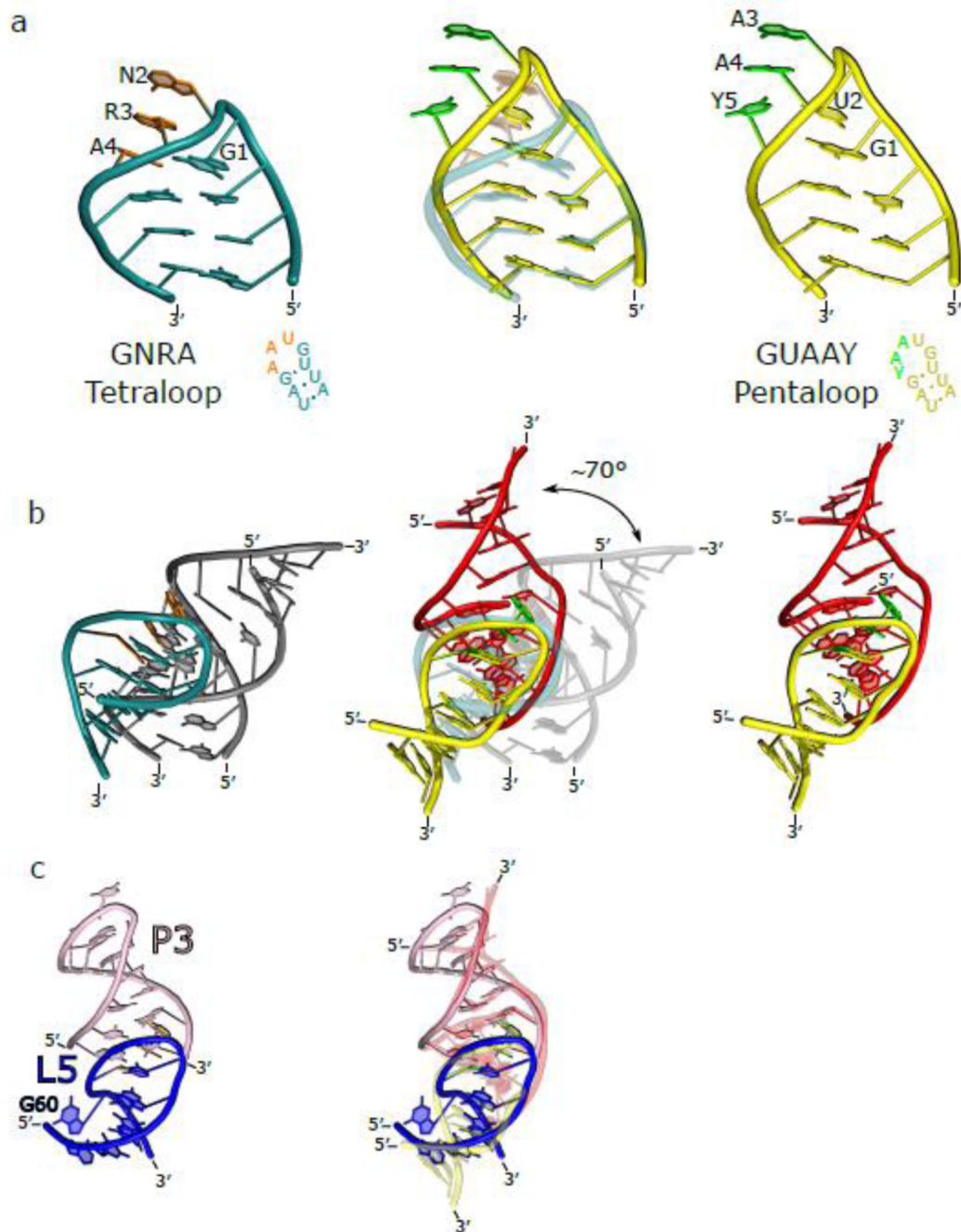


Figure 2. Comparison of the GNRA tetraloop versus the GUAAY pentaloop

A) The GNRA tetraloop (left) and the GUAAY pentaloop (right) are connected to a three base-pair helix and aligned to the bottom of the stem. Central panel shows the alignment of the two loops. This alignment shows that the the GUAAY pentaloop extends approximately one base pair higher and at a different orientation compared to the GNRA tetraloop. Nucleotides of the GNRA tetraloop and GUAAY pentaloop that interact with their cognate receptor are colored in orange and green, respectively. B) The GNRA tetraloop (left) nucleotides insert parallel to the base planes of the receptor, while the GUAAY pentaloop

(right) nucleotides are nearly perpendicular to the receptor base planes. This allows a greater number of nucleotides to insert into the minor groove. The alignment of the two structures (central panel) shows a $\sim 70^\circ$ deviation in terms of the angle of attack of the loop for the receptor. C) The structure of the equivalent GUAAY-containing hexanucleotide loop-receptor (left) from the thiamine pyrophosphate riboswitch (PDB 2CKY). The alignment with the GUAAY pentaloop-receptor from the group IIB intron (right) shows an almost identical mode of recognition between the two components of the interaction.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

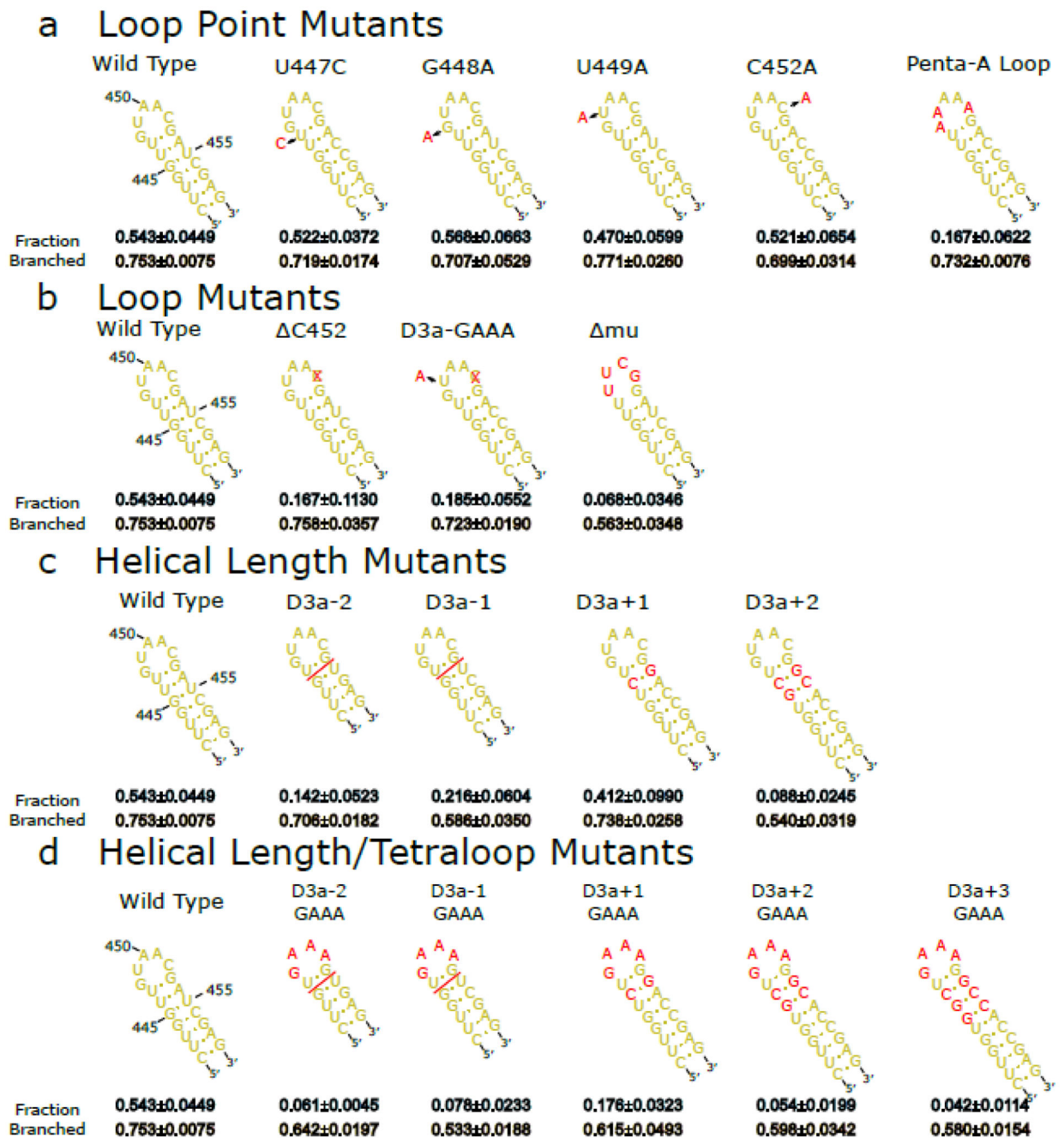


Figure 3. Secondary structure of *P.li.LSUI2* D3a mutants

Mutants were grouped into four categories: a) loop point mutants, b) loop mutants, c) helical length mutants, and d) helical length/tetraloop mutants. Red letters indicate changes to the wild-type sequence. Red lines indicate where helical deletions took place. Values below the mutants correspond to Figure 4 for *in vitro* splicing assays at 2.5mM (blue) and 10mM (orange) $MgCl_2$.

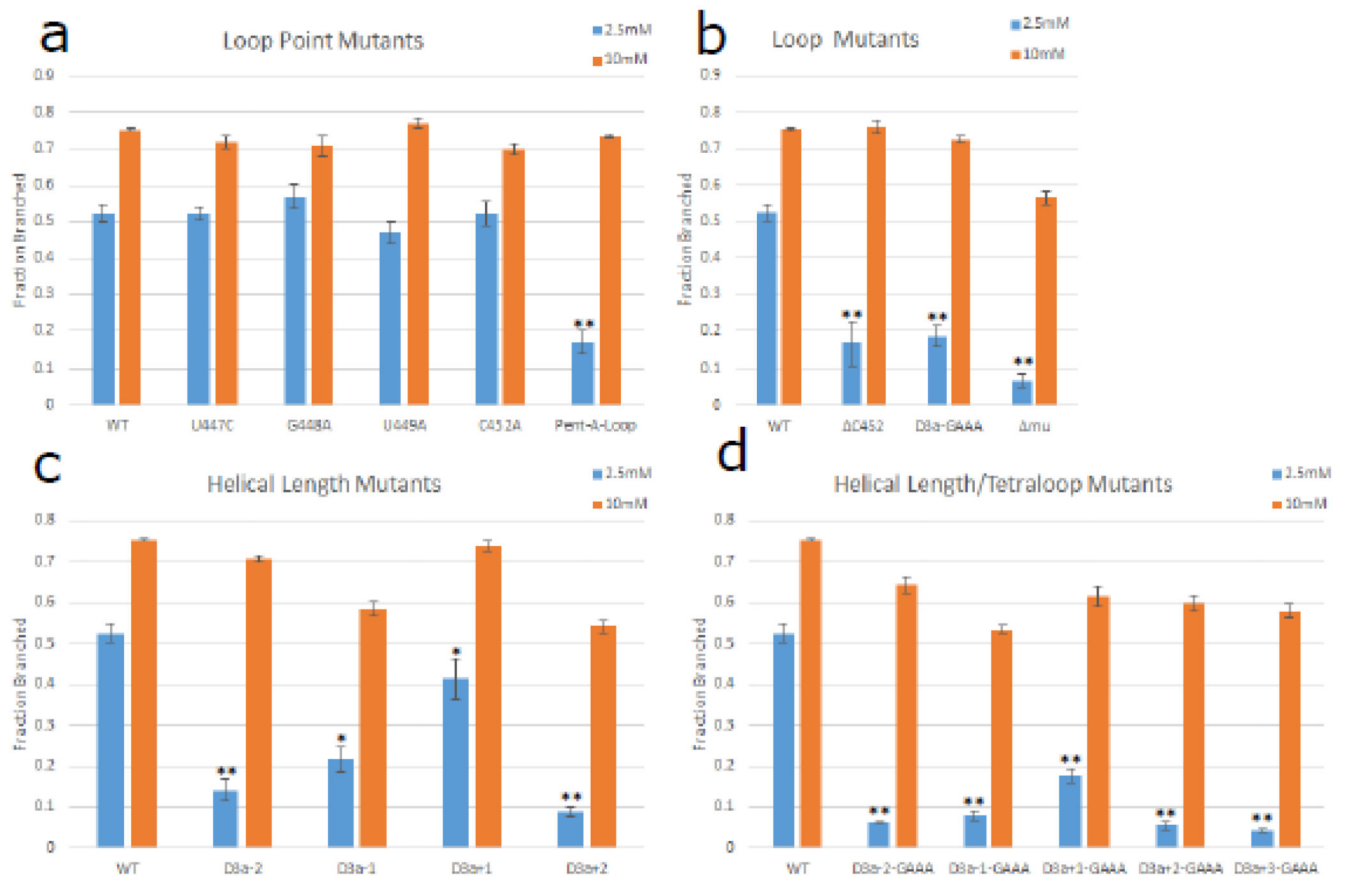


Figure 4. *In vitro* splicing assay of *P.li.LSUI2* D3a mutants

a) Point mutants made either in the closing base pair of the stem or in the pentaloop. The single point mutants retained activity comparable to wild-type. However, mutating the loop entirely to adenosines resulted in a drastic reduction in activity. b) Mutagenesis of the loop to any of the tetraloop sequences results in severe defects in splicing activity. c) Changing the helical length from the native seven base-pairs perturbs activity. d) The defects caused by the GAAA-tetraloop cannot be rescued with different helical length mutations. Thirty minute time points taken for all splicing assays. * <0.05 by two-tailed Student's t-test. ** <0.001 by two-tailed Student's t-test.

Table 1

Sequence alignment of the GUAAAY pentaloop in group IIB1 introns

Pti.I2 (*Pti.LSU12*) sequence is bolded. Sequence positions that are 100% conserved are denoted with a black background, while positions that are >75% conserved are denoted with a light grey background. Positions that were >95% pyrimidine and purine are denoted in light blue and light green, respectively.

Intron name	N	Y	U	G	U	A	A	Y	G	R	N
Pti.I2	G	U	U	G	U	A	A	C	G	A	U
Sh.fr.II	A	U	U	G	U	A	A	C	G	A	U
Vi.an.II	A	U	U	G	U	A	A	C	G	A	U
Vi.ha.II	A	U	U	G	U	A	A	C	G	A	U
Wo.sp.II	U	U	U	G	U	A	A	C	G	A	A
Wo.sp.I3	U	U	U	G	U	A	A	C	G	A	A
A.v.I3	C	C	U	G	U	A	A	C	G	G	G
A.v.II	C	C	U	G	U	A	A	C	G	G	G
Th.e.II	U	U	U	G	U	A	A	C	G	G	A
Al.sh.II	C	C	U	G	U	A	A	C	G	G	G
Cu.me.II	G	C	U	G	U	A	A	U	G	G	U
Cu.me.I2	G	C	U	G	U	A	A	U	G	G	U
Ag.r.II	G	C	U	G	U	A	A	U	G	G	U
Al.bo.II	C	C	U	G	U	A	A	U	G	G	G
Al.sh.I2	C	C	U	G	U	A	A	U	G	G	G
B.t.II	U	C	U	G	U	A	A	U	G	G	G
B.vu.II	C	C	U	G	U	A	A	U	G	G	G
B.vu.I2	U	C	U	G	U	A	A	U	G	G	G
Ciro.II	G	C	U	G	U	A	A	U	G	G	U
E.c.I5	U	C	U	G	U	A	A	U	G	G	A
P.st.II	C	C	U	G	U	A	A	U	G	G	G
H.s.II	A	C	U	G	U	A	A	U	G	G	U
Hm.ar.II	G	C	U	G	U	A	A	U	G	G	U
Kl.pn.I2	U	C	U	G	U	A	A	U	G	G	A
Le.pn.II	U	C	U	G	U	A	A	U	G	G	A

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Intron name	N	Y	U	G	U	A	A	A	Y	G	R	N
Le.pn.I2	G	C	U	G	U	A	A	A	U	G	G	C
P.st.I2	C	C	U	G	U	A	A	A	U	G	G	G
P.p.I2	C	C	U	G	U	A	A	A	U	G	G	G
P.p.I4	G	C	U	G	U	A	A	A	U	G	G	C
Po.sp.I3	U	C	U	G	U	A	A	A	U	G	G	A
Po.sp.I4	U	C	U	G	U	A	A	A	U	G	G	A
Po.sp.I1	U	C	U	G	U	A	A	A	U	G	G	A
Po.sp.I2	U	C	U	G	U	A	A	A	U	G	G	A
Re.sp.I1	G	C	U	G	U	A	A	A	U	G	G	C
Sh.sp.I2	U	C	U	G	U	A	A	A	U	G	G	A
The.I7	U	C	U	G	U	A	A	A	U	G	G	A
X.f.I1	C	C	U	G	U	A	A	A	U	G	G	G
The.I3	U	C	U	G	U	A	A	A	U	G	G	A
Eure.I1	C	U	U	G	U	A	A	A	U	G	A	G
An.pr.I1	C	U	U	G	U	A	A	A	U	G	A	G
Eu.si.I2	C	U	U	G	U	A	A	A	U	G	A	G
Fa.pr.I1	C	U	U	G	U	A	A	A	U	G	A	G
S.ce.I5	A	U	U	G	U	A	A	A	U	G	A	U
Sh.ba.I1	A	U	U	G	U	A	A	A	U	G	A	U
Sh.pi.I1	G	U	U	G	U	A	A	A	U	G	A	U
Sb.mo.I1	U	U	U	G	U	A	A	A	U	G	A	A
Sh.sp.I1	A	U	U	G	U	A	A	A	U	G	A	U
Vi.ha.I2	A	U	U	G	U	A	A	A	U	G	A	U
Sh.se.I1	A	U	U	G	U	A	A	A	U	G	A	U
E.c.I8	G	C	U	G	U	A	A	A	U	A	A	C
T.e.I8	U	G	U	G	U	A	A	A	U	A	C	A
R.pi.I1	G	C	U	G	U	A	A	A	U	C	G	U
P.ae.I2	G	C	U	G	U	A	A	A	U	C	G	U
Ch.lu.I1	C	C	U	G	U	A	A	A	U	C	G	G
Ch.ph.I1	C	C	U	G	U	A	A	A	U	C	G	G
Pr.ae.I4	C	C	U	G	U	A	A	A	U	G	G	G

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Intron name	N	Y	U	G	U	A	A	U	A	A	Y	G	R	N
S.eq.II	G	C	A	G	U	A	A	U	A	A	U	G	G	C
Hp.au.II	U	C	C	G	U	A	A	U	A	A	U	C	G	A