



Published in final edited form as:

Science. 2008 April 4; 320(5872): 77–82. doi:10.1126/science.1153803.

Crystal Structure of a Self-Spliced Group II Intron

Navtej Toor^{1,*}, Kevin S. Keating², Sean D. Taylor³, and Anna Marie Pyle^{1,4,*}

¹Department of Molecular Biophysics and Biochemistry, Yale University, 266 Whitney Avenue, Bass Building, New Haven, CT 06511, USA

²Interdepartmental Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT 06511, USA

³Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06511, USA

⁴Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA

Abstract

Group II introns are self-splicing ribozymes that catalyze their own excision from precursor transcripts and insertion into new genetic locations. Here we report the crystal structure of an intact, self-spliced group II intron from *Oceanobacillus iheyensis* at 3.1 angstrom resolution. An extensive network of tertiary interactions facilitates the ordered packing of intron subdomains around a ribozyme core that includes catalytic domain V. The bulge of domain V adopts an unusual helical structure that is located adjacent to a major groove triple helix (catalytic triplex). The bulge and catalytic triplex jointly coordinate two divalent metal ions in a configuration that is consistent with a two-metal ion mechanism for catalysis. Structural and functional analogies support the hypothesis that group II introns and the spliceosome share a common ancestor.

Group II introns are self-splicing ribozymes that catalyze their own excision from precursor-mRNAs (1). They also function as retroelements by associating with intron-encoded reverse transcriptases and invading DNA targets through reverse splicing reactions (2, 3). Group II introns are numerous in bacteria (4), and they are also found in the organellar genomes of plants, fungi, protists, and some animals (1, 5). They are considered to be the ancestors of nuclear introns and of the eukaryotic spliceosomal machinery, with which they share structural and sequence similarities (6, 7). Nuclear introns are regions of eukaryotic transcripts that are removed after transcription by a large ribonucleoprotein complex called the spliceosome. Despite the importance of group II introns and their spliceosomal relatives, there are no high-resolution crystal structures of these genetic elements in an intact state.

*To whom correspondence should be addressed. anna.pyle@yale.edu (A.M.P.); navtej.toor@yale.edu (N.T.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/320/5872/77/DC1

Materials and Methods

Figs. S1 to S8

Table S1

References

Most group II intron RNAs encode two basic components: a self-splicing ribozyme and an open reading frame (ORF) for expression of a reverse-transcriptase (RT) enzyme. The ribozyme is composed of six structural domains. The largest, domain I (DI), contains recognition sequences for binding the 5' and 3' exons and the branch site nucleophile. The 5' exon forms extended base-pairings with two binding sites in DI (EBS1 and EBS2) (8), and the 3' splice site is specified by a short pairing between EBS3 and the 3' exon (9). Domains II and III (DII and DIII) enhance the catalytic efficiency of splicing (10), and the linker that connects them (J2/3) is a major active-site component (11, 12). Domain IV (DIV) contains the reverse transcriptase ORF and structural motifs that bind the RT protein (13). The most highly conserved substructure in group II introns is domain V (DV), which is a short hairpin that contains a bulge essential for catalysis (14). The bulge is located near a "catalytic triad" of conserved nucleotides (usually AGC, but often CGC) at the base of domain V (15). Domain VI (DVI) contains an adenosine nucleotide that attacks the 5' splice site during the first step of splicing, forming lariat RNA (1). Although biochemical studies have revealed many aspects of group II intron architecture (1, 9, 16), the exact spatial organization of functional domains and molecular details of the ribozyme active site have remained obscure.

Here we describe the crystal structure of an intact group II intron from the halotolerant alkaliphile *Oceanobacillus iheyensis* at 3.1 Å resolution. This intron was identified in a screen of group II introns from various extremophilic bacteria and was chosen as a crystallization target because it exhibited robust splicing under conditions of low magnesium-ion concentration and high temperature. Indeed, the intron readily self-spliced during in vitro transcription by T7 RNA polymerase (fig. S1). The *O. iheyensis* intron belongs to the newly discovered group IIC class of introns (fig. S2), which are highly reactive and smaller (420 to 480 nucleotides) than the well-studied IIA and IIB classes (17, 18). The IIC introns are hypothesized to be the most primitive of the three classes (19), and they self-splice in vitro through a hydrolytic pathway (20), forming a linear intron (18, 21) instead of a cyclized lariat product.

The crystallization construct contains all of the six domains typically present in group II introns (Fig. 1A). The intron was crystallized in a postcatalytic state after undergoing both steps of splicing during in vitro transcription, which results in an intron RNA with homogeneously cleaved ends. The intron was then isolated in the native state, without the use of any denaturation steps that are typical for RNA purification (22). Native gel electrophoresis revealed the purified RNA to be conformationally homogeneous (fig. S3).

The *O. iheyensis* intron crystallized in space group $P2_12_12_1$ with one molecule per asymmetric unit. The structure was solved by multiwave-length anomalous dispersion (23). Phases were independently calculated from both Yb^{3+} and iridium hexamine (24) derivatives and combined, resulting in an experimental electron density map of high quality (Fig. 2 and figs. S4 to S6). The phylogenetically predicted secondary structure (17) served as a guide for building the model into the electron density. The final model had an R_{work} of 27.6% and an R_{free} of 31.0% (table S1). Density was observed for most of the intron nucleobases, although density for the κ region (Fig. 1A) and the internal loop of DIII was poor. Disordered regions of the intron included the first two nucleotides of the 5' end, DVI, and a small section of DI (Fig. 1A).

Overall structure

The dominant feature of the *O. iheyensis* intron structure is a network of intricate tertiary interactions that organize active-site components around the DV catalytic center (Fig. 1, C and D, and fig. S7A). Coaxial stacking of specific helices dictates the overall architectural form (Fig. 1B). For example, helices I(i) and I(ii) from DI are coaxially stacked on DII. Stem IC lies parallel to the I(i)–I(ii) helices in an orientation that is capped and stabilized by the θ - θ' interaction. Domains III and IV are also coaxially stacked, as are stems IA and IB in DI.

Long-range tertiary interactions

Most of the tertiary interactions proposed for group IIB introns also exist or have an analog in IIC introns. Some of these include α - α' (25), EBS1-IBS1 (8), EBS3-IBS3 (9), γ - γ' (26), ε - ε' (26), λ - λ' (27), θ - θ' (28), κ - κ' (29), and ζ - ζ' (30). Some of these interactions, such as θ - θ' , form a structure exactly as predicted from previous biochemical data (28). In contrast, other predicted interactions, such as α - α' and ζ - ζ' , have an unexpected structural form (see below). There are also some tertiary contacts that are missing in IIC introns. For example, μ - μ' (31) and the EBS2-IBS2 pairing are absent (18). In addition, there is a new long-range contact, ω - ω' , that only seems to exist in this primitive class of introns.

DI contains numerous conserved sequences and tertiary interactions that are important for maintaining the overall fold of the ribozyme. Consistent with previous folding studies, DI is structured as an autonomous scaffold that appears to organize the other domains (32). The most dominant long-range interaction within DI is the kissing loop between α and α' (Fig. 3A). This is a helix of seven consecutive Watson-Crick base pairs, followed and reinforced by a Watson-Crick pair (A50-U198) that points away from the α - α' helix. The coaxial stacking of stems IA and IB is essential for properly positioning the α sequence and for stabilizing the five-way junction in DI. This junction is reinforced by multiple stacking and pairing interactions between stem IA and the junction nucleotides (Fig. 3B and fig. S7B).

The θ - θ' interaction consists of a GCGA tetraloop from stem IC that docks into a receptor at the base of DII. This is a canonical GNRA tetraloop-receptor interaction (30) that positions the conserved ε' sequence within the core of the intron.

An unexpected interaction within DI, designated as ω - ω' (Fig. 3C), places EBS1 near catalytically essential motifs in DV and positions the 5' splice junction within the intron active site. The ω - ω' interaction involves the formation of a ribose zipper (33) between a small, conserved internal loop (ω') near EBS1 and stem IDI (ω).

Interactions between DI and DV

The active site is constructed through a network of interactions between the DI scaffold and DV. One of the most important contacts between these domains is ζ - ζ' , which consists of a GAAC tetra-loop from DV interacting with a receptor in DI (Fig. 1, A and C). This tetraloop-receptor interaction is conformationally unusual and consists only of base stacking between the AAC of the tetraloop and a single, bulged guanosine within the receptor.

Indeed, this interaction seems to be a new class of tetraloop receptor interaction found only in IIC introns and is identifiable by a GANC tetraloop sequence, in which the third nucleotide is variable.

Another important connection between DV and DI is κ - κ' , which serves to anchor DV within the DI scaffold (29). Although poor density in the κ loop prevents a detailed analysis of molecular features, the two elements of κ - κ' are close together and thus consistent with an interaction (Fig. 1C).

The z-anchor mediates structural integrity of the core

Nucleotides 106 to 111, from subdomain IC, form a series of base pairs, triples, and stacking interactions with three different regions of the intron. Two 160° kinks in the backbone, between nucleotides 109 to 111, cause nucleobases in this region to alternate from side to side (Fig. 4A), forming interactions with disparate parts of the intron and assembling them into a scaffold for the active site. The known ε - ε' and λ - λ' interactions are components of this larger, functional substructure, which we have named the “z-anchor.” The z-anchor has a form and function that is reminiscent of J8/7 in group I introns (34).

The z-anchor makes multiple contacts with the I(i) loop of DI and with the 5' end of the intron. For example, residues G108 and A110 form Watson-Crick pairs with nucleotides C11 and U259 of the I(i) loop (Figs. 1B and 4, A and B), whereas G107 forms a wobble pair with U4 (analogous to ε - ε'), thereby placing the z-anchor near the 5' splice site. In addition to these contacts, the I(i) loop and the 5' end of the intron interact directly through a set of Watson Crick/sugar-edge pairs (Figs. 1B and 4A).

The z-anchor ties this complex of DI constituents to DV via a base triple involving A106 and a base quartet involving G107. The base triple is analogous to the λ - λ' interaction that was chemogenetically identified in IIB introns (27). The sugar-edge of A106 forms multiple base and 2'-OH contacts to the sugar-edges of both C367 and G374 in DV (Fig. 4C). The unusual base quartet connects nucleobase and sugar functional groups of G107, U4, A366, and U375 (Fig. 4D). In addition, the I(i) loop and the 5' end of DI each make direct contacts with DV. The highly conserved G5 residue interacts with A376 in the DV bulge, allowing that nucleotide to extrude from the DV helix and thereby support the active-site structure. This elaborate network of tertiary contacts is crucial for intron function because it places the 5' splice site near the bulge of DV and mediates most of the interactions surrounding the DV bulge, thereby stabilizing the catalytic core of the intron.

J2/3 and bulge base C377 form a triple helix with the DV catalytic triad

J2/3 is a highly conserved linker region between domains II and III. Enzymological investigations have shown that J2/3 is important for efficient splicing by group II introns (11, 12), and photo-crosslinking studies have repeatedly placed J2/3 near the catalytic triad (16, 35). In the *O. iheyensis* intron, J2/3 is composed of nucleotides A287, G288, and C289. Nucleobases G288 and C289 insert into the major groove of DV, where they stack directly beneath bulge base C377 (Fig. 5, A and B, and fig. S7C). Each nucleotide of this stacked array (i.e., G288, C289, and C377) forms a base triple with the three nucleotides of the

catalytic triad located in DV (residues 358 to 360, Fig. 5B). In essence, the J2/3 strand and the DV stem completely merge, forming a triple helix that brings together catalytically essential residues of the intron (the catalytic triplex).

The bulge of DV forms an unusual helix

Base-stacking interactions between the DV bulge and residues G5, G288, and C289 distort the bulge region and induce formation of a short, single-stranded helical structure that spans nucleotides 375 to 379 (Fig. 5A and fig. S7C). This “mini-helix” is comparable in radius and pitch to the α helices commonly found in proteins.

This structure of DV differs greatly from the conformation of this domain when it is studied in isolation. Previous nuclear magnetic resonance (36) and crystallographic (37) data have indicated that the upper and lower stems of DV are essentially coaxial. However, in the intact intron, these two stems are at an angle of $\sim 45^\circ$ with respect to each other. This bend serves to move the backbones of the bulge and catalytic triad closer together, to a distance of ~ 3 Å. In addition to the angular orientation of the two helices, the bulge region of DV also differs from previous models in that it is extensively distorted by other intron components. This demonstrates that the entire group II intron structure is essential for inducing a catalytically relevant fold in DV.

DV is a metal ion-binding platform

The proximity of backbone moieties in the bulge and the catalytic triad results in the formation of a negatively charged pocket that binds two metal ions (M_1 and M_2) (Fig. 6). These two sites were assigned as Mg^{2+} ions on the basis of Yb^{3+} soaks. Yb^{3+} mimics the coordination state of Mg^{2+} and exhibits strong anomalous scattering, which allows its position to be easily identified within the structure.

Yb^{3+} was observed to bind at only four positions within the intron, three of which were located near the bulge of DV. Two of the Yb^{3+} sites correspond to electron density observed in the native map and were assigned as the aforementioned Mg^{2+} ions (Fig. 6). These two positions (M_1 and M_2) have very large anomalous signals of 75σ and 57σ . The other two sites do not correspond to any density in the native map and have smaller anomalous signals of only 7σ and 5σ . The third Yb^{3+} site is near the bulge, where it is located ~ 6 Å from M_1 and M_2 , while the fourth site is located outside the core of the intron altogether. The latter ions may represent low-occupancy Mg^{2+} sites or they may be specific only for Yb^{3+} .

Within the metal ion-binding pocket (Fig. 6), M_1 is bound by two inner-sphere ligands, consisting of the O1P of U375 and the O2P of C377. M_2 is bound by three inner-sphere contacts to the O1P of C358, the O2P of G359, and the O1P of C377. Previous biochemical studies on the $\alpha I5\gamma$ IIB intron have indicated that nonbridging phosphoryl oxygens at the corresponding DV positions are essential for both steps of splicing and, in at least one case, direct binding to metal ions (29, 38, 39), thereby providing functional evidence for the assignment of M_1 and M_2 as catalytic metal ions. The distance between metals M_1 and M_2 is 3.9 Å, which matches the ideal distance of 3.9 Å invoked for the classic two-metal ion reaction mechanism of the group I intron and protein phosphotransferases such as DNA and

RNA polymerase (40, 41). These two metal ions reside on the surface of DV and are therefore accessible to both the 5' and 3' splice sites (fig. S8).

Evidence that DIII is an allosteric effector

DIII contains a conserved, purine-rich internal loop that forms a curved, tightly wound RNA helix that is similar to loop motifs containing sheared G-A and A-A pairs (42). Density for many nucleobases in this region is weak, so a precise description of the pairing is not warranted. However, nucleotides in the conserved internal loop of DIII are in a position to form multiple base pairs and 2'-OH contacts with the basal regions of subdomain I(i) and DII. This is consistent with previous biochemical data indicating that DII helps position DIII (10). There is good base density for G320, indicating that it forms a trans sugar-edge/sugar-edge base pair (43) with G267, resulting in stacking of G320 upon A268 at the bottom of the DII stem. Thus, DIII is tied rigidly to the junction between DI and DII, which serves to position the adjacent J2/3 nucleotides within DV. Notably, a 140° kink in the backbone between the A290 of DIII and C289 of J2/3 (Fig. 5A) assists the insertion of J2/3 into DV. DIII also appears to stabilize the I(i) loop, which helps create a foundation for the active site through interactions with the z-anchor.

Enzymological studies have shown that DIII is not strictly required for catalysis, but it greatly accelerates the rate constants for splicing and for ribozyme reactions of the intron (44). The structure explains these observations by showing that DIII helps to organize the active site indirectly through its proximity to J2/3 and the 5' end of the intron. DIII is therefore an allosteric effector of catalysis and influences intron reactivity despite its distance from active-site moieties in DV.

Lack of electron density for domain VI

DVI contains the bulged adenosine that provides the 2'-OH nucleophile during lariat formation by group II introns (1). The crystal structure lacks electron density for DVI, even though this domain was included within the crystallization construct. Although DVI could not be visualized with the existing data, there is ample space for this domain within the structure, and it is likely to lie within the open cleft next to DV (fig. S8). This position would be consistent with biochemical evidence indicating that domains V and VI exist in a side-by-side arrangement (45).

Implications for catalysis

Group II introns have an absolute requirement for either Mg²⁺ (1) or Mn²⁺ (18) in order to catalyze both steps of splicing. The two-metal ion mechanism for RNA catalysis postulates that two divalent metal ions, located 3.9 Å apart, catalyze phospho-transfer reactions in self-splicing ribozymes (40). Crystallographic evidence indicates that this is the mechanism for group I intron splicing (41), and our structure suggests that a similar mechanism is also used in group II introns, consistent with biochemical data (46).

It is clear that the exon recognition machinery and the metal-binding pocket of DV are in close proximity within the ribozyme core. The 5' end of the intron and EBS1 are both

located near the DV bulge (Fig. 5A), thereby placing the 5' splice site in the active site. In addition, the γ nucleotide (A287), which pairs to the 3' terminal uridine (γ') of the intron, is located near the DV bulge (Fig. 5A), thereby positioning the 3' splice site for catalysis. The convergence of all these intron components on the two metal ions in DV and the close correspondence between biochemically determined contacts and the structure reported here suggest that this is a catalytically relevant structure. It is likely that this structure represents the “free retroelement” state of the intron that is capable of taking in substrate DNA for retrotransposition.

Evolutionary implications

The crystal structure provides a rationale for the strong phylogenetic conservation of DV. The constant separation of 5 base pairs (bp) between the catalytic triad and the bulge is required for the formation of the metal-binding platform. DV has remained conserved throughout higher organisms, and its basic form is used in the spliceosome. The intramolecular stem loop of the spliceosomal U6 small nuclear RNA (snRNA) also contains an AGC catalytic triad that is separated from a two-nucleotide bulge by 5 bp (47). Indeed, the U6 bulge binds metal ions in a manner similar to that of DV (48). DV can replace the U6atac snRNA in the U12-dependent spliceosome, providing functional evidence for parity between the DV and U6 motifs (47).

The exceptional conservation of J2/3 is now explained by the fact that it is an integral active-site motif, forming a catalytic triplex in the major groove of DV. This is evolutionarily relevant because J2/3 also has an apparent analog in the spliceosome (12): The phylogenetically invariant ACAGAGA box in U6 snRNA is refractory to mutagenesis and has been shown to interact near the bottom of the U6 stem (49). Thus, the spatial orientation of the AGA relative to the catalytic triad in U6 is very similar to that of J2/3 and DV in group II introns.

On the basis of these structural and functional analogies, it is most likely that the spliceosome is also a ribozyme and that it uses a two-metal ion mechanism for catalysis. These findings support the notion that group II introns evolved to colonize and shape the genomes of modern organisms. It has been suggested that the advent of spliceosomal introns from an ancestral group II intron resulted in the formation of the nuclear membrane and evolution of the eukaryotes (50). Introns also allowed eukaryotic genomes to breach the “one gene, one protein” barrier through alternative splicing. This crystal structure of a group II intron from an ancient lineage is therefore notable because it may represent a glimpse of the primordial catalyst that triggered the evolution of diverse life forms on Earth.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References and Notes

1. Pyle, AM.; Lambowitz, AM. *The RNA World*. 3. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY: 2006.

2. Zimmerly S, Guo H, Perlman P, Lambowitz A. *Cell*. 1995; 82:545. [PubMed: 7664334]
3. Robart A, Seo W, Zimmerly S. *Proc Natl Acad Sci USA*. 2007; 104:6620. [PubMed: 17420455]
4. Ferat J, Michel F. *Nature*. 1993; 364:358. [PubMed: 7687328]
5. Vallès Y, Halanych K, Boore J. *PLoS ONE*. 2008; 3:e1488. [PubMed: 18213396]
6. Sharp P. *Science*. 1991; 254:663. [PubMed: 1948046]
7. Cech T. *Cell*. 1986; 44:207. [PubMed: 2417724]
8. Jacquier A, Michel F. *Cell*. 1987; 50:17. [PubMed: 3297351]
9. Costa M, Michel F, Westhof E. *EMBO J*. 2000; 19:5007. [PubMed: 10990464]
10. Fedorova O, Mitros T, Pyle A. *J Mol Biol*. 2003; 330:197. [PubMed: 12823961]
11. Mikheeva S, Murray H, Zhou H, Turczyk B, Jarrell K. *RNA*. 2000; 6:1509. [PubMed: 11105751]
12. de Lencastre A, Pyle A. *RNA*. 2007; 14:11. [PubMed: 18039742]
13. Wank H, SanFilippo J, Singh R, Matsuura M, Lambowitz A. *Mol Cell*. 1999; 4:239. [PubMed: 10488339]
14. Schmidt U, Podar M, Stahl U, Perlman P. *RNA*. 1996; 2:1161. [PubMed: 8903346]
15. Boulanger S, et al. *Mol Cell Biol*. 1995; 15:4479. [PubMed: 7623838]
16. de Lencastre A, Hamill S, Pyle A. *Nat Struct Mol Biol*. 2005; 12:626. [PubMed: 15980867]
17. Toor N, Hausner G, Zimmerly S. *RNA*. 2001; 7:1142. [PubMed: 11497432]
18. Toor N, Robart A, Christianson J, Zimmerly S. *Nucleic Acids Res*. 2006; 34:6461. [PubMed: 17130159]
19. Rest J, Mindell D. *Mol Biol Evol*. 2003; 20:1134. [PubMed: 12777534]
20. Podar M, Chu V, Pyle A, Perlman P. *Nature*. 1998; 391:915. [PubMed: 9495347]
21. Granlund M, Michel F, Norgren M. *J Bacteriol*. 2001; 183:2560. [PubMed: 11274116]
22. Materials and methods are available as supporting material on *Science* Online.
23. Hendrickson W, Smith J, Sheriff S. *Methods Enzymol*. 1985; 115:41. [PubMed: 4079795]
24. Keel A, Rambo R, Batey R, Kieft J. *Structure*. 2007; 15:761. [PubMed: 17637337]
25. Harris-Kerr C, Zhang M, Peebles C. *Proc Natl Acad Sci USA*. 1993; 90:10658. [PubMed: 7504276]
26. Jacquier A, Michel F. *J Mol Biol*. 1990; 213:437. [PubMed: 2191139]
27. Boudvillain M, de Lencastre A, Pyle A. *Nature*. 2000; 406:315. [PubMed: 10917534]
28. Costa M, Déme E, Jacquier A, Michel F. *J Mol Biol*. 1997; 267:520. [PubMed: 9126835]
29. Boudvillain M, Pyle A. *EMBO J*. 1998; 17:7091. [PubMed: 9843513]
30. Costa M, Michel F. *EMBO J*. 1995; 14:1276. [PubMed: 7720718]
31. Fedorova O, Pyle A. *EMBO J*. 2005; 24:3906. [PubMed: 16252007]
32. Pyle A, Fedorova O, Waldsich C. *Trends Biochem Sci*. 2007; 32:138. [PubMed: 17289393]
33. Tamura M, Holbrook S. *J Mol Biol*. 2002; 320:455. [PubMed: 12096903]
34. Adams P, Stahley M, Kosek A, Wang J, Strobel S. *Nature*. 2004; 430:45. [PubMed: 15175762]
35. Podar M, et al. *RNA*. 1998; 4:151. [PubMed: 9570315]
36. Sigel R, et al. *Nat Struct Mol Biol*. 2004; 11:187. [PubMed: 14745440]
37. Zhang L, Doudna J. *Science*. 2002; 295:2084. [PubMed: 11859154]
38. Chanfreau G, Jacquier A. *Science*. 1994; 266:1383. [PubMed: 7973729]
39. Gordon P, Piccirilli J. *Nat Struct Biol*. 2001; 8:893. [PubMed: 11573097]
40. Steitz T, Steitz J. *Proc Natl Acad Sci USA*. 1993; 90:6498. [PubMed: 8341661]
41. Stahley M, Strobel S. *Science*. 2005; 309:1587. [PubMed: 16141079]
42. Baeyens K, De Bondt H, Pardi A, Holbrook S. *Proc Natl Acad Sci USA*. 1996; 93:12851. [PubMed: 8917508]
43. Leontis N, Westhof E. *RNA*. 2001; 7:499. [PubMed: 11345429]
44. Podar M, Dib-Hajj S, Perlman P. *RNA*. 1995; 1:828. [PubMed: 7493328]
45. Podar M, Perlman P. *RNA*. 1999; 5:318. [PubMed: 10024182]
46. Gordon P, Fong R, Piccirilli J. *Chem Biol*. 2007; 14:607. [PubMed: 17584608]

47. Shukla G, Padgett R. *Mol Cell*. 2002; 9:1145. [PubMed: 12049749]
48. Yean S, Wuenschell G, Termini J, Lin R. *Nature*. 2000; 408:881. [PubMed: 11130730]
49. Madhani H, Guthrie C. *Genes Dev*. 1994; 8:1071. [PubMed: 7926788]
50. Martin W, Koonin E. *Nature*. 2006; 440:41. [PubMed: 16511485]
51. www.pymol.org
52. We thank K. Rajashankar, N. Sukumar, and I. Kourinov of Northeastern Collaborative Access Team (NE-CAT) beamline ID-24 at the Advanced Photon Source (APS) of Argonne National Laboratory. We also thank J. Osipiuk and the staff at Structural Biology Center Collaborative Access Team (SBC-CAT) beamline ID-19 at APS, and C. Whalen, A. Héroux, A. Saxena, W. Shi, and H. Robinson at X25 and X29 at the National Synchrotron Light Source at Brookhaven National Laboratory. We thank J. Cochrane, M. Stahley, and P. S. Perlman for their advice and support. We thank S. A. Strobel, M. Stahley, J. Cochrane, J. Cabral, J. Li-Pook-Than, O. Fedorova, and G. P. Wagner for comments on the manuscript, R. Batey for the gift of iridium hexamine, and H. Takami for the *O. theyensis* strain. K.S.K. was supported by NIH training grant T15 LM07056. S.D.T. was supported by the U.S. Department of Defense through the National Defense Science and Engineering Graduate Fellowship Program and by a NSF Graduate Research Fellowship. This work was supported by the Howard Hughes Medical Institute (HHMI) and NIH grant GM50313 to A.M.P., who is an investigator of the HHMI. Coordinates, structure factors, and experimental phases of the group II intron have been deposited in the Protein Data Bank (accession number 3BWP, www.pdb.org) and the Nucleic Acid Database (accession number UR0130, <http://ndbserver.rutgers.edu>).

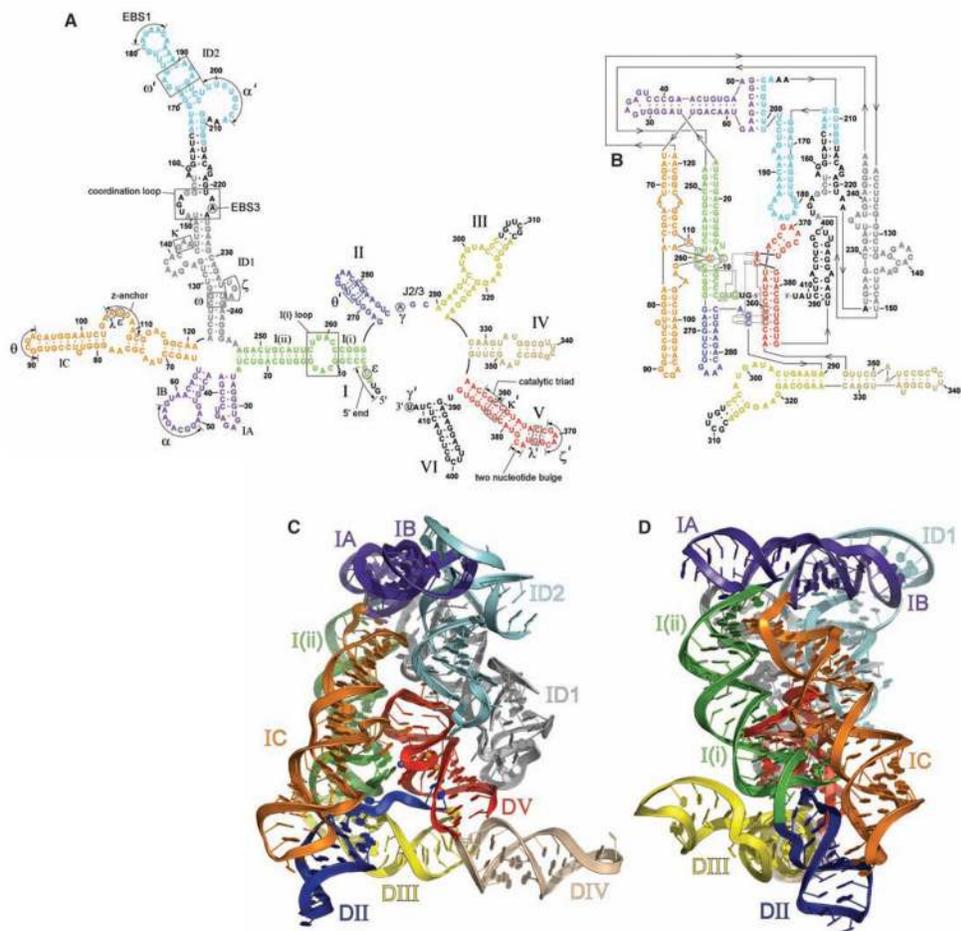


Fig. 1. Overall secondary and tertiary structure of the crystallized intron. **(A)** Secondary structure of the *O. iheyensis* intron in the canonical representation. Roman numerals and Greek letters indicate domains and tertiary interaction partners, respectively. The intron is depicted in the colors used for all subsequent figures. Disordered bases are shown in black. **(B)** Revised secondary structure reflecting the coaxial stacking and domain organization evident from the crystal structure. Watson-Crick tertiary interactions, base triples, and base-stacking contacts are shown only for the core of the intron and are represented as open circles, squares, and rectangles, respectively. **(C)** Overall structure of the intron in a ribbon representation. **(D)** A 90° rotation of the image shown in (C). Note that the UUCG tetraloop at the end of DIII was not well ordered, and bases are not shown. Figures were generated using PyMol (51).

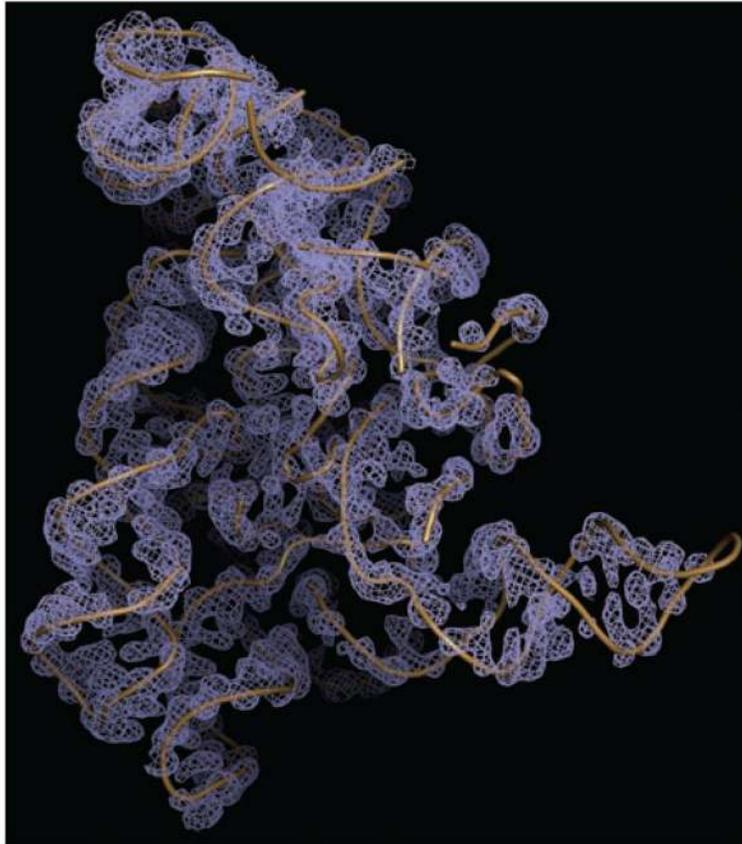


Fig. 2. Experimental, unbiased electron density map at 1.3σ . The structure is slightly rotated from Fig. 1C to better illustrate the quality of the map. A trace of the intron backbone is shown in beige.

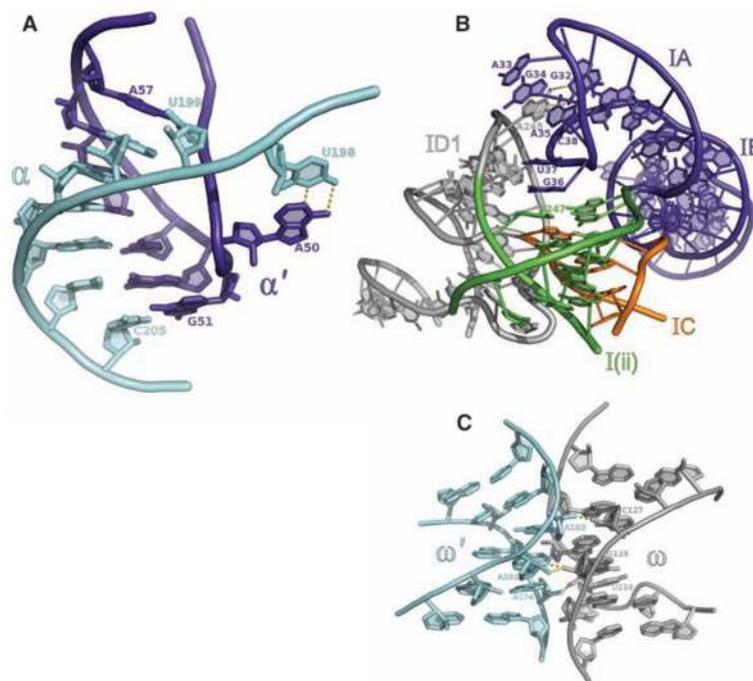


Fig. 3. Essential tertiary interactions within DI. **(A)** Kissing loop interaction between α and α' , stabilized by the extrahelical A50-U198 base pair. **(B)** The five-way junction in DI. Junction nucleotide A245 intercalates between G34 and A35 of stem IA and forms a trans Watson-Crick/sugar-edge base pair (43) with G32. Nucleotides G36 and U37 stack directly on top of A247. **(C)** The ω - ω' interaction, in which A192, A193, and A174 form a ribose zipper with C127, G128, and U238, respectively.

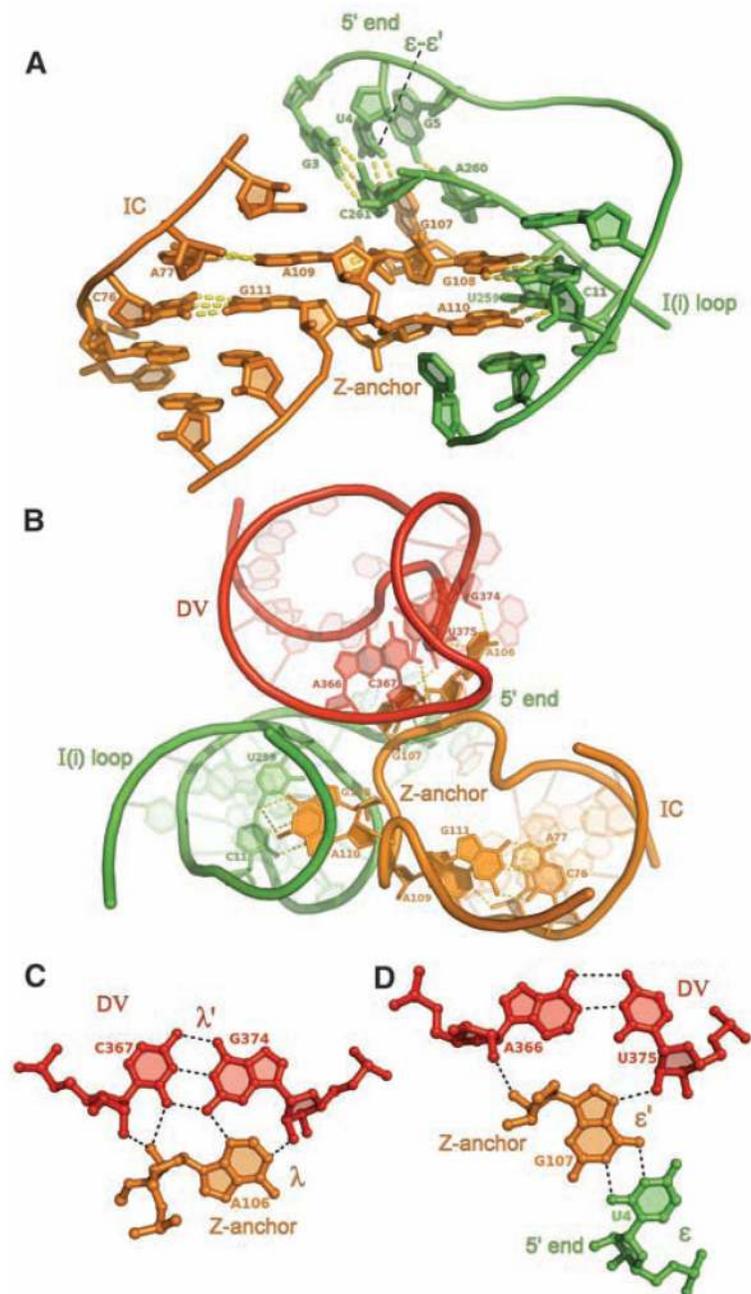


Fig. 4. The z-anchor and its network of interactions with DI(i) and DV. (A) Nucleotides 108 to 111 form an alternating “zig-zag” pattern which, together with 106 and 107, assemble the I(i) loop, the 5' end of the intron, and DV. (B) Top view of the z-anchor, illustrating its role in assembling multiple intron substructures. (C) Base triple between A106 and the C367-G374 base pair of DV (equivalent to λ - λ'). (D) Nucleotide quartet between G107, the 5' end, and DV.

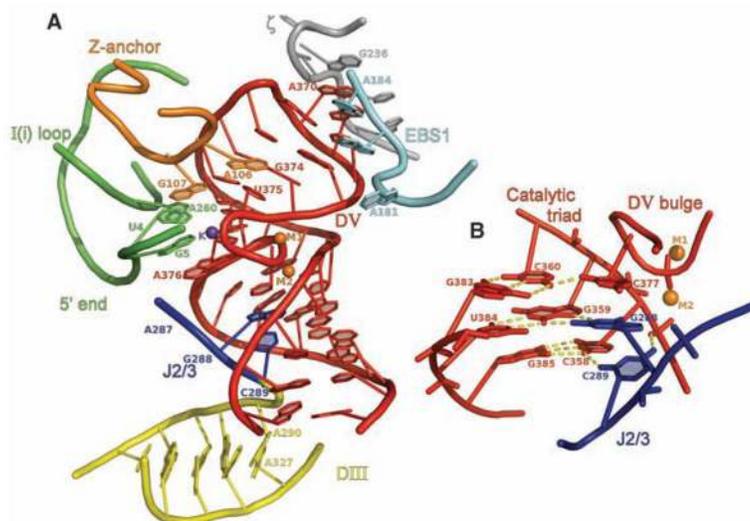


Fig. 5. DV and its local environment. **(A)** Interactions between DV, DI (including the z-anchor region and ζ - ζ'), and J2/3. For clarity, portions of DI (green and orange) are shown without bases. Only backbone density is visible for A287 (γ) and, thus, the base is not shown. Active-site Mg^{2+} ions are shown in orange and a K^+ ion is shown in purple. **(B)** The triple helix formed by J2/3, the catalytic triad, and the bulge of DV. Hydrogen bonds are shown as yellow dashed lines. C289 forms a base triple with C358-G385 through hydrogen bonding between the O2 and N4 of C289 and the N4 and O2P of C358, respectively. G288 forms a triple with G359-U384 via the N1 and N2 of G288 bonding with the N7 and O6 of G359. The final base triple involves the O2 and 2'-OH of C377 hydrogen bonding to the N4 and O2P of C360. J2/3 residues G288 and C289 are analogous to G588 and A589 in the $\alpha 15\gamma$ intron, respectively. In this work, O1P is the pro- S_p oxygen and O2P is the pro- R_p oxygen.

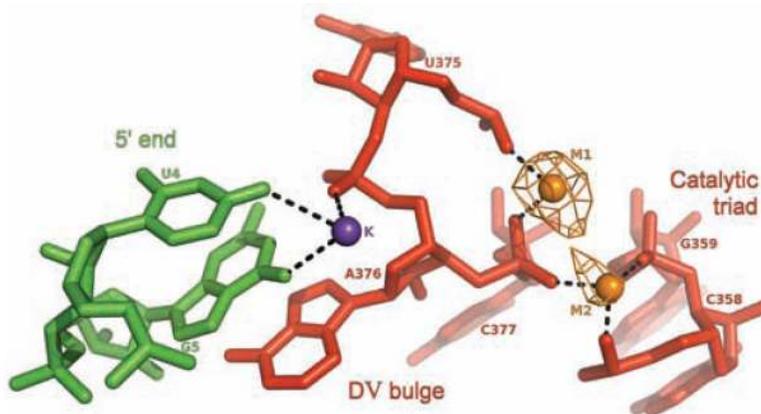


Fig. 6. Metals ions in the core of the intron. Yb³⁺ anomalous density is shown contoured at 42 σ for M₁ and M₂, which are spaced 3.9 Å apart. Additionally, a K⁺ ion interacts with the bulge and the bases of nucleotides 4 and 5. The K⁺ ion is located at a distance of 8 Å from M₁ and M₂. Black dashed lines indicate sites of inner-sphere coordination. Metal-binding residues C358, G359, C360, A376, and C377 from the *O. iheyensis* intron are analogous to residues A816, G817, C818, A838, and C839 from the *aI5* γ intron, respectively.