JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Communication
pubs.acs.org/JACS

A Role for a Single-Stranded Junction in RNA Binding and Specificity by the *Tetrahymena* Group I Ribozyme

Xuesong Shi, Sergey V. Solomatin, and Daniel Herschlag*

Department of Biochemistry, Stanford University, Stanford, California 94305, United States

Supporting Information

ABSTRACT: We have investigated the role of a single-stranded RNA junction, J1/2, that connects the substrate-containing P1 duplex to the remainder of the *Tetrahymena* group I ribozyme. Single-turnover kinetics, fluorescence anisotropy, and single-molecule fluorescence resonance energy transfer studies of a series of J1/2 mutants were used to probe the sequence dependence of the catalytic activity, the P1 dynamics, and the thermodynamics of docking of the P1 duplex into the ribozyme's catalytic core. We found that A29, the center A of three adenosine residues in J1/2, contributes 2 orders of magnitude to the overall ribozyme activity, and double-mutant cycles suggested that J1/2 stabilizes the docked state of P1 over the undocked state via a tertiary interaction involving A29 and the first base pair in helix P2 of the ribozyme, A31·U56. Comparative sequence analysis of this group I intron subclass suggests that the A29 interaction sets one end of a molecular ruler whose other end specifies the 5′-splice site and that this molecular ruler is conserved among a subclass of group I introns related to the *Tetrahymena* intron. Our results reveal substantial functional effects from a seemingly simple single-stranded RNA junction and suggest that junction sequences may evolve rapidly to provide important interactions in functional RNAs.

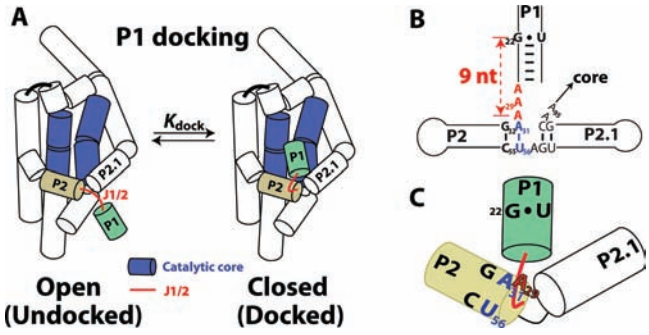

Figure 1. Schematic illustration of the *Tetrahymena* group I ribozyme and its J1/2 junction.[8] (A) The P1 duplex (green) docks into the catalytic core (blue); K_{dock} = [docked]/[undocked] is the equilibrium constant for the equilibrium between the undocked and docked states. The single-stranded J1/2 junction (red) connects the substrate containing the P1 duplex to the P2 duplex (beige), one of the several peripheral segments of the ribozyme. P denotes paired regions, and J denotes junctions that connect the paired regions. (B) J1/2 (red) and its surroundings. Individual residues that are mutated or discussed herein are highlighted. (C) Schematic model of the J1/2 conformation in the docked state (red line) based on results herein. J1/2 exits P2 and is bent to allow A29, the center A of J1/2 (shown in red), to form a tertiary interaction involving the first base pair in P2, A31·U56 (blue); also see Figure S3. The catalytic cleavage site is located a specified number of residues from this interaction in the 3′ direction (see Table 1 and the text).

While much focus has been placed on highly conserved regions of proteins and functional RNAs, these molecules also contain regions that have limited or no apparent conservation. There are regions of RNAs, such as telomeric RNA, SRP RNA, spliceosomal RNAs, and self-splicing introns, that are conserved only in subgroups,[1,2] and there are sequences of no obvious conservation that nevertheless occur in regions that might be expected to have functional consequences.[3] Here we report a multifaceted investigation of one such region, the J1/2 junction in the *Tetrahymena* group I intron (Figure 1).

J1/2 is not globally conserved in either length or sequence among different group I intron subgroups or within the IC1 subgroup that includes the *Tetrahymena* intron.[2,4,5] Nonetheless, J1/2 connects the substrate-containing P1 duplex to the rest of the intron (Figure 1A), and docking of the P1 duplex into tertiary interactions with the intron's catalytic core is a necessary step that precedes the chemical reaction[6,7] (Figure 1A). It was previously shown that shortening or lengthening J1/2 decreases the fidelity of splice-site selection in the *Tetrahymena* ribozyme reaction; the length changes weakened the docking of the P1 duplex into tertiary interactions in the correct register and thereby favored docking into and cleavage from alternative registers.[4,7] These mutational effects were not what would be expected for a simple tether, but the prior experiments could not distinguish whether tether flexibility, functional interactions with J1/2, non-native interactions with mutant J1/2 sequences, or steric constraints from the remainder of the ribozyme were responsible.

To understand the role of J1/2, we first used fluorescence polarization anisotropy (FPA) to assess the dynamics of the P1 duplex attached to ribozymes with J1/2 sequences that were systematically mutated. We related the effects of these J1/2 mutations on P1 dynamics to their functional consequences as assessed by single-molecule fluorescence resonance energy transfer (smFRET) assays of P1 docking and assays of catalytic activity. The results revealed a role for A29, the central A of the J1/2 AAA sequence (Figure 1B,C), and further mutational tests

Received: September 9, 2011
Published: January 5, 2012

dx.doi.org/10.1021/ja2083575 | *J. Am. Chem.Soc.* 2012, 134, 1910–1913

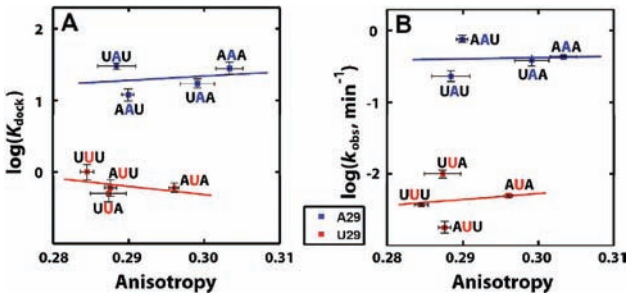

Figure 2. Effects of the J1/2 base sequence on (A) the docking equilibrium constant, K_{dock}, and (B) the catalytic activity, k_{obs}. The four sequences with the center residue (position 29) as A (blue) are represented by the blue lines, and the four sequences with A29 replaced by U (red) are represented by the red lines. K_{dock} in (A) was obtained from smFRET assays (see the SI). The activity in (B) was measured by following the reaction of $(E{\cdot}S{\cdot}G)_o$ using a ^{32}P-radiolabeled open-complex substrate under single-turnover conditions with saturating enzyme and saturating G (2 mM) in an assay that monitored both docking and the chemical step.[17] Conditions: 50 mM Na·MOPS (pH 7.0), 10 mM $MgCl_2$, and 25 °C (see the SI for details).

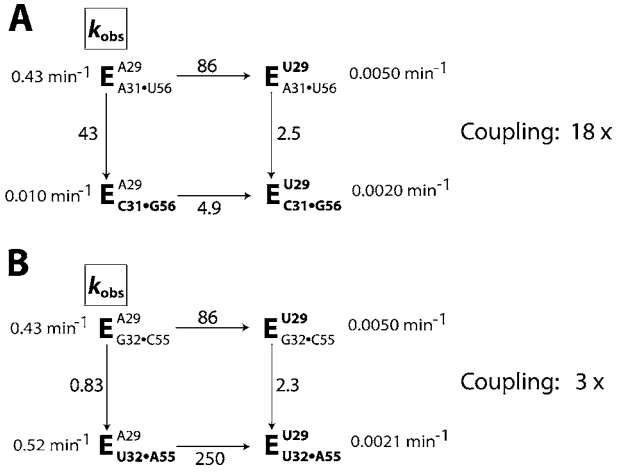

Figure 3. Testing a potential contact between A29 and two base pairs in P2, (A) A31·U56 and (B) G32·C55, via double-mutant cycles. The wild-type sequence has A29, A31·U56, and G32·C55. The introduced mutations are shown in bold. The numbers adjacent to the arrows are the fold decrease in k_{obs} resulting from introduction of the mutation. The amounts of coupling between the A29 interaction and the P2 mutation are shown at the right.

and sequence analyses provided support for a tertiary interaction involving this residue.

To modulate the flexibility of J1/2, we systemically replaced the A residues with U's, as U residues stack less well than A residues.[9,10] Eight ribozymes were investigated, with J1/2 sequences of AAA, AAU, AUA, UAA, AUU, UAU, UUA, and UUU. 6-Methyl isoxanthopterin (6-MI) was incorporated into the P1 duplex of each ribozyme [Figure S1A in the Supporting Information (SI)]. This fluorescent base analogue has the unusual property of maintaining a high quantum yield within helices and thus can be used to follow the dynamic properties of individual helices within complex RNAs.[11,12] In its open state (Figure 1A, left), the P1 helix is connected to the remainder of the ribozyme by J1/2 but appears to make no specific tertiary interactions.[12,13] As expected for the behavior of a tether, the anisotropy decreases, and thus the dynamics increase, as the number of U residues in J1/2 increases (Figure S1B). In other words, there is greater randomization of the position of P1 during the fluorescent lifetime of 6-MI as more U residues are introduced.

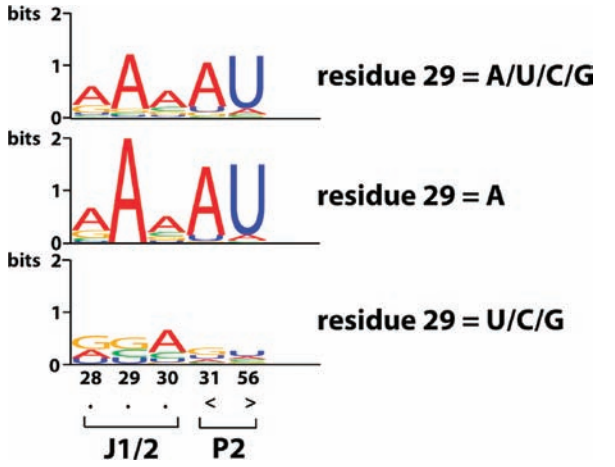

Figure 4. Co-conservation of A29 and the A31·U56 base pair.[20] The height of the letter for a base type i (i = A/U/G/C) at residue r in the sequence logos[23] is $f_{i,r}I_r$ in unit of bits, where $f_{i,r}$ is the fraction of residue r that is base type i and $I_r = 2 + \sum_i(f_{i,r} \log_2 f_{i,r})$ is the information content[24] of residue r. The top sequence logo includes all sequences that have a residue 29 (defined as the second base 5′ to P2 in the single-stranded region of J1/2; sequences with J1/2 of 2 nt or longer all have a residue 29; see also Tables 1 and Table S3). The middle sequence logo includes only the sequences with A at residue 29. The bottom sequence logo includes the sequences with residue 29 being anything other than A. The dots under residues 28–30 indicate unpaired residues, and the angle brackets under residues 31 and 56 indicate base pairing.

A simple prediction from a tether model would be that the effect of J1/2 mutations on increasing mobility in the open complex would inversely correlate with the stability of the docked complex, as increased conformational freedom would disfavor the more positioned docked complex. The docking equilibrium constants (K_{dock}) for the eight mutants were obtained by monitoring the open and closed states using an smFRET assay[14−16] (see the SI for details). We observed a striking discordance between the anisotropy and docking behaviors (Figure 2A). Whereas introduction of U residues at any position increased the mobility, the docking equilibrium constant was substantially affected only by substitution of A29. Mutation of either or both of the flanking A residues had effects of less than 3-fold on the docking (Figure 2A, blue), whereas mutation of the A29 decreased the docking ~30-fold (Figure 2A, red) regardless of the identity of the flanking residues. This same trend was observed for ribozyme activity in assays that monitored both docking and the chemical step (Figure 2B), but with about 3-fold larger effects for A29, either because of a small additional effect on the chemical step or because the magnitude of the docking effect differed slightly under the different assay conditions.

The simplest model to account for all of the data is that the J1/2 flexibility difference between 3A and 3U contributes little to P1 docking and that J1/2 stabilizes P1 docking through tertiary interactions involving A29. To test this model further, we carried out additional mutagenesis studies.

We first tested the base specificity of the putative A29 interaction by determining the catalytic activity of two additional J1/2 mutants with sequences of AGA and ACA. The values of k_{obs} for these mutants were within a factor of 3 of

dx.doi.org/10.1021/ja2083575 | J. Am. Chem.Soc. 2012, 134, 1910−1913

Table 1. Summary of a Comparative Sequence Analysis of Different Length Combinations of J1/2 and P1 for the IC1 Subgroup of Group I Introns[a]

J1/2 length, nt	P1 length, bp						
	1-2	3	4	5	6	7	8
6-11	0	0	1 (1.0, 1.0)	7 (0.29, 0.0)	0	0	0
5	0	0	1 (1.0, 1.0)	3 (0.67, 0.0)	2 (1.0, 0.0)	0	0
4	0	0	0	291 (0.97, 0.99)	10 (0.40, 0.75)	0	0
3	0	0	1 (0.0, −)	20 (0.75, 0.27)	27 (1.0, 0.96)	2 (1.0, 0.50)	0
2	0	9 (0.0, −)	3 (0.33, 1.0)	6 (0.33, 0.50)	21 (0.38, 0.0)	8 (1.0, 0.88)	0
1	1	2	89	112	1	0	0
0	1	1	3	178	6	3	1

[a]The number before a set of parentheses is the total number of sequences for a given J1/2 length (row) and P1 length (column; the P1 length includes the conserved G·U base pair that specifies the cleavage site). Sequences in the IC1 group I intron database[20] (3%) that do not have the conserved cleavage-site G·U base pair were excluded. The first of the two numbers in each set of parentheses is the fraction of residue 29 that is A, and the second is the fraction of residue 31 that is A when residue 29 is A. In our comparative sequence analysis of J1/2 residues, the residue equivalent to A29 is the second nucleotide upstream of P2. For J1/2 sequences of length 0 or 1, there is no residue equivalent of A29, so no values in parentheses are given. Subgroups with a combined P1 and J1/2 length of 9 nt are colored in red. Other highly populated subgroups are colored to represent a combined P1 and J1/2 length of 6 nt (green) or 5 nt (blue).

that for AUA mutant, with both reacting >80-fold slower than the wild-type AAA (Table S1 in the SI). Thus, the effect of residue 29 is specific to A.

We next wanted to identify potential interaction partners for A29. We first crudely assessed the geometrical accessibility of A29 to other residues. Using the structural model for the *Tetrahymena* group I intron,[8] we considered residues within a sphere with its origin at A31 and a radius of 11 Å, which is roughly the length of an extended 2 nucleotide (nt) linker.[18] Among the accessible residues, we found that two base pairs in the P2 stem, A31·U56 and G32·C55 (Figure 1B), exhibited some degree of sequence coconservation with A29 (i.e., the second residue of J1/2 that is 5′ of P2; see below). Double-mutant cycles were used to test for interactions[19] between A29 and these P2 base pairs. Briefly, the effects of mutation of each of the putative interaction partners (A29 and each of the P2 base pairs) were determined alone and together (Table S1). If there were an interaction, then a lessened effect would be expected with the other mutation present. Such a dependence was observed with A29 and the A31·U56 base pair (Figure 3A). Mutation of either A29 or A31·U56 alone gave 40−80-fold effects, whereas each mutation in a background in which the other mutation had already been made gave an effect of <5-fold. In contrast, mutation of G32·C55 had no significant effect, and there was a similar large effect of mutation of A29 whether the G32·C55 base pair was wild-type or mutant (Figure 3B). While the simplest model for the functional interaction between A29 and the A31·U56 base pair would be a base triple, additional mutagenesis tests provided no evidence for an isosteric base triple (Table S4). Specifically, double-mutant cycles revealed that whereas there was energetic coupling of A29 with the A31·U56 mutant to C·G (Figure 3A), as noted above, no energetic coupling was observed with G·C and U·A base pairs (Table S4). Thus, more complex models involving additional interactions and/or conformation rearrangements must be invoked.

We mutated additional residues potentially in the vicinity of A29. Modest coupling was observed for residue A95 (3-fold), A304 (2-fold) and A270 (≥10-fold enhancement of the A29 effect) (Tables S4 and S5 and unpublished results). Conversely, A269 and the first two base pairs of P2.1 had no energetic interaction with A29 (Table S5 and unpublished results). These results support a model in which A29 is situated close to and possibly interacting with A31·U56 and also provide evidence for an extended network of indirect interactions that extends to the catalytic core (A270 and A304; Figure S3). The absence of a larger anisotropy effect for the A29 mutants than for the other J1/2 mutants (Figure S1B and Table S1) suggests that this interaction network does not include A29 in the undocked state (Figure S1B). Thus, the A29 interaction very likely forms along with docking of P1.

The functional interaction between A29 of J1/2 and the A31·U56 base pair led us to look more closely at potential phylogenetic relationships using the extensive sequence database for group I introns.[20] We found that within the IC1 subgroup of introns, the mutual information (MI)[21] between residue 29, the second residue upstream of P2, and residues 31 and 56 (MI = 0.29 and 0.19, respectively), which compose the first base pair in P2, is significantly higher than the MI between residue 29 and random residues in the rest of the intron sequence (MI = 0.06 ± 0.06; see Table S2 for more information). The high MI comes from a strong coconservation between A29 and the A31·U56 base pair and is further illustrated in Figure 4 using sequence logos.[23] The observed sequence coconservation between A29 and the A31·U56 base pair is consistent with a functional interaction, as supported by the double-mutant cycles described above. Despite the strong coconservation, there is no covariation or observed isosteric three-base combinations of residue 29, 31 and 56, also in agreement with the experimental functional tests described above (Table S4).

A conserved number of residues equal to 12 from a GNRA tetraloop at the end of the P2 helix to the conserved G·U wobble cleavage site in P1 was previously observed in several group I intron subgroups (IC3, IB2, IB4, IA1, IA2, and IA3)[2,22,23] but not in the IC1 subgroup that lacks this tetraloop.[2] We asked whether there might be an analogous relationship for the IC1 subgroup that includes the *Tetrahymena* intron (Table 1). Many introns within the IC1 subgroup (327 of 810; see the red diagonal in Table 1) have a combined length of 9 for J1/2 and P1 (up to the conserved G·U pair; Figure 1), and nearly all of these introns (314 of 327; Table 1, red) have an A residue two residues upstream of P2 and an A·U base pair as the first base pair of P2. Thus, a

dx.doi.org/10.1021/ja2083575 | J. Am. Chem.Soc. 2012, 134, 1910−1913

measuring mechanism is suggested, consistent with prior observations of reduced 5′-splice-site fidelity upon lengthening or shortening of J1/2 of the *Tetrahymena* intron.[4,7]

Interestingly, a significant number of IC1 introns do not follow this "rule of nine", with some of these other introns having and some not having the above-noted A·U base pair (e.g., Table 1, blue and green). It appears that there have been multiple solutions within this subgroup for ensuring accurate 5′-splice-site selection, and there is more to be learned about the evolutionary journey of these fascinating catalytic RNAs.

Counterintuitively, junction sequences without apparent conservation may be used liberally to optimize RNA function. We have shown that J1/2 plays sophisticated roles in RNA function. J1/2 is important for substrate binding and ribozyme activity through interactions involving A29, the center A of J1/2 of the *Tetrahymena* intron. A29 interacts with an extended network of residues connecting J1/2 with the RNA core. Further, A29 and the constant J1/2 and P1 length allow distal tuning of the ribozyme activity and specificity via a molecular ruler. These results suggest that regions known as "linkers" can be important for function. Junctions that are not constrained to form secondary or tertiary structures may be able to explore sequence and conformational space extensively and thus may rapidly evolve new functional interactions. The ability of J1/2, a seemingly nonconserved linker, to be co-opted for function provides an important precedent and model for dissection of the function of other structured RNAs.

■ ASSOCIATED CONTENT

Supporting Information

Methods; anisotropy, K_{dock} and k_{obs} data; additional comparative sequence analysis and P1 docking data; heterogeneity in smFRET measurements; P1 anisotropy results; typical smFRET traces; and functional effects of mutations. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

herschla@stanford.edu

■ ACKNOWLEDGMENTS

We thank Pablo Cordero, Alain Laederach, and members of the Herschlag lab for comments on the manuscript. This work was supported by NIH Grant GM49243.

■ REFERENCES

(1) Chen, J. L.; Blasco, M. A.; Greider, C. W. *Cell* **2000**, *100*, 503−514.
(2) Michel, F.; Westhof, E. *J. Mol. Biol.* **1990**, *216*, 585−610.
(3) Bonen, L.; Vogel, J. *Trends Genet.* **2001**, *17*, 322−331.
(4) Young, B.; Herschlag, D.; Cech, T. R. *Cell* **1991**, *67*, 1007−1019.
(5) Che, A. J.; Knight, T. F. Jr. *Nucleic Acids Res.* **2010**, *38*, 2748−2755.
(6) Bevilacqua, P. C.; Kierzek, R.; Johnson, K. A.; Turner, D. H. *Science* **1992**, *258*, 1355−1357.
(7) Herschlag, D. *Biochemistry* **1992**, *31*, 1386−1399.
(8) Lehnert, V.; Jaeger, L.; Michel, F.; Westhof, E. *Chem. Biol.* **1996**, *3*, 993−1009.
(9) Tso, P. O. P.; Melvin, I. S.; Olson, A. C. *J. Am. Chem. Soc.* **1963**, *85*, 1289−1296.
(10) Holcomb, D. N.; Tinoco, I. *Biopolymers* **1965**, *3*, 121−133.
(11) Hawkins, M. E.; Pfleiderer, W.; Balis, F. M.; Porter, D.; Knutson, J. R. *Anal. Biochem.* **1997**, *244*, 86−95.
(12) Shi, X.; Mollova, E. T.; Pljevaljcic, G.; Millar, D. P.; Herschlag, D. *J. Am. Chem. Soc.* **2009**, *131*, 9571−9578.
(13) Grant, G. P. G.; Boyd, N.; Herschlag, D.; Qin, P. Z. *J. Am. Chem. Soc.* **2009**, *131*, 3136−3137.
(14) Bartley, L. E.; Zhuang, X. W.; Das, R.; Chu, S.; Herschlag, D. *J. Mol. Biol.* **2003**, *328*, 1011−1026.
(15) Greenfeld, M.; Pavlichin, D. S.; Mabuchi, H.; Herschlag, D. *PLoS One* **2012**, in press.
(16) Zhuang, X. W.; Bartley, L. E.; Babcock, H. P.; Russell, R.; Ha, T. J.; Herschlag, D.; Chu, S. *Science* **2000**, *288*, 2048−2050.
(17) Herschlag, D.; Eckstein, F.; Cech, T. R. *Biochemistry* **1993**, *32*, 8299−8311.
(18) Smith, S. B.; Cui, Y. J.; Bustamante, C. *Science* **1996**, *271*, 795−799.
(19) Horovitz, A. *Folding Des.* **1996**, *1*, R121−R126.
(20) Zhou, Y.; Lu, C.; Wu, Q.-J.; Wang, Y.; Sun, Z.-T.; Deng, J.-C.; Zhang, Y. *Nucleic Acids Res.* **2008**, *36*, D31−D37.
(21) Gutell, R. R.; Power, A.; Hertz, G. Z.; Putz, E. J.; Stormo, G. D. *Nucleic Acids Res.* **1992**, *20*, 5785−5795.
(22) Haugen, P.; Simon, D. M.; Bhattacharya, D. *Trends Genet.* **2005**, *21*, 111−119.
(23) Schneider, T. D.; Stephens, R. M. *Nucleic Acids Res.* **1990**, *18*, 6097−6100.
(24) Schneider, T. D.; Stormo, G. D.; Gold, L.; Ehrenfeucht, A. *J. Mol. Biol.* **1986**, *188*, 415−431.