

## Review

# Group II introns: structure, folding and splicing mechanism

**Olga Fedorova<sup>1,2,\*</sup> and Nora Zingler<sup>2,\*</sup>**

<sup>1</sup>Howard Hughes Medical Institute, Yale University, New Haven, CT 06520, USA

<sup>2</sup>Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA

\*Corresponding authors

e-mail: olga.fedorova@yale.edu; nora.zingler@yale.edu

## Abstract

Group II introns are large autocatalytic RNAs found in organellar genomes of plants and lower eukaryotes, as well as in some bacterial genomes. Interestingly, these ribozymes share characteristic traits with both spliceosomal introns and non-LTR retrotransposons and may have a common evolutionary ancestor. Furthermore, group II intron features such as structure, folding and catalytic mechanism differ considerably from those of other large ribozymes, making group II introns an attractive model system to gain novel insights into RNA biology and biochemistry. This review explores recent advances in the structural and mechanistic characterization of group II intron architecture and self-splicing.

**Keywords:** catalytic RNA; protein-assisted splicing; ribozyme.

## Introduction

Self-splicing introns are large RNA molecules that are able to catalyze their own excision from a pre-mRNA while at the same time covalently joining the flanking exonic sequences. Group II introns represent one of the major classes of these autocatalytic RNAs, which also include group I and recently discovered group I-like ‘capping’ introns (Cech, 1990; Nielsen et al., 2005). They are found in bacteria and the organelles of various eukaryotes (mostly fungi and plants), but are absent from animal genomes. Their total number is difficult to gauge, but a recent database search yielded almost 600 mitochondrial group II introns (Lang et al., 2007), and a website dedicated to bacterial introns lists over 100 eubacterial and 14 archaeal group II introns (<http://www.fp.ucalgary.ca/group2introns/>). Despite their very diverse primary sequences, group II introns are defined by a highly conserved secondary structure. This generally consists of six domains radiating from a central wheel (Figure 1). Domain 1 (D1), the largest of all intronic domains, serves as a scaffold for assembly of the other domains and is indispensable for exonic substrate recognition (Qin and Pyle, 1998; Pyle and Lambowitz, 2006). D2 and D3 are not absolutely required for group II intron function (Koch

et al., 1992), but their presence enhances the catalytic efficiency (Qin and Pyle, 1998). D4 is the most variable region of the intron. In some introns it contains an open reading frame coding for a multifunctional intron-encoded protein (IEP) that facilitates splicing (maturase function) and copies and inserts the intron sequence into a new DNA target site (mobility function) (Qin and Pyle, 1998; Lambowitz and Zimmerly, 2004). D5 is the catalytic center of the intron, while D6 provides the bulged adenosine residue that serves as the branch point in the branching pathway of splicing (Qin and Pyle, 1998; Lehmann and Schmidt, 2003). Based on structural characteristics and their IEPs, group II introns can be further subdivided into families; the three main group II intron families are IIA, IIB and IIC (Michel et al., 1989; Toor et al., 2001).

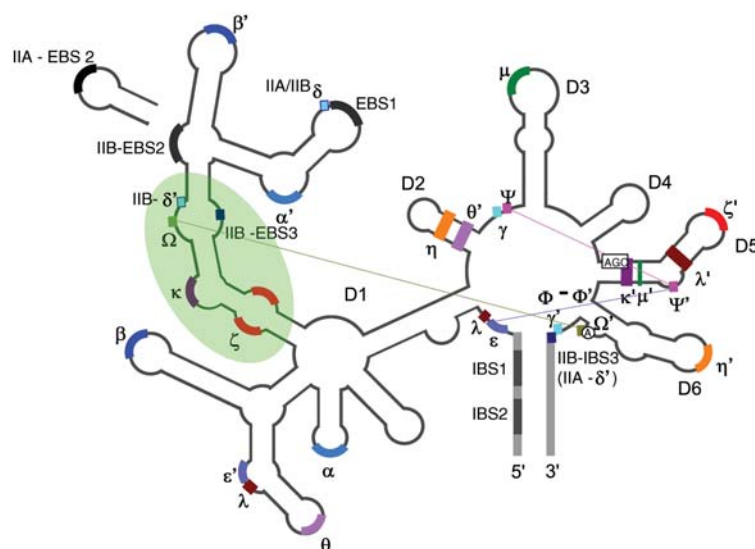
Similarities in the splicing mechanism suggest that group II introns and nuclear spliceosomal introns may share a common evolutionary ancestor (Michel and Ferat, 1995). In addition, their ability to act as autonomous mobile elements strongly suggests that group II introns are predecessors of modern non-LTR retrotransposons (Lambowitz and Zimmerly, 2004; Robart and Zimmerly, 2005). For an in-depth discussion of group II intron evolution and their role as mobile elements, we would like to direct the reader's attention to recent reviews on these topics (Lambowitz and Zimmerly, 2004; Robart and Zimmerly, 2005; Pyle and Lambowitz, 2006). In this article, we focus on the aspects of group II intron splicing *per se*.

## Splicing mechanisms

Group II introns generally excise from pre-mRNA as a lariat, a lasso-like structure that is also adopted by spliceosomal introns (van der Veen et al., 1986). This makes them a valuable model system to examine the splicing mechanism in detail (Jacquier, 1990). While *in vivo* studies have supplied important information on group II intron biology and mechanism, the most enlightening results were derived from biochemical analysis of their catalytic activity *in vitro*. In these experiments, group II introns need unphysiologically high salt concentrations and temperatures for efficient catalytic activity (see section ‘Folding of Group II introns’ below), but none of the available data suggest chemical or mechanistic differences between *in vitro* and *in vivo* reactivity once the intron is folded.

## Forward splicing

Group II intron splicing proceeds through two sequential transesterification reactions (Peebles et al., 1987; Jarrell et al., 1988). It is initiated by the 2'-hydroxyl group of a



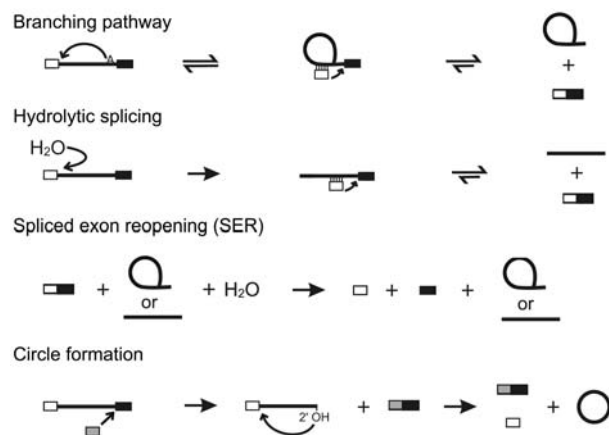
**Figure 1** Secondary structure and long-range tertiary contacts in a group II intron.

Most of the information is based on results obtained from the ai5γ intron. Long-range tertiary contacts are color-coded and labeled by small Greek letters. Proximity contacts identified by UV crosslinking are labeled by Greek capital letters and shown as dotted lines. The folding control element is highlighted by a green oval. Boxed letters denote the AGC triad in D5 and the branch-point adenosine in D6.

highly conserved bulged adenosine within D6. The nucleophile attacks the phosphate at the 5'-end of the intron, and in a typical  $S_N2$  reaction (Padgett et al., 1994) releases the 5'-exon while forming a lariat structure. This intermediate consists of an RNA circle with a 2'-5' linkage at the branch-site adenosine, and a 3'-tail still covalently attached to the 3'-exon. Since the released 5'-exon is bound to the intron through base pairing interactions, its free 3'-OH is then positioned correctly to attack the 3'-splice site in the second step of splicing. In the ensuing  $S_N2$  displacement reaction (Padgett et al., 1994; Podar et al., 1995b), the 5'- and 3'-exon are joined together, and the lariat intron is released (Figure 2).

### Reverse splicing

Both steps of splicing are reversible reactions (Augustin et al., 1990; Mörl and Schmelzer, 1990). The rate constants for the forward and reverse reactions of the first



**Figure 2** Mechanisms of splicing-related reactions catalyzed by group II introns.

Lines represent the intron RNA, rectangles the exons (white, 5'-exon; black, 3'-exon; gray, free 5'-exon in *trans*).

step of splicing are comparable (Chin and Pyle, 1995). It was hypothesized that one reason for this ready reversibility was an inherent proofreading mechanism (Chin and Pyle, 1995), but recent experiments with branched RNA molecules mimicking mis-spliced intermediates demonstrated that this is not the case (Wang and Silverman, 2006). Typically, the second step of forward splicing is much faster than the branching reaction, making the first step rate-limiting (Daniels et al., 1996). The reverse reaction is considerably slower, so that reverse splicing is typically an inefficient process. However, certain reaction conditions or subsequent coupled reactions (e.g., reverse transcription) can render the backward reaction rather efficient (Müller et al., 1991; Aizawa et al., 2003). Remarkably, however, this reaction is not limited to RNA substrates, but also works efficiently with DNA targets (Mörl et al., 1992; Griffin et al., 1995). This exceptional versatility in substrate choice is biologically relevant in intron mobility, in which reverse splicing into the target DNA is a crucial step of the homing reaction.

### Hydrolytic splicing

Early *in vitro* studies of group II intron splicing suggested that, in addition to the lariat splicing pathway, the intron can excise via an alternative pathway: water or a hydroxyl ion is used as a nucleophile in the first splicing step (van der Veen et al., 1987; Jarrell et al., 1988; Daniels et al., 1996). The second step then proceeds as in the branching pathway of splicing, and the products of this reaction are ligated exons and a linear intron (Figure 2). The balance between branching and hydrolytic splicing is strongly influenced by the choice of monovalent cation in the reaction (Daniels et al., 1996). Therefore, it was not clear at first whether the hydrolytic pathway was just an *in vitro* artifact, but in 1998 it was shown *in vivo* that introns with branch-point mutations retain splicing activity through this pathway (Podar et al., 1998a). The discovery of introns that naturally lack a branch-point

adenosine and are still active has revealed that hydrolytic splicing is an important, biologically relevant variation of group II intron splicing (Vogel and Börner, 2002).

### Unusual splicing reactions

While the classical lariat excision pathway is fairly well characterized, alternative and unusual splicing pathways and their significance *in vivo* have recently begun to attract more attention.

**Spliced exon reopening** *In vitro* characterization of the splicing reaction also led to the observation of an additional side reaction, spliced exon reopening (Jarrell et al., 1988; Daniels et al., 1996) (SER, Figure 2). Here the excised intron (either in its lariat or linear form) recognizes the ligated exons and hydrolyzes them exactly at the 5'-3' splice site junction. This reaction appears to be specific to group IIB introns; it is hardly observed in group IIA introns, even under optimal conditions (Schmidt et al., 1990; Hebbar et al., 1992). The biological relevance of this reaction is not yet clear, but it has been speculated that it might be involved in the generation of intron circles (see below).

**Circle formation** The formation of circles during splicing was first described over 20 years ago (Osiewacz and Esser, 1984): initially it was thought that circles rather than lariats were the product of splicing. With the discovery of lariat structures (van der Veen et al., 1986), circles were almost forgotten, but recently they have been rediscovered as byproducts of splicing reactions.

The first detailed characterization of group II intron circles utilized group II intron mutants lacking the branch-point adenosine (Murray et al., 2001). Both *in vitro* and *in vivo*, RNA circle formation was observed. Recently, however, several wild-type introns in both bacteria (Molina-Sanchez et al., 2006) and plant mitochondria (Li-Pook-Than and Bonen, 2006) have been shown to naturally form circular splicing products *in vivo*, suggesting that circle formation is not a random artifact of a few intron species, but a conserved side reaction with an unknown function.

The circular molecules usually consist of the complete intron sequence and seem to be linked by a 2'-5' bond at the cyclization junction (Murray et al., 2001; Molina-Sanchez et al., 2006). Therefore, they must be generated by a mechanism different from the well-established cyclization mechanism of group I introns, which takes place after splicing, cleaves off the 5'-end of the intron and leads to a 3'-5' linkage at the cyclization site (Zaug et al., 1983). The proposed mechanism for group II introns (Murray et al., 2001) involves attack of the 3'-junction by a free 5'-exon (possibly generated by SER) *in trans*. The resulting 5'-exon/intron intermediate undergoes an intramolecular transesterification reaction, with the 2'-hydroxyl group of the 3'-terminal nucleotide attacking the 5'-splice site (Figure 2). However, intron RNA circles could also be derived from linear excision products generated by the hydrolytic pathway and then circularized by a host-encoded RNA ligase. This notion is corroborated by the observation of intron circles that have non-encod-

ed nucleotides or missing nucleotides at the supposed circularization site (Li-Pook-Than and Bonen, 2006).

Intriguingly, when RNA intron circles are observed *in vivo*, DNA intron circles can usually also be detected (Murray et al., 2001; Molina-Sanchez et al., 2006), although it is still unknown how they are produced and what their function might be. In this context, it is interesting that circular intron DNA molecules in *Podospira anserina* have been shown to increase in abundance in senescent cells (Osiewacz and Esser, 1984).

**Aberrant/alternative splicing** Recent findings have identified the possibility that the fidelity of splicing is not always as high as observed for the well-characterized introns ai5 $\gamma$  and LtrB. For example, the bacterial group IIB3 intron Rmlnt1 displays unexpected *in vitro* activities that are thought to originate from interaction of exon binding site 1 (EBS1) with an alternative intron binding site 1 (IBS1) a few bases downstream of the 3'-intron/exon junction (Costa et al., 2006a). In addition, EBS2 of the same intron seems to be capable of interacting with both the canonical IBS2 and an adjacent alternative upstream site, a feature conserved in related bacterial introns (Costa et al., 2006b). The bacterial IIC intron B.h.11 from *Bacillus halodurans* represents the first example of a group II intron using 5'-splice sites upstream of the canonical intron structure (Toor et al., 2006). As a caveat it should be mentioned that the reactions described above were all conducted *in vitro*, and so far have not been observed *in vivo*. They might therefore be artifacts only occurring under unusual salt conditions and in the absence of the cognate IEP. However, intron B.c.14 from *Bacillus cereus* naturally splices 56 nucleotides downstream of the expected canonical 3'-splice site (Tourasse et al., 2005). There is a high probability that more unusual splicing reactions will be found as more and more introns are biochemically characterized.

## Structural organization of group II introns

Each of the six intronic domains has a specific role in folding, conformational rearrangements or catalysis. The native conformation of a group II intron is sustained by intra- and interdomain long-range tertiary interactions, which are critical either for folding of the intron to the native state or for its catalytic activity. The specific functions of the intronic domains and all identified tertiary contacts are discussed below.

### Domain 1

It has been long known that D1 is absolutely essential for catalysis (Michel et al., 1989; Michel and Ferat, 1995). It serves as a scaffold for assembly of other domains into the catalytically active structure. D1 is indispensable for exonic substrate recognition and is involved in several tertiary interactions critical for catalytic function (Qin and Pyle, 1998; Pyle and Lambowitz, 2006).

**Exonic substrate recognition and splice site selection** Typically, D1 contains two 5'-exonic substrate recognition sequences (EBS1 and EBS2), which interact with

corresponding complementary regions at the 3'-end of the 5'-exon (IBS1 and IBS2) and help to define the 5'-splice site (Qin and Pyle, 1998; Pyle and Lambowitz, 2006). The mechanism of 5'-splice site recognition and substrate specificity of group II introns has been studied in detail by taking advantage of group II intron modularity. A series of ribozyme constructs comprising different intron domains was provided *in trans* to a short RNA oligo containing the last 17 nucleotides of the 5'-exon (including both IBS1 and IBS2) and the first seven nucleotides of the intron (Michels and Pyle, 1995; Xiang et al., 1998; Su et al., 2001). Fluorescence studies have demonstrated that an energetic penalty is paid upon binding of the 5'-exonic substrate by D1, which greatly increases the substrate specificity of group II ribozymes (Qin and Pyle, 1999). Mutational analysis of EBS1-IBS1 and EBS2-IBS2 interactions revealed that even a single mismatch may result in a significant decrease in substrate hydrolysis efficiency (Xiang et al., 1998). It has been hypothesized that the cleavage site is determined by the nucleotide paired with the 5'-terminal residue of EBS1 (G329 in the  $\alpha 5\gamma$  intron) (Jacquier and Jacquesson-Breuleux, 1991). (Please note that nucleotide numbers refer to the  $\alpha 5\gamma$  intron.) However, further studies showed that, instead of recognizing specific nucleotides at the cleavage site, the EBS1 region attempts to form the most thermodynamically stable duplex with the exonic substrate. Cleavage then occurs at the junction between the single-stranded and double-stranded regions (Su et al., 2001). Interestingly, some group II introns (class IIC) do not have an EBS2 site and appear to rely solely on EBS1 for 5'-splice site recognition (Toor et al., 2001; Toor et al., 2006). These ancient bacterial introns are generally located downstream of transcriptional terminator motifs, and for at least one of the IIC introns (B.h.I. 1 from *Bacillus halodurans*) the stem-loop terminator motif participates in defining the 5'-splice site, thus partially compensating for the absence of EBS2-IBS2 interaction.

In addition to 5'-exonic substrate recognition sites, D1 also contains elements that contribute to recognition of the 3'-exonic substrate by interacting with the first nucleotide of the 3'-exon [EBS3-IBS3 interaction (subgroup IIB) or  $\delta$ - $\delta'$  interaction (subgroup IIA), Figure 1] (Jacquier and Jacquesson-Breuleux, 1991; Costa et al., 2000).

#### **Tertiary contacts involving D1 that are critical for catalysis**

Phylogenetic analysis and nucleotide analog interference suppression (NAIS) studies identified three tertiary interactions between D1 and the 'catalytic center' of the intron, D5:  $\zeta$ - $\zeta'$  and  $\kappa$ - $\kappa'$ , which are important for D5 docking (Costa and Michel, 1995; Boudvillain and Pyle, 1998), and  $\lambda$ - $\lambda'$ , which positions D5 in close proximity to the 5'-splice site and is directly involved in catalysis (Boudvillain et al., 2000; Figure 1). Multiple studies have demonstrated that the intradomain interaction  $\varepsilon$ - $\varepsilon'$  (Figure 1), which was identified by phylogenetic analysis, is also important for the catalytic function of the intron (Jacquier and Michel, 1987, 1990).

Interestingly, the  $\varepsilon$ - $\varepsilon'$  nucleotides exhibit strong nucleotide analog interference effects in the sugar-phosphate backbone, suggesting that there may be a third com-

ponent of this interaction elsewhere in the intron (Boudvillain and Pyle, 1998). Another phylogenetically identified long-range tertiary interaction,  $\alpha$ - $\alpha'$ , has been implicated in facilitating binding of the 5'-exonic substrate (Qin and Pyle, 1998); however, it also plays an important role in intron folding (Waldsich and Pyle, 2007). It has been suggested that the tetraloop-receptor interaction  $\theta$ - $\theta'$  (Figure 1) plays a role in structural stabilization of the native structure rather than being directly involved in catalysis (Costa et al., 1997). This interaction is also important for recruiting the catalytic effector D3 and the phylogenetically conserved interdomain joiner J2/3 into the active site (Fedorova et al., 2003). However, the precise nature of its interaction with D3 or J2/3 remains unknown.

**Docking site for D6** D6 is critical for the branching pathway of the first step of splicing, which is common to both group II intron and spliceosomal splicing (Lehmann and Schmidt, 2003; Pyle and Lambowitz, 2006). Both branch-point recognition and splice-site selection are very accurate (Chu et al., 2001), suggesting that the branch-point adenosine must be precisely positioned in the active site of the intron. Surprisingly, neither phylogenetic analysis nor NAIM/NAIS studies have been successful in determining local surroundings of the branch-point adenosine in the ribozyme active site. This was achieved by site-specifically incorporating crosslinkers at or near the branch point in chemically synthesized RNAs containing D5 and D6, and reacting them with the rest of the intron (exD123 RNA) (Hamill and Pyle, 2006). The docking site for the branch point has been identified in the asymmetric loop in D1 ('coordination loop') (Hamill and Pyle, 2006) (Figure 1), which also contains residues involved in positioning of the 3'-splice site (EBS3-IBS3 interaction; Costa et al., 2000). Mutational analysis confirmed that the coordination loop is important for branching (Hamill and Pyle, 2006), but the exact nature of the interaction between this substructure and the branch point remains enigmatic. It has been hypothesized that the branch point is recognized not by means of conventional hydrogen bonding, but via a network of van der Waals' contacts and stacking interactions that are sensitive to the shape of the branch-site region (Liu et al., 1997; Hamill and Pyle, 2006). A similar hypothesis has been proposed for spliceosomal branch-site recognition by protein components of the spliceosome (Schellenberg et al., 2006; Spadaccini et al., 2006).

#### **Structural elements in D1 that are critical for intron folding**

Over the past 10 years, multiple studies have demonstrated that D1 folding is independent of other intronic domains (Qin and Pyle, 1997). Furthermore, recent studies have shown that folding of D1 is a rate-limiting step in the folding of the entire intron (Su et al., 2005). Strikingly, NAIM studies using group II ribozyme compaction as a selection step have conclusively demonstrated that most of the functional groups critical for intron compaction and folding are clustered in a relatively small substructure designated as a folding control element (Waldsich and Pyle, 2007). This substructure harbors docking sites for D5 ( $\kappa$ - and  $\zeta$ -elements) and D6 (coordination loop) (Figure 1), implying that compaction

and folding of the intron hinges on proper formation of the docking site for the catalytic domain, thus ensuring specificity and accuracy of group II ribozyme catalysis (Waldsich and Pyle, 2007). In addition, the formation of  $\alpha$ - $\alpha'$  and  $\beta$ - $\beta'$  contacts are important for ribozyme compaction (Waldsich and Pyle, 2007), although it was previously shown that  $\beta$ - $\beta'$  is dispensable for group II intron function (Chu, 2000).

**Protein-binding sites** The i-stem of D1 in some group II introns contains a large insertion, hypothesized to be a potential protein-binding site (Adamidi et al., 2003). It has recently been shown that this insertion in maize group II intron *atpF* is indeed a binding site for the protein co-factor CRS1 (Figure 4), suggesting that this may also be the case for other group II introns with a similar feature (Ostersetzer et al., 2005). Chemical probing data have identified potential LtrA maturase-binding sites in D1 of the *L. lactis* LtrB intron. It is proposed that these are located in the c2 and d4 substructures and represent a secondary binding site for this protein, with the primary binding site located in D4 (Figure 4; Matsuura et al., 2001).

## Domain 2

D2 forms two essential long-range tertiary contacts with D1 ( $\theta$ - $\theta'$ , see above) and D6 ( $\eta$ - $\eta'$ ) (Chanfreau and Jacquier, 1996; Costa et al., 1997). The latter is a tetraloop-receptor interaction that is structurally conserved between different IIA and IIB introns (Costa et al., 1997). Interestingly, in group IIA introns the tetraloop is located in D2 and the receptor in D6, whereas in group IIB introns the location of the tetraloop and the receptor is reversed: the tetraloop is in D6 and the receptor is in D2 (Costa et al., 1997). It has been proposed that the  $\eta$ - $\eta'$  interaction serves as a conformational switch between the two steps of splicing (Chanfreau and Jacquier, 1996). However, recent UV crosslinking studies have suggested that all intronic elements important for both steps of splicing are located in close proximity before the first step occurs, thus implying that both steps likely have a single active site and that no significant conformational rearrangement between the two steps is required (de Lencastre et al., 2005). Since little is known about the functionalities required for the second step of splicing, further studies are necessary to shed more light on this issue.

## Domain 3

D3 is generally referred to as a catalytic effector (Qin and Pyle, 1998; Pyle and Lambowitz, 2006). It is not strictly required for catalysis (Koch et al., 1992), but its presence remarkably enhances reaction rates of group II-derived ribozyme constructs (Griffin et al., 1995; Xiang et al., 1998; Fedorova et al., 2003). D3 binds the rest of the intron with relatively high affinity (Podar et al., 1995a), and some of its regions are highly protected against hydroxyl radicals (Swisher et al., 2001), suggesting that it forms an extensive network of tertiary interactions with the rest of the intron. However, there is no phylogenetic co-variation between D3 and other domains, suggesting that

these contacts do not involve Watson-Crick base-pairing or simple and easily identifiable tetraloop-receptor interactions. Early chemical probing and modification interference studies suggested that the basal stem, internal bulge and pentaloop of D3 are important for catalysis and may be involved in tertiary contacts with D5 and other domains (Jestin et al., 1997). The first tertiary contact between D3 and D5 ( $\mu$ - $\mu'$ ) has recently been identified by NAIS analysis (Fedorova and Pyle, 2005). This interaction, which is very different from a conventional tetraloop-receptor interaction, involves two adenosine residues from the D3 pentaloop and the 2'-hydroxyl group of G844 in D5, which was previously shown to be important for catalysis (Fedorova and Pyle, 2005). The most conserved part of D3 is the internal bulge (Michel and Ferat, 1995). Recent NAIM studies have suggested that bulged nucleotides form tandem sheared *trans*-Hoogsteen-sugar edge base pairs, which may be involved in a tertiary contact (Fedorova and Pyle, 2005). This non-canonical pairing is phylogenetically conserved and may be part of a larger structural motif. However, its potential interaction partner is still unknown.

## Domain 4

In some introns, D4 contains an open reading frame coding for the maturase protein that facilitates intron splicing under physiological conditions and is required for intron mobility (Lambowitz and Zimmerly, 2004). D4 also contains the primary binding site for the maturase protein (Matsuura et al., 2001; Singh et al., 2002; Watanabe and Lambowitz, 2004) and other intron-specific splicing factors (Ostersetzer et al., 2005). As it is located on the surface of the folded intron, it likely plays a more general role as a protein-binding element and may interact with various protein co-factors, facilitating intron splicing and mobility.

## Domain 5

D5 is the second intronic domain absolutely required for group II catalysis besides D1. This small (34 nt) stem-loop structure is the most phylogenetically conserved part of the entire intron (Michel and Ferat, 1995). It contains many functionalities critical either for its binding to D1 or for catalysis (Chanfreau and Jacquier, 1994; Boulanger et al., 1995; Costa and Michel, 1995; Peebles et al., 1995; Abramovitz et al., 1996; Schmidt et al., 1996; Boudvillain and Pyle, 1998; Konforti et al., 1998; Boudvillain et al., 2000). The atoms and functional groups involved in D5 docking lie on one side of the molecule (binding face), and functionalities involved in catalysis are located on the chemical face of D5 (Abramovitz et al., 1996). D5 docks to the D1 scaffold via two tetraloop-receptor interactions:  $\zeta$ - $\zeta'$  (Costa and Michel, 1995) and  $\kappa$ - $\kappa'$  (Boudvillain and Pyle, 1998), both of which involve the binding face of D5. In two major classes of group II introns, IIA and IIB,  $\zeta'$  is a canonical GNRA tetraloop (Michel and Ferat, 1995; Toor et al., 2001). However, in IIC introns it is an unusual GAAC tetraloop. The  $\zeta$ -receptor in IIC introns is also non-canonical (Toor et al., 2001, 2006), suggesting that the  $\zeta$ - $\zeta'$  interaction in IIC introns does not belong to any known type of tetraloop-receptor

interaction. Another D1-D5 interaction ( $\lambda$ - $\lambda'$ ) brings the chemical face of D5 and the 5'-splice site together (Boudvillain et al., 2000). It was also recently shown that D5 directly interacts with the catalytic effector D3 via  $\mu$ - $\mu'$  contact, which possibly helps to anchor D3 in the catalytic core of the ribozyme (Fedorova and Pyle, 2005). One of the regions in D5 which is extremely important for catalysis is the dinucleotide bulge (Schmidt et al., 1996). Terbium cleavage studies suggested that it harbors a magnesium ion-binding site that may be important for catalytic activity (Sigel et al., 2000). Importantly, a magnesium binding site at the same location was also found in the structure of U6 RNA, and is believed to contribute to spliceosomal catalysis (Yean et al., 2000).

Neither phylogenetic analysis nor NAIS studies were successful in identifying its interaction partner. UV cross-linking data suggest that this area lies in close proximity to the two most phylogenetically conserved residues (G588 and A589) in the linker between D2 and D3 (J2/3) and the  $\epsilon$  region (nucleotide C4), which are also important for the catalytic function of the intron (de Lencastre et al., 2005). Recent NMR studies by Sigel et al. (2004) suggest that the dinucleotide bulge forms a new RNA folding motif: U823 is paired neither with A838, as suggested in early publications (Michel et al., 1989), nor with G840, as first proposed by Michel and coworkers (Costa et al., 1998) and later by crystallographic studies (Zhang and Doudna, 2002). Instead, the four nucleotides of the bulge (U823, A838, C839 and G840) are unpaired and three of the four form stacking interactions with the helical nucleotides. The fourth, G840, has an unusual *syn*-conformation and is flipped out into the major groove, which would make it accessible for long-range tertiary interactions with other intronic elements. The differences between the NMR and X-ray structures in this case may be attributed to the bulge nucleotides forming intermolecular lattice contacts in the crystal. However, since both studies deal with free D5 without its interaction partners within the intron, it still remains to be confirmed whether either of the two proposed conformations is adopted in the context of the fully formed active site.

One of the most conserved regions in D5 is the AGC triad, also frequently referred to as the 'catalytic triad' (Figure 1), in which G is an invariant residue critical for both *in vivo* and *in vitro* splicing (Boulanger et al., 1995; Peebles et al., 1995; Konforti et al., 1998). X-Ray studies suggested that this region is dynamic and somewhat disordered (Zhang and Doudna, 2002); however, the NMR structure does not support this observation (Sigel et al., 2004). This trinucleotide has also been shown to harbor a magnesium-binding site, which has been proposed to mediate D5 binding to D1 (Gordon and Piccirilli, 2001; Sigel et al., 2004). The AGC triad is a feature that group II introns share with U6 snRNA from the spliceosome. However, the precise role and interaction partners of the catalytic triad are still unknown.

#### Domain 6

D6 provides the bulged adenosine residue (Figure 1) that serves as a branch point in the branching pathway of splicing that is predominant *in vivo* (Lehmann and Schmidt, 2003; Pyle and Lambowitz, 2006). Multiple

studies have demonstrated that branch-site selection by group II introns is generally very precise. For example, introducing a base-paired rather than a bulged A does not result in cryptic branching (Chu et al., 1998). Paradoxically, however, this important domain does not appear to be phylogenetically conserved (except for the branch-point adenosine) (Michel et al., 1989; Michel and Ferat, 1995; Toor et al., 2001). Recent NMR studies have shown that the branch-point adenosine is not flipped out, but resides within the helical structure, where it is partially stacked between flanking GU pairs (Erat et al., 2007), although this is somewhat controversial (Schlatterer et al., 2006). Two studies report that mutating the branch-point adenosine to a different nucleotide results in a drastic decrease in splicing efficiency, although there appears to be a disagreement on whether such mutations result in cryptic branching (Gaur et al., 1997; Liu et al., 1997). Systematic chemical mutagenesis of the branch-point adenosine followed by kinetic assay in a *cis*- or *trans*-branching construct (Liu et al., 1997) suggested that the ribozyme recognizes an exocyclic amino group of the adenine residue and sterically excludes non-adenine features of other nucleobases (Liu et al., 1997). Further studies have demonstrated that the exceptional accuracy of branch-site selection by group II introns is ensured by a combination of several partially redundant structural determinants, including the 4-bp basal stem of D6, the 3-nt linker between D5 and D6 in IIB introns, and a G-U pair upstream of the branch-point adenosine (Chu et al., 2001). None of these features is absolutely required for branching accuracy, but their combination guarantees proper branch-point selection (Chu et al., 2001). Redundancy of structural determinants for branch-site selection may explain why group II introns are more specific than spliceosomal splicing (Ruskin et al., 1985; Query et al., 1994).

#### Domain 7

Recently an unusual type of group II introns has been identified in *Bacillus cereus*. The B2-like intron B.c.I4 splices 56 nucleotides downstream of the expected canonical 3'-splice site (Tourasse et al., 2005). Notably, in this case the 'canonical' 3'-splice site (three nucleotides after D6) can form neither  $\gamma$ - $\gamma'$  nor EBS3-IBS3 tertiary contacts, but both interactions are formed by the actual downstream 3'-splice site. However, branching still occurs at the normal site in D6 (Stabell et al., 2007), which is in good agreement with previous results suggesting that selection of the branch site and choice of the proper 3'-splice site are decoupled. The extra 56 nucleotides form a stem-loop structure that does not seem to have an important role in splicing. The intron appears to have adapted to the extra substructure, which could be referred to as D7.

#### Interdomain joiners

Most nucleotides comprising the central wheel of a group II intron are highly conserved and play an important role in group II intron catalysis, although their function is not completely understood. A combination of phylogenetic and mutational analyses, as well as NAIM and NAIS stud-

ies, has demonstrated the importance of the nucleotides at the beginning of the intron: residues G3, C4 and G5 have been shown to be involved in catalytically critical  $\varepsilon$ - $\varepsilon'$  and  $\lambda$ - $\lambda'$  tertiary contacts, respectively (see above). Multiple studies underscored the important role of joiner J2/3 between D2 and D3 in group II catalysis (Podar et al., 1998b; Mikheeva et al., 2000; Fedorova et al., 2003; de Lencastre et al., 2005; Fedorova and Pyle, 2005). It harbors two of the most conserved nucleotides in the entire intron, G588 and A589, which have been proposed to be important for both the first step and the second step of splicing (Mikheeva et al., 2000; Fedorova and Pyle, 2005). Recent studies have placed these two nucleotides in proximity with the catalytically critical 2-nt bulge of D5 (de Lencastre et al., 2005). A587 is involved in a phylogenetically conserved tertiary contact with the last nucleotide of the intron ( $\gamma$ - $\gamma'$ ), which is particularly important for the second step of splicing and accurate 3'-splice site selection (Jacquier and Michel, 1990; Jacquier and Jacquesson-Breuleux, 1991). Intriguingly, the catalytic function of J2/3 strictly depends on its physical connection to the basal stem of D2, which implies that the D2 stem may facilitate formation or positioning of the J2/3 structure (Fedorova et al., 2003).

The 3-nt linker between D5 and D6 is one of the determinants for branch-site selection by group IIB1 introns (Chu et al., 2001). This feature is shared by IIC introns, but is absent in IIA and chloroplast-like class 2 introns (Toor et al., 2001). Either it is not a branch-site selection determinant in the latter two groups, or there are compensatory changes in the branch-point docking site.

### Structural differences of group II intron families

Although all group II introns share the basic six-domain architecture, there are characteristic differences defining families IIA, IIB and IIC. The IIA and IIB families differ in the location of the EBS2 site, the length and composition of interdomain joiners, and the positioning of tertiary interactions involved in recognition of the 3'-splice site. The latter function is fulfilled by the  $\delta$ - $\delta'$  contact in group IIA introns and the EBS3-IBS3 interaction in the group IIB family (Figure 1) (Toor et al., 2001). In addition, the two families have some differences in secondary structure (e.g., in IIB introns the  $\varepsilon'$  and  $\lambda$  nucleotides are located in a small asymmetric bulge, but they are part of a much larger single-stranded region in IIA introns) (Toor et al., 2001).

The IIC family is believed to be the most ancient and is possibly ancestral to other group II families (Pyle and Lambowitz, 2006). Introns from this family are smaller and differ significantly from IIA and IIB families: the most conserved element of any group II intron, D5, is lacking two base pairs, has a different 5'-sequence, and is capped by an unconventional tetraloop. In addition, IIC introns are missing some of the common group II intron features, such as the C2 substructure in D1 or the EBS2-IBS2 interaction, and they have a very different  $\varepsilon'$ - $\lambda'$  region compared to other families (Toor et al., 2001). There are also substantial differences in 5'-splice site recognition by IIC introns in comparison with the IIA and IIB families. For a detailed phylogenetic analysis and

classification of group II introns, see Michel et al. (1989), Michel and Ferat (1995), and Toor et al. (2001).

### Folding of group II introns

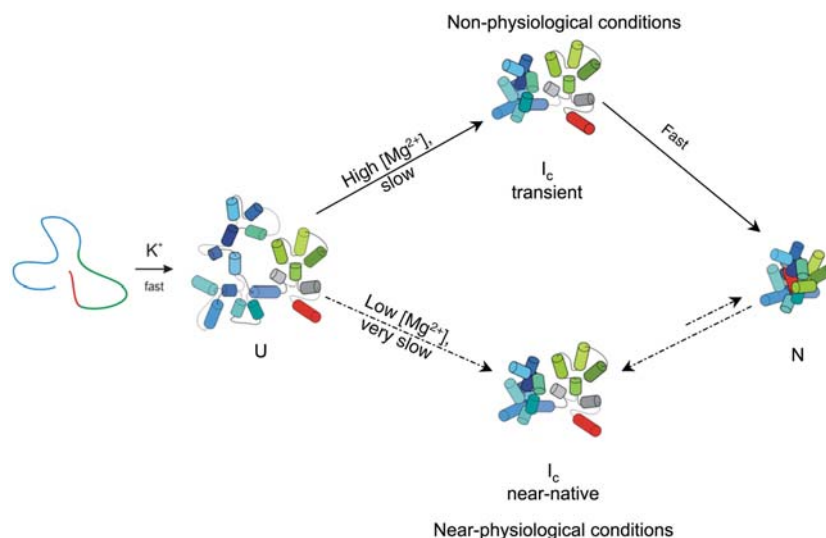
To properly perform all catalytic functions, group II intron RNAs must adopt the correct native conformation. Folding of group II introns has been extensively studied using the ai5 $\gamma$  intron-derived ribozyme D135 (Su et al., 2001) as a model system (Swisher et al., 2001, 2002; Su et al., 2003, 2005). For a detailed review about the group II intron folding mechanism in comparison to other RNAs, see Pyle et al. (2007). Here we concentrate on the most recent findings.

### Folding under conditions optimal for *in vitro* catalysis

Monitoring the evolution of hydroxyl radical footprints as a function of magnesium ion concentration or time suggested that solvent-inaccessible regions consistent with the established native tertiary contacts are formed in a concerted manner. The rate of their formation is equal to the rate of global compaction of the D135 ribozyme, which was determined by analytical ultracentrifugation and native gel analysis (Swisher et al., 2002; Su et al., 2005). In addition, urea denaturation studies have shown that unfolding and folding of the D135 ribozyme have similar thermodynamic parameters and that unfolding from the native state is fully reversible (Swisher et al., 2002; Su et al., 2003). These data suggest that the D135 ribozyme folds by an apparent two-step mechanism. Further folding studies with other ai5 $\gamma$ -derived ribozyme constructs have shown that folding of D1 is a rate-limiting step in folding of the D135 RNA, and that it requires as much Mg<sup>2+</sup> as folding of the whole D135 RNA (Su et al., 2005). Taken together, these results suggest the formation of a transient on-pathway intermediate in D1 (Figure 3).

It has long been thought that folding of a large RNA is unavoidably hindered by formation of stable misfolded structures, usually referred to as 'kinetic traps' (Treiber et al., 1998). Acceleration of folding to the native state in the presence of subdenaturing concentrations of urea is generally considered a hallmark of a kinetic trap (Treiber and Williamson, 1999). In this case, the rate-limiting step in the folding pathway is escape from a kinetic trap, and folding is slow (Treiber and Williamson, 1999). This phenomenon has been observed for both *Tetrahymena* and RNase P ribozymes (Pan and Sosnick, 1997; Woodson, 2000, 2005). In contrast, folding of the C-domain of the RNase P ribozyme (Fang et al., 2002; Sosnick and Pan, 2003) and a smaller group I ribozyme from *Azoarcus* (Pan et al., 2000; Rangan et al., 2003; Perez-Salas et al., 2004) is devoid of kinetic traps and occurs on a much faster time scale.

One of the most puzzling features of D135 ribozyme folding is the combination of the slow rate (1 min<sup>-1</sup>) (Swisher et al., 2002) and the lack of experimental evidence of a kinetic trap (Swisher et al., 2001, 2002). One hypothesis is that folding is slowed by high contact order, i.e., the need to form long-range tertiary contacts between



**Figure 3** Folding pathways of ribozymes derived from the ai5 $\gamma$  group II intron under high-salt (top) and near-physiological (bottom) conditions.

regions located far apart from each other in the secondary structure (Swisher et al., 2002). Recently, a NAIM study using compaction of this RNA as a selection step was used to identify atoms and functional groups in the D135 ribozyme that are critical for compaction (Waldsich and Pyle, 2007). Regions involved in intra-domain  $\alpha$ - $\alpha'$  and  $\beta$ - $\beta'$  tertiary contacts and stem-loop a, as well as the three-way junction connecting the c, c1 and c2 stems of D1, also contain functional groups important for ribozyme compaction (Waldsich and Pyle, 2007), although some of these regions have previously been shown to be dispensable for catalysis. Surprisingly, the largest cluster of functionalities critical for D135 compaction has been located not in the areas forming long-range tertiary contacts, but in a small region of D1 that contains docking sites for D5 and D6 (the folding control element; Waldsich and Pyle, 2007). This result indicates that proper folding and/or positioning of this element likely triggers compaction and folding of the entire intron.

These NAIM data suggest that high contact order is an unlikely reason for the slow folding of the group II intron. As the folding rate is not affected by denaturant and the compact D1 intermediates and native structures require unusually high magnesium concentrations (Swisher et al., 2002; Su et al., 2005), folding could be governed by consolidation around magnesium ion-binding sites, as previously proposed for some large RNAs unburdened by kinetic traps (Sosnick and Pan, 2003). However, additional studies are required to determine the nature of the rate-limiting step in D135 ribozyme folding.

### Folding under near-physiological conditions

It has been shown for other RNAs, especially the *Tetrahymena* ribozyme, that changing ionic conditions and/or temperature can affect the folding pathway (Woodson, 2005). Therefore, we cannot assume that under near-physiological conditions (30°C and low  $Mg^{2+}$  concentration) the ai5 $\gamma$ -derived ribozyme will follow the same pathway as under high-salt and temperature conditions. In addition, it has been proposed that under near-physio-

logical conditions the ai5 $\gamma$  folding pathway is rugged (Mohr et al., 2006). Interestingly, instead of reaching rapid equilibrium between unfolded and compact states, at 30°C the ribozyme reaches complete compaction, even at low magnesium concentrations (3 mM), which has not been observed at 42°C (Fedorova et al., 2007). However, compaction under these conditions is two orders of magnitude slower than under optimal conditions. Under all conditions tested, addition of subdenaturing urea did not increase the rate of compaction, suggesting that the folding pathway may be free of kinetic traps (Fedorova et al., 2007). Furthermore, it has been shown that the compact state formed at low  $Mg^{2+}$  concentrations is an on-pathway folding intermediate (the near-native state, Figure 3), which can be easily chased to the native state by increasing the magnesium concentration. The near-native state contains most of the D1 structural elements (except for the  $\epsilon$ - $\epsilon'$  and  $\theta$ - $\theta'$  interactions), but D5 and D3 are undocked (Fedorova et al., 2007). This is in good agreement with data previously obtained at 42°C and suggests that the near-native state at 30°C is structurally similar to the transient intermediate formed at 42°C.

Urea titration experiments demonstrated that the compact intermediate is not very stable at low  $Mg^{2+}$  concentrations. However, after the native state is formed at high magnesium concentrations, dilution of magnesium to much lower concentrations results in the formation of only the compact intermediate, but not an unfolded state. In addition, as mentioned above, compaction is extremely slow at low magnesium concentrations. The combination of these results suggests that the compact intermediate has kinetic stability due to a large energy barrier between unfolded and compact states, that is, compaction occurs under kinetic control (Fedorova et al., 2007).

In conclusion, under near-physiological ionic conditions the ribozyme has two major folding problems: first, formation of the near-native intermediate is very slow, and second, the native state is not stable and requires additional stabilization by either high salt concentration or a protein co-factor.

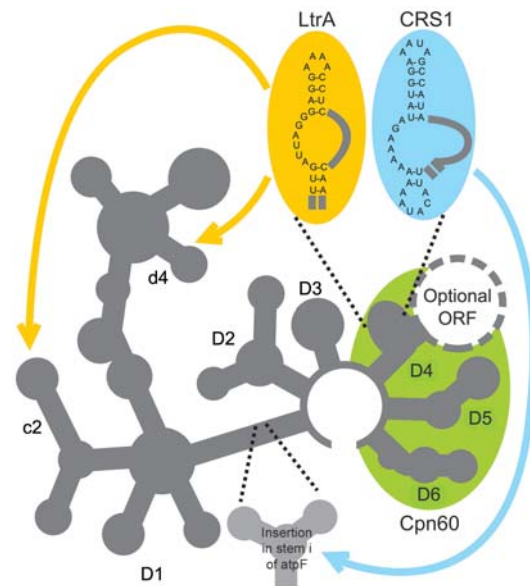
## Protein-assisted splicing

As mentioned above, most *in vitro* experiments on group II intron mechanism and folding suffer from the fact that they have to be executed under non-physiological conditions to observe reactivity. *In vivo*, various proteins have been found to assist group II intron splicing. Some of these act indirectly and unspecifically, e.g., the mitochondrial membrane proteins Mrs2p and Lpe10p that are assumed to assist group II intron splicing in yeast mitochondria by raising the intraorganellar  $Mg^{2+}$  concentrations (Gregan et al., 2001; Weghuber et al., 2006). However, even with this activity, mitochondrial magnesium concentrations are still one to two orders of magnitude lower than the amounts typically used for *in vitro* experiments. Clearly, other factors are required to promote splicing *in vivo*, e.g., other ions, small molecules or proteins. A wide variety of proteins are known to interact directly (and often specifically) with group II intron RNA and promote splicing by stabilizing the catalytically active RNA structure or by enabling the RNA to fold properly (chaperone function) at low magnesium concentrations.

## Maturases

Maturases are proteins that have been acquired by introns and are usually encoded within D4 of the intron sequence. These multifunctional proteins specifically bind their own mRNA, promote the formation of critical tertiary RNA contacts, and thus enable splicing (Matsura et al., 2001). In addition, maturases also play a crucial role in retrohoming and retrotransposition through target-primed reverse transcription, the replication mechanism of group II introns: they usually contain an endonuclease (EN) domain that recognizes and cleaves the antisense strand of the cognate DNA target. Moreover, after reverse splicing of the intron RNA into the sense strand, the maturase-encoded reverse transcriptase (RT) domain uses the free 3'-end of the cleaved target as a primer to synthesize the intron cDNA (Lambowitz and Zimmerly, 2004).

While EN and RT are easily identifiable domains with a phylogenetically conserved sequence, the regions responsible for maturase function are more diffuse. Mutational studies and threading (Cui et al., 2004) have shown that the RT X domain, in conjunction with other RNA binding regions of the RT, seems to be responsible for maturase function (Mohr et al., 1993). The N-terminal domain of the LtrB maturase specifically binds group II intron RNA in a region of D4 that overlaps with the 5'-end of its coding sequence (Cui et al., 2004; Blocker et al., 2005). Through additional contacts, it then stabilizes the correctly folded intron structure (Figure 4) (Matsuura et al., 2001). Careful biochemical analysis of the intron/maturase interaction revealed that the active ribonucleoprotein particle (RNP) contains a maturase dimer and is formed by mutual stabilization of the RNA and the protein component (mutually induced fit model; Rambo and Doudna, 2004). A recent study on the Rmlnt1 intron (Molina-Sanchez et al., 2006) suggests that the maturase may also promote formation of the correct IBS/EBS inter-



**Figure 4** Protein binding sites in group II introns.

A generic group II intron structure is schematically shown in gray. Colored ovals represent the intron-associated proteins LtrA (yellow), CRS1 (blue) and Cpn60 (green). Arrows indicate interactions with secondary binding sites.

action, and even implicates it as a regulator controlling the balance between lariat and circle formation.

## Host-encoded splicing factors

Many group II introns lack a functional open reading frame and thus do not encode their own maturase. Frequently, these introns do not use IEPs from other introns, but recruit host-encoded proteins in order to splice efficiently. A number of proteins have been implicated in group II intron splicing. Some are degenerate maturase proteins with mutated and/or lost RT and EN domains that presumably have retained their maturase function (e.g., matK, nMat; Vogel et al., 1999; Mohr and Lambowitz, 2003). However, proteins identified as splicing factors often seem to be derived from different host-encoded proteins. Group II intron splicing in chloroplasts illustrates the wide variety of proteins that have been 'adopted' by different introns in various organisms to ensure efficient and specific splicing: For example, psaA mRNA in *Chlamydomonas reinhardtii* is known to require at least 14 nuclear genes for splicing of its two introns (Girard et al., 1980; Goldschmidt-Clermont et al., 1990). Only a subset has been characterized so far. This includes Rat1 and Rat2, proteins that show homology to poly(ADP-ribose) polymerases (Balczun et al., 2005), Raa2 (previously called Maa2), which is derived from a pseudouridine synthase (Perron et al., 1999), and Raa1, a bifunctional protein in which two distinct domains are responsible for splicing of two different introns (Merenino et al., 2006). Despite their general importance for splicing, however, it is still unclear whether any of these factors act directly or indirectly on the intron RNA.

It has been proposed that Cpn60 is a general splicing factor in *C. reinhardtii* based on its specific binding to the heterologous mitochondrial intron r11 (Bunse et al., 2001). It resembles a heat shock protein with two GroEL-like

ATPase domains and has been shown to bind both homologous and heterologous group II intron RNA in a region narrowed down to D4–D6 (Figure 4). However, its mechanism of action is not yet clear (Balczun et al., 2006). Similarly, it has been shown that cNAPL, a protein resembling nuclear nucleosome assembly proteins, interacts with group II intron RNA (Glanz et al., 2006), but neither its binding sites nor its function have been conclusively examined to date.

CRS2 is a maize protein derived from tRNA peptidyl hydrolases (Jenkins and Barkan, 2001) and co-sediments with group II intron RNA. However, it does not exhibit specific intron binding activity, but instead interacts with two proteins, CAF1 and CAF2, which convey selectivity for individual introns (Ostheimer et al., 2003, 2006). The latter proteins belong to a family defined by the RNA binding module CRM (chloroplast RNA splicing and ribosome maturation) (Barkan et al., 2007). Another important member of the CRM family is CRS1, a splicing factor that specifically promotes splicing of the intron *atpF* by binding and compacting the RNA. CRS1 interacts with the region in D4 that is homologous to the maturase binding site in LtrB, and makes additional specific contacts to an insertion in D1 (Figure 4). Interestingly, *atpF* RNA does not appear to fold properly *in vitro*, even under high-salt conditions, indicating that in this case, the host protein not only lowers the magnesium requirements, but is absolutely required for proper folding (Ostersetzer et al., 2005).

None of the proteins mentioned above appears to be able to stimulate splicing *in vitro*, indicating that they are not sufficient to promote splicing and require additional co-factors. Indeed, many of these proteins were identified as components of RNP complexes as large as 1700 kDa (Jenkins and Barkan, 2001; Rivier et al., 2001; Till et al., 2001; Perron et al., 2004). Recently, however, it was shown that two closely related DEAD-box proteins (Cyt19 from *Neurospora crassa* and Mss116 from *Saccharomyces cerevisiae*) alone are sufficient to stimulate group II intron splicing *in vitro* under near-physiological conditions (Mohr et al., 2006; Solem et al., 2006; Halls et al., 2007). It has long been known that Mss116 is essential for splicing *in vivo* (Tzagoloff et al., 1975; Faye and Simon, 1983; Huang et al., 2005). The phenotype of an Mss116 knockout mutant of *S. cerevisiae* could be rescued by overexpression of Cyt19 (Huang et al., 2005), indicating that both proteins use a similar mechanism to promote splicing. A new *in vitro* assay allowed examination of this mechanism. Cyt19- and Mss116-promoted splicing is dependent on ATP hydrolysis (Mohr et al., 2006; Solem et al., 2006; Halls et al., 2007) and thus differs substantially from the maturase mechanism that acts through ATP-independent RNA binding and stabilization. Since both proteins belong to the family of DEAD-box helicases, and Mss116 has demonstrated helicase activity *in vitro* (Solem et al., 2006; Halls et al., 2007), it was first thought that these proteins promote splicing by unwinding kinetic traps (Seraphin and Rosbash, 1989; Mohr et al., 2006; Halls et al., 2007). Surprisingly, however, subsequent experiments with Mss116 mutants that are unable to unwind, but still retain their ATPase function, indicate that unwinding activity is not required to

promote splicing (Solem et al., 2006). Still, the mechanism by which these proteins act on group II introns is far from being elucidated. Unexpectedly, the *S. cerevisiae* DEAD-box protein Ded1, a translation initiation factor that is not implicated in intron self-splicing, was also shown to stimulate splicing *in vitro* (Solem et al., 2006; Halls et al., 2007). Mss116, Cyt19 and Ded1 rather un-specifically facilitate splicing of a number of non-cognate introns (both group I and group II), but yield variable results under specific assay settings (Halls et al., 2007). Together, these findings indicate that, in spite of their complicated secondary and tertiary structure, group II introns do not necessarily require a specific protein for activity. Further characterization of these protein-RNA interactions will soon allow new insights into group II intron folding and reactivity.

## Acknowledgments

We thank Amanda Solem, Michael Roitzsch and Jennifer Li-Pook-Than for critically reading the manuscript. O.F. thanks the Howard Hughes Medical Institute for funding.

## References

- Abramovitz, D.L., Friedman, R.A., and Pyle, A.M. (1996). Catalytic role of 2'-hydroxyl groups within a group II intron active site. *Science* 271, 1410–1413.
- Adamidi, C., Fedorova, O., and Pyle, A.M. (2003). A group II intron inserted into a bacterial heat-shock operon shows autocatalytic activity and unusual thermostability. *Biochemistry* 42, 3409–3418.
- Aizawa, Y., Xiang, Q., Lambowitz, A.M., and Pyle, A.M. (2003). The pathway for DNA recognition and RNA integration by a group II intron retrotransposon. *Mol. Cell* 11, 795–805.
- Augustin, S., Müller, M.W., and Schweyen, R.J. (1990). Reverse self-splicing of Group II intron RNAs *in vitro*. *Nature* 343, 383–386.
- Balczun, C., Bunse, A., Hahn, D., Bennoun, P., Nickelsen, J., and Kück, U. (2005). Two adjacent nuclear genes are required for functional complementation of a chloroplast *trans*-splicing mutant from *Chlamydomonas reinhardtii*. *Plant J.* 43, 636–648.
- Balczun, C., Bunse, A., Schwarz, C., Piotrowski, M., and Kück, U. (2006). Chloroplast heat shock protein Cpn60 from *Chlamydomonas reinhardtii* exhibits a novel function as a group II intron-specific RNA-binding protein. *FEBS Lett.* 580, 4527–4532.
- Barkan, A., Klipcan, L., Ostersetzer, O., Kawamura, T., Asakura, Y., and Watkins, K.P. (2007). The CRM domain: an RNA binding module derived from an ancient ribosome-associated protein. *RNA* 13, 55–64.
- Blocker, F.J., Mohr, G., Conlan, L.H., Qi, L., Belfort, M., and Lambowitz, A.M. (2005). Domain structure and three-dimensional model of a group II intron-encoded reverse transcriptase. *RNA* 11, 14–28.
- Boudvillain, M. and Pyle, A.M. (1998). Defining functional groups, core structural features and inter-domain tertiary contacts essential for group II intron self-splicing: a NAIM analysis. *EMBO J.* 17, 7091–7104.
- Boudvillain, M., de Lencastre, A., and Pyle, A.M. (2000). A new RNA tertiary interaction that links active-site domains of a group II intron and anchors them at the site of catalysis. *Nature* 406, 315–318.

- Boulanger, S.C., Belcher, S.M., Schmidt, U., Dib-Hajj, S.D., Schmidt, T., and Perlman, P.S. (1995). Studies of point mutants define three essential paired nucleotides in the domain 5 substructure of a group II intron. *Mol. Cell. Biol.* **15**, 4479–4488.
- Bunse, A.A., Nickelsen, J., and Kuck, U. (2001). Intron-specific RNA binding proteins in the chloroplast of the green alga *Chlamydomonas reinhardtii*. *Biochim. Biophys. Acta* **1519**, 46–54.
- Cech, T.R. (1990). Self-splicing of group I introns. *Annu. Rev. Biochem.* **59**, 543–568.
- Chanfreau, G. and Jacquier, A. (1994). Catalytic site components common to both splicing steps of a group II intron. *Science* **266**, 1383–1387.
- Chanfreau, G. and Jacquier, A. (1996). An RNA conformational change between the two chemical steps of group II self-splicing. *EMBO J.* **15**, 3466–3476.
- Chin, K. and Pyle, A.M. (1995). Branch-point attack in group II introns is a highly reversible transesterification, providing a possible proof-reading mechanism for 5'-splice site selection. *RNA* **1**, 391–406.
- Chu, V.-T. (2000) Mechanism of branch-point selection in a catalytic group II intron (New York, USA: Columbia University).
- Chu, V.-T., Liu, Q., Podar, M., Perlman, P.S., and Pyle, A.M. (1998). More than one way to splice an RNA: branching without a bulge and splicing without branching in group II introns. *RNA* **4**, 1186–1202.
- Chu, V.T., Adamidi, C., Liu, Q., Perlman, P.S., and Pyle, A. (2001). Control of branch-site choice by a group II intron. *EMBO J.* **20**, 6866–6876.
- Costa, M. and Michel, F. (1995). Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J.* **14**, 1276–1285.
- Costa, M., Deme, E., Jacquier, A., and Michel, F. (1997). Multiple tertiary interactions involving domain II of group II self-splicing introns. *J. Mol. Biol.* **267**, 520–536.
- Costa, M., Christian, E.L., and Michel, F. (1998). Differential chemical probing of a group II self-splicing intron identifies bases involved in tertiary interactions and supports an alternative secondary structure model of domain V. *RNA* **4**, 1055–1068.
- Costa, M., Michel, F., and Westhof, E. (2000). A three-dimensional perspective on exon binding by a group II self-splicing intron. *EMBO J.* **19**, 5007–5018.
- Costa, M., Michel, F., Molina-Sanchez, M.D., Martinez-Abarca, F., and Toro, N. (2006a). An alternative intron-exon pairing scheme implied by unexpected *in vitro* activities of group II intron Rmlnt1 from *Sinorhizobium meliloti*. *Biochimie* **88**, 711–717.
- Costa, M., Michel, F., and Toro, N. (2006b). Potential for alternative intron-exon pairings in group II intron Rmlnt1 from *Sinorhizobium meliloti* and its relatives. *RNA* **12**, 338–341.
- Cui, X., Matsuura, M., Wang, Q., Ma, H., and Lambowitz, A.M. (2004). A group II intron-encoded maturase functions preferentially in cis and requires both the reverse transcriptase and X domains to promote RNA splicing. *J. Mol. Biol.* **340**, 211–231.
- Daniels, D., Michels, W.J., and Pyle, A.M. (1996). Two competing pathways for self-splicing by group II introns; a quantitative analysis of *in-vitro* reaction rates and products. *J. Mol. Biol.* **256**, 31–49.
- de Lencastre, A., Hamill, S., and Pyle, A.M. (2005). A single active-site region for a group II intron. *Nat. Struct. Mol. Biol.* **12**, 626–627.
- Erat, M.C., Zerbe, O., Fox, T., and Sigel, R.K. (2007). Solution structure of domain 6 from a self-splicing group II intron ribozyme: a Mg<sup>2+</sup> binding site is located close to the stacked branch adenosine. *ChemBioChem* **8**, 306–314.
- Fang, X., Thiyagarajan, P., Sosnick, T., and Pan, T. (2002). The rate-limiting step in the folding of a large ribozyme without kinetic traps. *Proc. Natl. Acad. Sci. USA* **99**, 8518–8523.
- Faye, G. and Simon, M. (1983). Processing of the Oxi-3 pre-messenger RNA in yeast. In: *Mitochondria 1983: Nucleo-Mitochondrial Interactions*, R.J. Schweyen, K. Wolf, and F. Kaudewitz, eds. (Berlin, Germany: de Gruyter), pp. 433–439.
- Fedorova, O. and Pyle, A.M. (2005). Linking the group II intron catalytic domains: tertiary contacts and structural features of domain 3. *EMBO J.* **24**, 3906–3916.
- Fedorova, O., Mitros, T., and Pyle, A.M. (2003). Domains 2 and 3 interact to form critical elements of the group II intron active site. *J. Mol. Biol.* **330**, 197–209.
- Fedorova, O., Waldsich, C., and Pyle, A.M. (2007). Group II intron folding under near-physiological conditions: collapsing to the near-native State. *J. Mol. Biol.* **366**, 1099–1114.
- Gaur, R.K., McLaughlin, L.W., and Green, M.R. (1997). Functional group substitutions of the branchpoint adenosine in a nuclear pre-mRNA and a group II intron. *RNA* **3**, 861–869.
- Girard, J., Chua, N.H., Bennoun, P., Schmidt, G., and Delosme, M. (1980). Studies on mutants deficient in the photosystem I reaction centers in *Chlamydomonas reinhardtii*. *Curr. Genet.* **2**, 215–221.
- Glanz, S., Bunse, A., Wimbert, A., Balczun, C., and Kück, U. (2006). A nucleosome assembly protein-like polypeptide binds to chloroplast group II intron RNA in *Chlamydomonas reinhardtii*. *Nucleic Acids Res.* **34**, 5337–5351.
- Goldschmidt-Clermont, M., Girard-Bascou, J., Choquet, Y., and Rochaix, J.D. (1990). *Trans*-splicing mutants of *Chlamydomonas reinhardtii*. *Mol. Gen. Genet.* **223**, 417–425.
- Gordon, P.M. and Piccirilli, J.A. (2001). Metal ion coordination by the AGC triad in domain 5 contributes to group II intron catalysis. *Nat. Struct. Biol.* **8**, 893–898.
- Gregan, J., Kolisek, M., and Schweyen, R.J. (2001). Mitochondrial Mg<sup>2+</sup> homeostasis is critical for group II intron splicing *in vivo*. *Genes Dev.* **15**, 2229–2237.
- Griffin, E.A., Qin, Z.-F., Michels, W.A., and Pyle, A.M. (1995). Group II intron ribozymes that cleave DNA and RNA linkages with similar efficiency, and lack contacts with substrate 2'-hydroxyl groups. *Chem. Biol.* **2**, 761–770.
- Halls, C., Mohr, S., Del Campo, M., Yang, Q., Jankowsky, E., and Lambowitz, A.M. (2007). Involvement of DEAD-box proteins in group I and group II intron splicing. Biochemical characterization of Mss116p, ATP hydrolysis-dependent and -independent mechanisms, and general RNA chaperone activity. *J. Mol. Biol.* **365**, 835–855.
- Hamill, S. and Pyle, A.M. (2006). The receptor for branch-site docking within a group II intron active site. *Mol. Cell* **23**, 831–840.
- Hebbar, S.K., Belcher, S.M., and Perlman, P.S. (1992). A maturase-encoding Group IIA intron of yeast mitochondria self-splices *in vitro*. *Nucleic Acids Res.* **20**, 1747–1754.
- Huang, H.R., Rowe, C.E., Mohr, S., Jiang, Y., Lambowitz, A.M., and Perlman, P.S. (2005). The splicing of yeast mitochondrial group I and group II introns requires a DEAD-box protein with RNA chaperone function. *Proc. Natl. Acad. Sci. USA* **102**, 163–168.
- Jacquier, A. (1990). Self-splicing group II and nuclear pre-mRNA introns: how similar are they? *Trends Biochem. Sci.* **15**, 351–354.
- Jacquier, A. and Jacquesson-Breuleux, N. (1991). Splice site selection and role of the lariat in a Group II intron. *J. Mol. Biol.* **219**, 415–428.
- Jacquier, A. and Michel, F. (1987). Multiple exon-binding sites in class II self-splicing introns. *Cell* **50**, 17–29.
- Jacquier, A. and Michel, F. (1990). Base-pairing interactions involving the 5'- and 3'-terminal nucleotides of group II self-splicing introns. *J. Mol. Biol.* **213**, 437–447.
- Jarrell, K.A., Peebles, C.L., Dietrich, R.C., Romiti, S.L., and Perlman, P.S. (1988). Group II intron self-splicing: alternative reaction conditions yield novel products. *J. Biol. Chem.* **263**, 3432–3439.
- Jenkins, B.D. and Barkan, A. (2001). Recruitment of a peptidyl-tRNA hydrolase as a facilitator of group II intron splicing in chloroplasts. *EMBO J.* **20**, 872–879.

- Jestin, J.-L., Deme, E., and Jacquier, A. (1997). Identification of structural elements critical for inter-domain interactions in a group II self-splicing intron. *EMBO J.* **16**, 2945–2954.
- Koch, J.L., Boulanger, S.C., Dib-Hajj, S.D., Hebbar, S.K., and Perlman, P.S. (1992). Group II introns deleted for multiple substructures retain self-splicing activity. *Mol. Cell. Biol.* **12**, 1950–1958.
- Konforti, B.B., Abramovitz, D.L., Duarte, C.M., Karpeisky, A., Beigelman, L., and Pyle, A.M. (1998). Ribozyme catalysis from the major groove of group II intron domain 5. *Mol. Cell* **1**, 433–441.
- Lambowitz, A.M. and Zimmerly, S. (2004). Mobile group II introns. *Annu. Rev. Genet.* **38**, 1–35.
- Lang, B.F., Laforest, M.J., and Burger, G. (2007). Mitochondrial introns: a critical view. *Trends Genet.* **23**, 119–125.
- Lehmann, K. and Schmidt, U. (2003). Group II introns: structure and catalytic versatility of large natural ribozymes. *Crit. Rev. Biochem. Mol. Biol.* **38**, 249–303.
- Li-Pook-Than, J. and Bonen, L. (2006). Multiple physical forms of excised group II intron RNAs in wheat mitochondria. *Nucleic Acids Res.* **34**, 2782–2790.
- Liu, Q., Green, J.B., Khodadadi, A., Haeberli, P., Beigelman, L., and Pyle, A.M. (1997). Branch-site selection in a group II intron mediated by active recognition of the adenine amino group and steric exclusion of non-adenine functionalities. *J. Mol. Biol.* **267**, 163–171.
- Matsuura, M., Noah, J.W., and Lambowitz, A.M. (2001). Mechanism of maturase-promoted group II intron splicing. *EMBO J.* **20**, 7259–7270.
- Merendino, L., Perron, K., Rahire, M., Howald, I., Rochaix, J.D., and Goldschmidt-Clermont, M. (2006). A novel multifunctional factor involved in *trans*-splicing of chloroplast introns in *Chlamydomonas*. *Nucleic Acids Res.* **34**, 262–274.
- Michel, F. and Ferat, J.-L. (1995). Structure and activities of group II introns. *Annu. Rev. Biochem.* **64**, 435–461.
- Michel, F., Umesono, K., and Ozeki, H. (1989). Comparative and functional anatomy of group II catalytic introns – a review. *Gene* **82**, 5–30.
- Michels, W.J. and Pyle, A.M. (1995). Conversion of a group II intron into a new multiple-turnover ribozyme that selectively cleaves oligonucleotides: elucidation of reaction mechanism and structure/function relationships. *Biochemistry* **34**, 2965–2977.
- Mikheeva, S., Murray, H.L., Zhou, H., Turczyk, B.M., and Jarrell, K.A. (2000). Deletion of a conserved dinucleotide inhibits the second step of group II intron splicing. *RNA* **6**, 1509–1515.
- Mohr, G. and Lambowitz, A.M. (2003). Putative proteins related to group II intron reverse transcriptase/maturases are encoded by nuclear genes in higher plants. *Nucleic Acids Res.* **31**, 647–652.
- Mohr, G., Perlman, P.S., and Lambowitz, A.M. (1993). Evolutionary relationships among group II intron-encoded proteins and identification of a conserved domain that may be related to maturase function. *Nucleic Acids Res.* **21**, 4991–4997.
- Mohr, S., Matsuura, M., Perlman, P.S., and Lambowitz, A.M. (2006). A DEAD-box protein alone promotes group II intron splicing and reverse splicing by acting as an RNA chaperone. *Proc. Natl. Acad. Sci. USA* **103**, 3569–3574.
- Molina-Sanchez, M.D., Martinez-Abarca, F., and Toro, N. (2006). Excision of the *Sinorhizobium meliloti* group II intron Rmlnt1 as circles *in vivo*. *J. Biol. Chem.* **281**, 28737–28744.
- Mörl, M. and Schmelzer, C. (1990). Integration of group II intron bl1 into a foreign RNA by reversal of the self-splicing reaction *in vitro*. *Cell* **60**, 629–636.
- Mörl, M., Niemer, I., and Schmelzer, C. (1992). New reactions catalyzed by a Group II intron ribozyme with RNA and DNA substrates. *Cell* **70**, 803–810.
- Müller, M.W., Stocker, P., Hetzer, M., and Schweyen, R.J. (1991). Fate of the junction phosphate in alternating forward and reverse self-splicing reaction of group II intron RNA. *J. Mol. Biol.* **222**, 145–150.
- Murray, H.L., Mikheeva, S., Coljee, V.W., Turczyk, B.M., Donahue, W.F., Bar-Shalom, A., and Jarrell, K.A. (2001). Excision of group II introns as circles. *Mol. Cell* **8**, 201–211.
- Nielsen, H., Westhof, E., and Johansen, S. (2005). An mRNA is capped by a 2',5' lariat catalyzed by a group I-like ribozyme. *Science* **309**, 1584–1587.
- Osiewacz, H.D. and Esser, K. (1984). The mitochondrial plasmid of *Podospora anserina*: a mobile intron of a mitochondrial gene. *Curr. Genet.* **8**, 299–305.
- Ostersetzer, O., Cooke, A.M., Watkins, K.P., and Barkan, A. (2005). CRS1, a chloroplast group II intron splicing factor, promotes intron folding through specific interactions with two intron domains. *Plant Cell* **17**, 241–255.
- Ostheimer, G.J., Williams-Carrier, R., Belcher, S., Osborne, E., Gierke, J., and Barkan, A. (2003). Group II intron splicing factors derived by diversification of an ancient RNA-binding domain. *EMBO J.* **22**, 3919–3929.
- Ostheimer, G.J., Rojas, M., Hadjivassiliou, H., and Barkan, A. (2006). Formation of the CRS2-CAF2 group II intron splicing complex is mediated by a 22-amino acid motif in the COOH-terminal region of CAF2. *J. Biol. Chem.* **281**, 4732–4738.
- Padgett, R.A., Podar, M., Boulanger, S.C., and Perlman, P.S. (1994). The stereochemical course of group II intron self-splicing. *Science* **266**, 1685–1688.
- Pan, J., Deras, M., and Woodson, S. (2000). Fast folding of a ribozyme by stabilizing core interactions: evidence for multiple folding pathways in RNA. *J. Mol. Biol.* **296**, 133–144.
- Pan, T. and Sosnick, T.R. (1997). Intermediates and kinetic traps in the folding of a large ribozyme revealed by circular dichroism and UV absorbance spectroscopies and catalytic activity. *Nat. Struct. Biol.* **4**, 931–938.
- Peebles, C.L., Benatan, E.J., Jarrell, K.A., and Perlman, P.S. (1987). Group II intron self-splicing: development of alternative reaction conditions and identification of a predicted intermediate. *Cold Spring Harbor Symp. Quant. Biol.* **52**, 223–232.
- Peebles, C.L., Zhang, M., Perlman, P.S., and Franzen, J.F. (1995). Identification of a catalytically critical trinucleotide in Domain 5 of a group II intron. *Proc. Natl. Acad. Sci. USA* **92**, 4422–4426.
- Perez-Salas, U.A., Rangan, P., Krueger, S., Briber, R.M., Thirumalai, D., and Woodson, S.A. (2004). Compaction of a bacterial group I ribozyme coincides with the assembly of core helices. *Biochemistry* **43**, 1746–1753.
- Perron, K., Goldschmidt-Clermont, M., and Rochaix, J.D. (1999). A factor related to pseudouridine synthases is required for chloroplast group II intron *trans*-splicing in *Chlamydomonas reinhardtii*. *EMBO J.* **18**, 6481–6490.
- Perron, K., Goldschmidt-Clermont, M., and Rochaix, J.D. (2004). A multiprotein complex involved in chloroplast group II intron splicing. *RNA* **10**, 704–711.
- Podar, M., Dib-Hajj, S., and Perlman, P.S. (1995a). A UV-induced Mg<sup>2+</sup>-dependent cross-link traps an active form of domain 3 of a self-splicing group II intron. *RNA* **1**, 828–840.
- Podar, M., Perlman, P.S., and Padgett, R.A. (1995b). Stereochemical selectivity of group II intron splicing, reverse-splicing and hydrolysis reactions. *Mol. Cell. Biol.* **15**, 4466–4478.
- Podar, M., Chu, V.T., Pyle, A.M., and Perlman, P.S. (1998a). Group II intron splicing *in vivo* by first step hydrolysis. *Nature* **391**, 915–918.
- Podar, M., Zhou, J., Zhang, M., Franzen, J.S., Perlman, P.S., and Peebles, C.L. (1998b). Domain 5 binds near a highly-conserved dinucleotide in the joiner linking domains 2 and 3 of a group II intron. *RNA* **4**, 151–166.
- Pyle, A.M. and Lambowitz, A.M. (2006). Group II introns: ribozymes that splice RNA and invade DNA. In: *The RNA World*, R.F. Gesteland, C.T.R., and A.J.F., eds. (Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press), pp. 469–506.

- Pyle, A.M., Fedorova, O., and Waldsich, C. (2007). Folding of group II introns: a model system for large, multidomain RNAs? *Trends Biochem. Sci.* **32**, 138–145.
- Qin, P.Z.F. and Pyle, A.M. (1997). Stopped-flow fluorescence spectroscopy reveals that domain 1 of a group II intron is an independent folding unit. *Biochemistry* **36**, 4718–4730.
- Qin, P.Z. and Pyle, A.M. (1998). The architectural organization and mechanistic function of group II intron structural elements. *Curr. Opin. Struct. Biol.* **8**, 301–308.
- Qin, P.Z. and Pyle, A.M. (1999). Antagonistic substrate binding by a group II intron ribozyme. *J. Mol. Biol.* **297**, 15–27.
- Query, C.C., Moore, M.M., and Sharp, P.A. (1994). Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model. *Genes Dev.* **8**, 587–597.
- Rambo, R.P. and Doudna, J.A. (2004). Assembly of an active group II intron-maturase complex by protein dimerization. *Biochemistry* **43**, 6486–6497.
- Rangan, P., Masquida, B., Westhof, E., and Woodson, S.A. (2003). Assembly of core helices and rapid tertiary folding of a small bacterial group I ribozyme. *Proc. Natl. Acad. Sci. USA* **100**, 1574–1579.
- Rivier, C., Goldschmidt-Clermont, M., and Rochaix, J.D. (2001). Identification of an RNA-protein complex involved in chloroplast group II intron *trans*-splicing in *Chlamydomonas reinhardtii*. *EMBO J.* **20**, 1765–1773.
- Robart, A.R. and Zimmerly, S. (2005). Group II intron retroelements: function and diversity. *Cytogenet. Genome Res.* **110**, 589–597.
- Ruskin, B., Greene, J.M., and Green, M.R. (1985). Cryptic branch point activation allows accurate *in vitro* splicing of human  $\beta$ -globin intron mutants. *Cell* **41**, 833–844.
- Schellenberg, M.J., Edwards, R.A., Ritchie, D.B., Kent, O.A., Golas, M.M., Stark, H., Luhrmann, R., Glover, J.N., and MacMillan, A.M. (2006). Crystal structure of a core spliceosomal protein interface. *Proc. Natl. Acad. Sci. USA* **103**, 1266–1271.
- Schlatterer, J.C., Crayton, S.H., and Greenbaum, N.L. (2006). Conformation of the Group II intron branch site in solution. *J. Am. Chem. Soc.* **128**, 3866–3867.
- Schmidt, U., Riederer, B., Morl, M., Schmelzer, C., and Stahl, U. (1990). Self-splicing of the mobile group II intron of the filamentous fungus *Podospira anserina* (COI 1) *in vitro*. *EMBO J.* **9**, 2289–2298.
- Schmidt, U., Podar, M., Stahl, U., and Perlman, P.S. (1996). Mutations of the two-nucleotide bulge of D5 of a group II intron block splicing *in vitro* and *in vivo*: phenotypes and suppressor mutations. *RNA* **2**, 1161–1172.
- Seraphin, B. and Rosbash, M. (1989). Identification of functional U1 snRNP-premRNA complexes committed to spliceosome assembly and splicing. *Cell* **59**, 349–358.
- Sigel, R., Vaidya, A., and Pyle, A. (2000). Metal ion binding sites in a group II intron core. *Nat. Struct. Biol.* **7**, 1111–1116.
- Sigel, R.K., Sashital, D.G., Abramovitz, D.L., Palmer, A.G., Butcher, S.E., and Pyle, A.M. (2004). Solution structure of domain 5 of a group II intron ribozyme reveals a new RNA motif. *Nat. Struct. Mol. Biol.* **11**, 187–192.
- Singh, R.N., Saldanha, R.J., D'Souza, L.M., and Lambowitz, A.M. (2002). Binding of a group II intron-encoded reverse transcriptase/maturase to its high affinity intron RNA binding site involves sequence-specific recognition and autoregulates translation. *J. Mol. Biol.* **318**, 287–303.
- Solem, A., Zingler, N., and Pyle, A.M. (2006). A DEAD protein that activates intron self-splicing without unwinding RNA. *Mol. Cell* **24**, 611–617.
- Sosnick, T.R. and Pan, T. (2003). RNA folding: models and perspectives. *Curr. Opin. Struct. Biol.* **13**, 309–316.
- Spadaccini, R., Reidt, U., Dybkov, O., Will, C., Frank, R., Stier, G., Corsini, L., Wahl, M.C., Luhrmann, R., and Sattler, M. (2006). Biochemical and NMR analyses of an SF3b155-p14-U2AF-RNA interaction network involved in branch point definition during pre-mRNA splicing. *RNA* **12**, 410–425.
- Stabell, F.B., Tourasse, N.J., Ravnun, S., and Kolstø, A.B. (2007). Group II intron in *Bacillus cereus* has an unusual 3' extension and splices 56 nucleotides downstream of the predicted site. *Nucleic Acids Res.*, in press; doi:10.1093/nar/gkm1031.
- Su, L., Qin, P., Michels, W., and Pyle, A. (2001). Guiding ribozyme cleavage through motif recognition: the mechanism of cleavage site selection by a group II intron ribozyme. *J. Mol. Biol.* **306**, 665–668.
- Su, L.J., Brenowitz, M., and Pyle, A.M. (2003). An alternative route for the folding of large RNAs: apparent two-state folding by a group II intron ribozyme. *J. Mol. Biol.* **334**, 639–652.
- Su, L.J., Waldsich, C., and Pyle, A.M. (2005). An obligate intermediate along the slow folding pathway of a group II intron ribozyme. *Nucleic Acids Res.* **33**, 6674–6687.
- Swisher, J., Duarte, C., Su, L., and Pyle, A. (2001). Visualizing the solvent-inaccessible core of a group II intron ribozyme. *EMBO J.* **20**, 2051–2061.
- Swisher, J., Su, L., Brenowitz, M., Anderson, V., and Pyle, A. (2002). Productive folding to the native state by a group II intron ribozyme. *J. Mol. Biol.* **315**, 297–310.
- Till, B., Schmitz-Linneweber, C., Williams-Carrier, R., and Barkan, A. (2001). CRS1 is a novel group II intron splicing factor that was derived from a domain of ancient origin. *RNA* **7**, 1227–1238.
- Toor, N., Hausner, G., and Zimmerly, S. (2001). Coevolution of group II intron RNA structures with their intron-encoded reverse transcriptases. *RNA* **7**, 1142–1152.
- Toor, N., Robart, A.R., Christianson, J., and Zimmerly, S. (2006). Self-splicing of a group IIC intron: 5' exon recognition and alternative 5' splicing events implicate the stem-loop motif of a transcriptional terminator. *Nucleic Acids Res.* **34**, 6461–6471.
- Tourasse, N.J., Stabell, F.B., Reiter, L., and Kolstø, A.B. (2005). Unusual group II introns in bacteria of the *Bacillus cereus* group. *J. Bacteriol.* **187**, 5437–5451.
- Treiber, D.K. and Williamson, J.R. (1999). Exposing the kinetic traps in RNA folding. *Curr. Opin. Struct. Biol.* **9**, 339–345.
- Treiber, D., Rook, M., Zarrinkar, P., and Williamson, J.R. (1998). Kinetic intermediates trapped by native interactions in RNA folding. *Science* **279**, 1943–1946.
- Tzagoloff, A., Akai, A., Needleman, R.B., and Zulch, G. (1975). Assembly of the mitochondrial membrane system. Cytoplasmic mutants of *Saccharomyces cerevisiae* with lesions in enzymes of the respiratory chain and in the mitochondrial ATPase. *J. Biol. Chem.* **250**, 8236–8242.
- van der Veen, R., Arnberg, A.C., van der Horst, G., Bonen, L., Tabak, H.F., and Grivell, L.A. (1986). Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing *in vitro*. *Cell* **44**, 225–234.
- van der Veen, R., Kwakman, J.H.J.M., and Grivell, L.A. (1987). Mutations at the lariat acceptor site allow self-splicing of a group II intron without lariat formation. *EMBO J.* **6**, 3827–3831.
- Vogel, J. and Börner, T. (2002). Lariat formation and a hydrolytic pathway in plant chloroplast group II intron splicing. *EMBO J.* **21**, 3794–3803.
- Vogel, J., Börner, T., and Hess, W. (1999). Comparative analysis of splicing of the complete set of chloroplast group II introns in three higher plant mutants. *Nucleic Acids Res.* **27**, 3866–3874.
- Waldsich, C. and Pyle, A.M. (2007). A folding control element for tertiary collapse of a group II intron ribozyme. *Nat. Struct. Mol. Biol.* **14**, 37–44.
- Wang, Y. and Silverman, S.K. (2006). Experimental tests of two proofreading mechanisms for 5'-splice site selection. *ACS Chem. Biol.* **1**, 316–324.
- Watanabe, K. and Lambowitz, A.M. (2004). High-affinity binding

- site for a group II intron-encoded reverse transcriptase/maturase within a stem-loop structure in the intron RNA. *RNA* 10, 1433–1443.
- Weghuber, J., Dieterich, F., Froschauer, E.M., Svidova, S., and Schweyen, R.J. (2006). Mutational analysis of functional domains in Mrs2p, the mitochondrial Mg<sup>2+</sup> channel protein of *Saccharomyces cerevisiae*. *FEBS J.* 273, 1198–1209.
- Woodson, S.A. (2000). Recent insights on RNA folding mechanisms from catalytic RNA. *Cell. Mol. Life Sci.* 57, 796–808.
- Woodson, S.A. (2005). Metal ions and RNA folding: a highly charged topic with a dynamic future. *Curr. Opin. Chem. Biol.* 9, 104–109.
- Xiang, Q., Qin, P.Z., Michels, W.J., Freeland, K., and Pyle, A.M. (1998). The sequence-specificity of a group II intron ribozyme: multiple mechanisms for promoting unusually high discrimination against mismatched targets. *Biochemistry* 37, 3839–3849.
- Yean, S.L., Wuenschell, G., Termini, J., and Lin, R.J. (2000). Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature* 408, 881–884.
- Zaug, A.J., Grabowski, P.J., and Cech, T.R. (1983). Autocatalytic cyclization of an excised intervening sequence RNA is a cleavage-ligation reaction. *Nature* 301, 578–583.
- Zhang, L. and Doudna, J.A. (2002). Structural insights into group II intron catalysis and branch-site selection. *Science* 295, 2084–2088.