Crystal Structure of a Self-Spliced Group II Intron
Navtej Toor, et al.
Science 320, 77 (2008);
DOI: 10.1126/science.1153803

The following resources related to this article are available online at www.sciencemag.org (this information is current as of April 3, 2008):

Updated information and services, including high-resolution figures, can be found in the online version of this article at: http://www.sciencemag.org/cgi/content/full/320/5872/77

Supporting Online Material can be found at: http://www.sciencemag.org/cgi/content/full/320/5872/77/DC1

This article cites 48 articles, 20 of which can be accessed for free: http://www.sciencemag.org/cgi/content/full/320/5872/77#otherarticles

This article appears in the following subject collections:
Biochemistry
http://www.sciencemag.org/cgi/collection/biochem

Information about obtaining reprints of this article or about obtaining permission to reproduce this article in whole or in part can be found at: http://www.sciencemag.org/about/permissions.dtl
Crystal Structure of a Self-Spliced Group II Intron

Navtej Toor,1,* Kevin S. Keating,2 Sean D. Taylor,3 Anna Marie Pyle1,4

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 Oceanaobacillus 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-miRNAs (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 splicing 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.

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 (DIV 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 (DV) 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 Oceanaobacillus 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 P212121 with one molecule per asymmetric unit. The structure was solved by multiwavelength anomalous dispersion (23). Phases were independently calculated from both Yb3+ 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 Rwork of 27.6% and an Rfree of 31.0% (table S1). Density was observed for most of the intron nucleobases, although density for the k 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, α-α', 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 ε-ε' (Fig. 3C), places EBS1 near catalytically essential motifs in DV and posi-

---

**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).
tions 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 tetraloop 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 e-e′ 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 e-e′), 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 native map and were assigned as Mg2+ ions on the basis of Yb3+ soaks. The other two sites do not correspond to electron density observed in the structural data but indicate a metal ion site that 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 aI5 IIIB intron have indicated that nonbridging phosphoril 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 M1 and M2 as catalytic metal ions. The distance between metals M1 and M2 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).

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 DIII. 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 struc-
nature 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

catalysis and influences intron reactivity des-

pite its distance from active-site moieties in DV.

Lack of electron density for domain VI.

DV contains the bulged adenosine that pro-

vides the 2′-OH nucleophile during lariat forma-

tion 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 biochem-
ic 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 machin-
y 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 be-

tween 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 struc-
ture provides a rationale for the strong phy-
genetic conservation of DV. The constant sep-

aration of 5 base pairs (bp) between the cata-
ytic 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 splicesomal 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 func-
tional 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 func-
tional 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 splicesomal 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.

References and Notes
22. Materials and methods are available as supporting material on Science Online.
51. www.pymol.org
52. We thank K. Rajashankar, N. Sukumar, and J. 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. Osiipuk 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. 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 iodine heximide, and H. Takami for the G. ibexiensis 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 GMS0131 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://dsbserver.rutgers.edu).

Supporting Online Material
www.sciencemag.org/cgi/content/full/320/5872/77/DC1
Materials and Methods
Figs. 51 to 58
Table S1
References
5 December 2007; accepted 13 February 2008
10.1126/science.1153803

**REPORTS**

Revealing Magnetic Interactions from Single-Atom Magnetization Curves

Focko Meier,* Lihui Zhou, Jens Wiebe, † Roland Wiesendanger

The miniaturization of magnetic devices toward the limit of single atoms calls for appropriate tools to study their magnetic properties. We demonstrate the ability to measure magnetization curves of individual magnetic atoms adsorbed on a nonmagnetic metallic substrate with use of a scanning tunneling microscope with a spin-polarized tip. We can map out low-energy magnetic interactions on the atomic scale as evidenced by the oscillating indirect exchange between a Co adatom and a nanowire on Pt(111). These results are important for the understanding of variations that are found in the magnetic properties of apparently identical adatoms because of different local environments.

Magnetic nanostructures consisting of a few atoms on nonmagnetic substrates (adatoms) are explored as model systems for miniaturized data storage and spintronic devices and for the implementation of quantum computing. Because these structures are well defined and controllable on the atomic scale, they are ideally suited to study the fundamentals of magnetic interactions that are the ingredients of today’s and future memory and computation technology.

Since the early days of modern research in magnetism, the magnetization in response to an external magnetic field (a magnetization curve) has been recorded to gather information on the basic properties of magnetic samples (1). Such curves can be used to deduce the sample’s magnetic moment and magnetic anisotropy energy. The downscaling of samples from bulk over thin films and nanowires to nanodots requires an ever-increasing sensitivity of this method. It has been demonstrated that x-ray absorption spectroscopy with polarization analysis is able to measure magnetization curves of adatoms on a nonmagnetic substrate, albeit limited to large ensembles (2). Different approaches are potentially able to detect individual spins with nanometer spatial resolution ranging from magnetic resonance measurements (3) over magnetic exchange force microscopy (4) to scanning tunneling microscopy and spectroscopy (STM and STS) (5–10). Spin-averaged STS has been used to indirectly deduce the properties of single and coupled spins via the Kondo effect (5), the detection of noise (6, 7), or the observation of exchange splittings (8, 9). Inelastic electron tunneling has been adopted to measure the magnetic moments and anisotropy of individual atoms by spin-flip spectroscopy (10). This approach is complementary to the detection of magnetization curves but does not provide information on the dynamics of the spin and is so far restricted to adatoms on insulating layers. The method of choice for various substrates, spin-polarized STS (SP-STS), has been proven to detect single spins stabilized by direct exchange to