Critical Reviews in Biochemistry and Molecular Biology, 38(3):249–303 (2003)

Group II Introns: Structure and Catalytic Versatility of Large Natural Ribozymes

*Karola Lehmann and Udo Schmidt**

Department of Microbiology and Genetics, Faculty III, University of Technology Berlin, Gustav-Meyer-Allee 25, 13355 Berlin, Germany

* **Corresponding author**. e-mail: U.Schmidt@lb.tu-berlin.de, Phone: +49-30-314-72752, Fax: +49-30-314-72922

ABSTRACT: Group II introns are large, natural catalytic RNAs or ribozymes that were discovered in organelles of certain protists, fungi, algae, and plants and more recently also in prokaryotic organisms. *In vitro*, some members were found to self-splice from their pre-RNAs by two consecutive transesterification reactions joining the flanking exons and releasing the intron in a typical lariat form. Apart from self-splicing, a variety of other *in vitro* activities have been detected for group II introns demonstrating their amazing catalytic versatility. Group II introns fold into a conserved secondary structure consisting of six domains radiating from a central wheel that brings the 5' and 3' splice junction into close proximity. Domain 1 is the largest domain that is assumed to deliver the molecular scaffold assembling the intron in its active structure, while domain 5 is the phylogenetically most conserved part that represents the active site of the ribozyme. *In vivo,* the splicing reaction of many, if not all group II introns is assisted by proteins either encoded by the introns themselves (maturases), or encoded by other genes of the host organisms. The host proteins known to date have additional cellular functions and seem to have been adapted for splicing during evolution. Some of the protein-encoding group II introns were also shown to act as mobile genetic elements. They can integrate efficiently into intronless alleles of the same gene (homing) and at much lower frequencies into ectopic sites (transposition). The mobility process depends on intron encoded protein functions (endonuclease and reverse transcriptase) and on the intron RNA. This review provides a comprehensive survey of the structure/function relationships and the reaction potential of group II introns, the structurally most complicated, but also most fascinating ribozymes when looking at their catalytic repertoire *in vitro* and *in vivo*.

KEY WORDS: group II intron, ribozyme**,** self-splicing, protein-dependent splicing, maturase, homing, transposition, reverse transcriptase, endonuclease

I. INTRODUCTION

The living cell is a highly complex system, the function of which depends on innumerable chemical reactions. Most of these reactions occur spontaneously at a very slow rate and have to be accelerated to support and maintain the processes of life. It was believed for a long time that biological catalysis depended exclusively on special proteins called enzymes.

One of the most important findings in molecular biology in the last 25 years was the discovery that protein-enzymes are not the only biomolecules that are able to catalyze chemical reactions in living cells. At the beginning of the 1980s, two labs could show independently that RNA molecules can also possess catalytic activity leading to sequence-specific intra- or intermolecular cleavage of phosphodiester bonds.

During the analysis of the nuclear gene for the large ribosomal RNA of the ciliated protozoa *Tetrahymena thermophila*, Thomas R. Cech and coworkers from the University of Colorado in Boulder found that the single intron of the gene has an autocatalytic splicing activity *in vitro*. The intron can excise itself from the pre-RNA and the flanking exons are joined to the mature rRNA independent of

1040-9238/03/$.50

proteins or additional energy (Kruger *et al*., 1982). Shortly afterward, Sidney Altman and his team from Yale University showed that the RNA component of RNase P from *Escherichia coli* is able to process its substrate, a pre-tRNA, in the absence of its protein subunit (Guerrier-Takada *et al*., 1983). In 1989 Cech and Altman were awarded the Nobel Prize in chemistry for the discoveries on the catalysis of RNA.

An extensive search for further catalytic active RNAs began soon after the discovery of the first RNA molecules with catalytic activities. It appeared that RNA-catalytic processes are widespread in nature, particularly in plants, lower eukaryotes, bacteria, and viruses. RNA catalysis was found in the genomes of viroids and virusoids, in satellite RNAs of certain viruses, in different salamander species, and the filamentous fungi *Neurospora crassa*, in the pre-tRNA processing machinery of prokaryotes, as well as in a couple of group I and group II introns.

Following the term "enzymes" for catalytic active proteins, RNA molecules with catalytic activity are called *RNA enzymes* or *ribozymes*. In addition, most ribozymes have turned out to be typical metal-dependent enzymes: binding of divalent cations (usually Mg^{2+}) is crucial to fold into their active state (for details see Pyle, 1993; Feig and Uhlenbeck, 1999; Hanna and Doudna, 2000; Pyle, 2002). In general, naturally occurring ribozymes catalyze *in cis*, which leads to their own modification. This characteristic contradicts the real definition of an enzyme. Nevertheless, the notion of "ribozyme" is justified as the natural catalytic RNAs can be engineered to specifically cleave or modify other RNAs *in trans* without becoming altered themselves. In some cases, a variation of the *in vitro* reaction conditions is already sufficient for this type of reaction (for examples see Sections VI.A and VI.C).

The different types of naturally occurring ribozymes are presented in the next section. The main part of this review concentrates on a comprehensive representation of the structurally most complicated, but also the most fascinating, ribozymes when looking at their various catalytic abilities *in vitro* and *in vivo*: the group II introns.

II. NATURAL RIBOZYMES — AN INTRODUCTION

Seven structurally different naturally occurring types of ribozymes have been identified to date (see Table 1; reviewed in Tanner, 1999). However, this might be not the final number. It is assumed also that the RNA components of many ribonucleoprotein complexes (RNPs) in the cell also possess autocatalytic activities, as, for example, the rRNA of ribosomes, snRNA components of the nuclear spliceosome, snoRNAs in the rRNA processing machinery, or guideRNAs needed for RNA editing.

Due to their differences in size and reaction mechanism, the seven types of ribozymes already known can be divided into two different groups, the small and the large ribozymes.

A. Small Ribozymes

The group of small ribozymes includes the hammerhead and hairpin motif, the HDV RNA and the VS ribozyme. These four different catalytic RNAs were found in a size range from about 40 nt up to 154 nt (see Table 1).

The hammerhead is the most frequently found catalytic motif in plant pathogenic viroids, virusoids, and other satellite RNAs, whereas the hairpin ribozyme was only demonstrated in some virusoids (reviewed in Symons, 1997). In contrast, the hepatitis delta virus is a viroid-like satellite virus of the human hepatitis B virus (HBV). HDV leads to an exceptionally strong type of hepatitis in infected patients (Lai, 1995; Gerin, 2001). The appearance of catalytic motifs in viroids and satellite RNAs is based on their typical replication mechanism. The monomeric circular molecules of these RNA species are used as a template for replication via the "rolling circle" mechanism, which leads to the formation of long linear transcripts consisting of *in tandem* joined monomers. These so-called RNA concatemers are cut into the monomeric structures by their catalytic motifs and recirculated for the next round of replication.

The fourth small ribozyme is a 154 nt part of the 881 nt VS-RNA. The VS-RNA is transcribed from the circular mono- or multimeric Varkud-plasmid localized in the mitochondria of some strains of the filamentous fungus *Neurospora crassa* (Saville and Collins, 1990; Guo *et al*., 1993). The resulting multimeric transcripts are cleaved site-specifically, similar to the RNA concatemers of viroids and satellite RNAs and ligated to monomeric RNA circles of 881 nt. The RNA monomers are used as a template to synthesize the double-stranded plasmid by reverse transcription and second strand synthesis (Kennell *et al*., 1995).

Table 1
The Seven Different Types of Naturally Occurring Ribozymes.

ribozyme	source	size	function	reaction products
self-cleaving RNAs				
hammerhead motif	plant viroids and satellite RNAs, salamander	≈ 40 nt	replication	5'-OH; 2', 3'-cyclic phosphate
hairpin motif	plant satellite RNAs	≈ 60 nt	replication	5'-OH; 2', 3'-cyclic phosphate
HDV	hepatitis delta virus (human)	≈ 80 nt	replication	5'-OH; 2', 3'-cyclic phosphate
VS ribozyme	*Neurospora crassa* mitochondria	154 nt	replication	5'-OH; 2', 3'-cyclic phosphate
RNase P RNAs	eukaryotes (nucleus, organelles), prokaryotes	140-490 nt	tRNA processing	products with 5'-phosphate and 3'OH
self-splicing RNAs				
group I introns	eukaryotes (nucleus, organelles), prokaryotes, bacteriophages	200-1500 nt	splicing	intron with 5'-guanosin and 3'-OH; 5'/3' ligated exons
group II introns	eukaryotes (organelles), prokaryotes	300-3000 nt	splicing	intron with 2'-5' lariat and 3'-OH; 5'/3' ligated exons

Due to their molecular function described above, the four groups of small ribozymes are also called *self-cleaving* RNAs. The cleavage of phosphodiester bonds seems to follow the same chemical reaction mechanism in all small ribozymes described to date. The products and the stereochemistry of the cleavage reaction assume a transesterification by a so-called "in-line" S_N2 mechanism (Figure 1A): The internal 2'-OH group of the ribose next to the phosphodiester bond to be cleaved attacks the phosphate, leading to an inversion of the configuration around the phosphorus. The incoming group is in line with the hydroxyl group in the transition state leaving the reaction center and the phosphorus takes up a typical trigonal bipyramidal geometry. The reaction yield a product with a 2', 3'-cyclic phosphate and a product with a 5'-OH terminus (Lilley, 1999b; McKay and Wedekind, 1999).

A more detailed presentation on the structure and the catalytic characteristics of small ribozymes is given in reviews by Lilley (1999b), McKay and Wedekind (1999), and Doudna and Cech (2002). Furthermore, there are also some comprehensive reviews dealing with one group of small ribozymes like the hammerhead motif (Scott and Klug, 1996; Birikh *et al*., 1997; McKay and Wedekind, 1999; Hammann and Lilley, 2002), the hairpin motif (Walter and Burke, 1998; Lilley, 1999a; Fedor, 2000; Lilley, 2001), the hepatitis delta virus (Been and Wickham, 1997; Gerin, 2001; Shih and Been, 2002), and the VS ribozyme (Collins, 2002; Lafontaine *et al*., 2002).

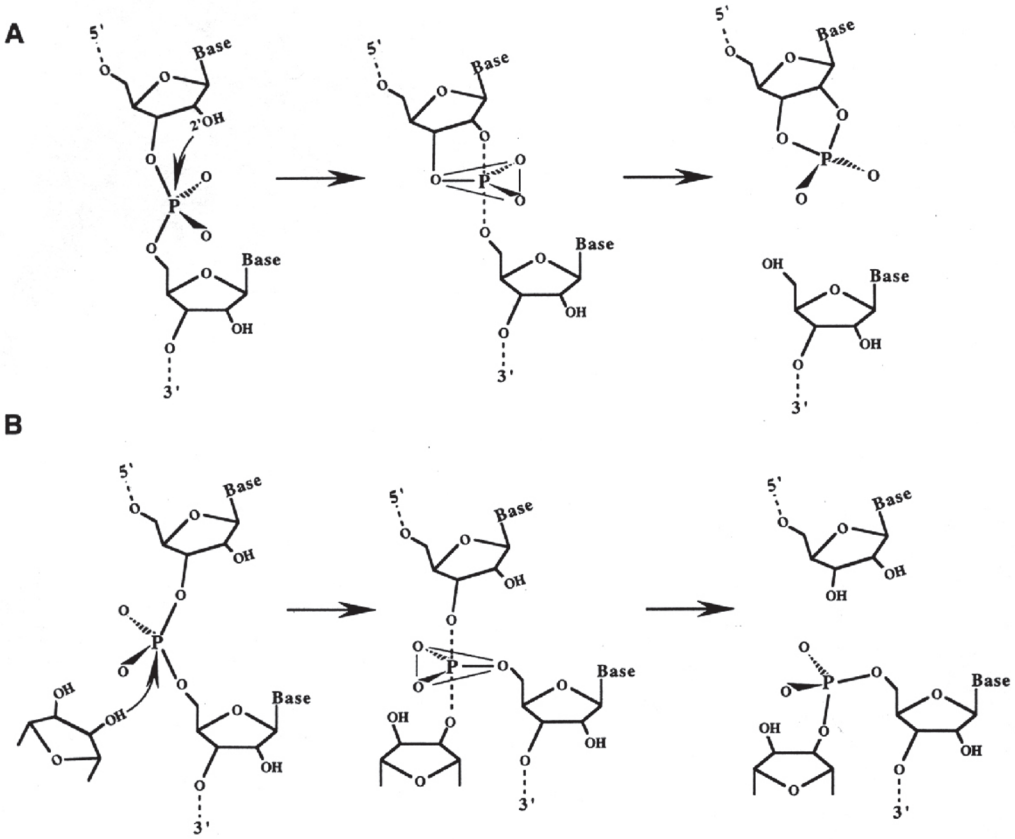

FIGURE 1. Mechanisms of cleavage of phosphodiester bonds by ribozymes. The attack of the nucleophile, the transition state, and the final products of the reactions of small (**A**) and large (**B**) ribozymes are shown. In both cases the chemical reaction is an "in line" S_N2 nucleophilic substitution with an inversion of the configuration at the phosphate. For details see the text. (Figure reprinted from *Biochimie* 78, A. Jacquier, Group II introns: elaborate ribozymes, pp. 474-487, Copyright (1996), with permission from Elsevier Science.)

B. Large Ribozymes

The RNA component of ribonuclease P and the members of the group I and group II intron family belong to the group of large ribozymes (Cech, 1993; Tanner, 1999). The molecules vary in size from a few 100 nt up to about 3000 nt (see Table 1).

RNase P is a site-specific endonuclease that processes the 5' ends of tRNAs. The endonuclease is an essential factor for the functional maintenance of the translation machinery, explaining its ubiquitous distribution. In almost all organisms investigated so far, RNase P is a ribonucleoprotein complex with a significant lower protein content in prokaryotes as in the nucleus or in the mitochondria of eukaryotes. In contrast, the RNase P from chloroplasts lacks an RNA component and appears to function as a catalytic protein.

Interestingly, the RNA component of RNase P was demonstrated to be catalytically active in only a few eubacteria. The prokaryotic RNase P consists of a 350 to 400 nt RNA and a small basic 14-kDa protein. While the RNA component shows autocatalytic activity *in vitro* under high salt conditions, *in vivo* the basic protein is essential for the ribozyme activity. Apparently, the protein supports the binding of the RNA enzyme to the substrate by electrostatic shielding. However, it is yet not clear how the substrate is bound for precise cleavage. Although an acceptor region and a binding motif was localized on the ribozyme, the primary sequence seems to be relatively unimportant. It is assumed that predominantly tertiary interactions between the ribozyme and the substrate are important for the correct catalytic function (for details see Frank and Pace, 1998; Altmann and Kirsebom, 1999; Schön, 1999; Kurz and Fierke, 2000; Thomas *et al*., 2000; Gopalan *et al*., 2002; Xiao *et al*., 2002).

The two remaining groups of ribozymes belong to the introns or intervening sequences. Introns are segments of sequences that interrupt the coding sequence of a gene, the exons. To ensure the correct expression of a gene, the introns have to be removed from the pre-RNA, a process called *splicing*. Introns can be divided into five main classes based on the differences in structure and splicing mechanisms. Apart from group I and group II introns (reviewed in Saldanha *et al*., 1993; Michel and Ferat, 1995; Lambowitz *et al*., 1999; Bonen and Vogel, 2001), the introns of nuclear tRNAs (Abelson *et al*., 1998), the introns of archaebacteria (Lykke-Andersen *et al*., 1997), as well as nuclear mRNAs (Burge *et al*., 1999; Reed, 2000) make up their own class, respectively. However, only a few representatives of group I and group II introns are known to catalyze their own excision from the pre-mRNA *in vitro* (*self-splicing introns*).

Group I introns are widely spread ribozymes and can be found in almost all organisms with the exception of higher eukaryotes. Although only a few highly conserved nucleotides and sequence elements are present, all group I introns show a typical phylogenetically conserved secondary structure in their central part, which is characterized by the base-paired regions P1-P10. The mechanism of splicing was first described for the intron in the gene for the large rRNA of *Tetrahymena* (surveys in Cech, 1990; Cech, 1993). This mechanism was principally confirmed for all group I introns analyzed to date, for example, the self-splicing reaction of the second intron in the mitochondrial cytochrome b gene from the filamentous fungus *Podospora anserina* (bI2) (Schmidt *et al*., 1992). In contrast to the small ribozymes described above, the catalytic activity of group I introns is not characterized by only one reaction, but at least by two consecutive transesterification steps. In the first step, the 5' splicing site is cleaved by the attack of an external guanosine co-factor, leading to the release of the 5' exon and an intron/3' exon intermediate with an additional G at the 5' end. In the second step, the free 5' exon attacks the 3' splicing site, followed by the ligation of the exons and the release of the linear intron. In many cases the intron is still reactive and a third transesterification reaction follows. A short sequence from the 5' end is cleaved off leading to a circularization of the intron.

As in the case of group I introns, the catalytic activity of group II introns is typically characterized by two transesterification steps. In contrast to group I introns, the first reaction step is initiated by the attack of an unpaired intron-internal adenosine, located close to the 3' end of the intron (A* in Figure 2), on the 5' splicing site, resulting in a free 5' exon and a branched intron/3' exon intermediate. In the second reaction step, the 5' exon attacks the 3' splicing site, leading to the ligation of the exons and the release of the intron in a lariat form typical for group II introns (Michel and Ferat, 1995; Jacquier, 1996). A more detailed description of the catalytic features of group II introns is given in Section VI of this review.

The three groups of large ribozymes differ from small ribozymes not only by their catalytic activities, but also in their chemical reaction mechanisms. As

described above, the catalysis of RNase P, group I and group II introns is also a nucleophilic "in line" S_N2 type substitution. However, in contrast to the small ribozymes the attack is not initiated by an adjacent internal, but by an external or a far distant internal nucleophile (Figure 1B). In the case of the hydrolytic reaction catalyzed by RNase P, the attacking hydroxyl is water and products with 5'-phosphate and 3'-OH ends are formed. Group I introns use the 3'-OH of the free guanosine and group II introns the 2'-OH of the internal adenosine as the nucleophile for the first splicing step. The second step in both intron groups is initiated by the attack of the 3'-OH from the released 5' exon on the 3' splice site. The resulting products are also carrying 5'-phosphate and 3'-OH ends (Cech, 1993; Jacquier, 1996).

III. DISTRIBUTION OF GROUP II INTRONS

In 1982 the introns of group II were defined for the first time as an independent structural intron class using only four known sequences from two different mitochondrial genes of the baker's yeast *Saccharomyces cerevisiae* (Michel *et al*., 1982). The number of known group II introns has grown to more than 200 members and is still growing rapidly. Group II introns were found in low frequency in the mitochondrial genomes of fungi, sporadically in the organellar genomes of algae, and are numerous in the organellar genomes of higher plants (Michel *et al*., 1989). The plastid DNA of *Euglena gracilis* represents a special case, in which the enormous number of 91 group II introns could be identified (Doetsch *et al*., 1998). Furthermore, group II introns were discovered in proteobacteria and blue algae, which are regarded as potential "ancestors" of mitochondria and chloroplasts (Ferat and Michel, 1993; Ferat *et al*., 1994; Mills *et al*., 1996; Shearman *et al*., 1996). New data predominantly derived from different sequencing projects of prokaryotic organisms show that group II introns are also surprisingly widespread in the bacterial world (reviewed in Martinez-Abarca and Toro, 2000a; Dai and Zimmerly, 2002), including the archaebacteria (Dai and Zimmerly, 2003).

IV. GENERAL STRUCTURE OF GROUP II INTRONS

When group II introns were defined as an independent class, a first common secondary structure model was established on the basis of the few known phylogenetic data (Michel *et al*., 1982; Schmelzer *et al*., 1982; Michel and Dujon, 1983; Schmelzer *et al*., 1983). This early model was mainly confirmed by biochemical analysis (Kwakman *et al*., 1990; Chanfreau and Jacquier, 1994). The secondary structure of group II introns is characterized by six typical stem-loop structures, also called domains I to VI or D1 to D6 (Figure 2). The domains radiate from a central core that brings the 5' and 3' splice junctions into close proximity. The proximal helix structures of the six domains are connected by a few nucleotides in the central region (linker or joiner sequences). Due to its enormous size, the domain 1 was divided further into subdomains a, b, c, and d.

On the basis of 70 published sequences, some anatomic and sequence differences of group II introns were identified which led to a further division into subgroups IIA and IIB. Each subgroup could be further divided into two subfamilies (A1, A2, B1, B2) (Michel *et al*., 1989). With the fast-growing number of sequenced group II introns, additional structural diversity became apparent leading to the definition of four new, less typical classes of bacterial group II introns. These are classified as A (with group IIA and IIB features), B (group IIB2-like with many unique features), C (hybrid intron structure with unusual features), and D (group IIB-like with some IIA features) (for more details see Toor *et al.,* 2001). Nevertheless, the classification into the subgroups IIA and IIB will also be used in the following sections for two reasons: (1) All group II introns that have been objects for a detailed functional and structural analysis belong to one of the two subgroups. (2) The structural differences of group IIA and group IIB introns reflect remarkable differences in the autocatalytic splicing behavior of the introns (see Section VI.A.1 and VI.A.2).

As can be seen in Figure 2, group II introns possess only a very few conserved nucleotides, and the nucleotides important for the catalytic function are spread over the complete intron structure. The few strictly conserved primary sequences are the consensus at the 5' and 3' splicing site (…↓GUGYG… and …AY↓…), some of the nucleotides of the central core (joiner sequences), a relatively high number of nucleotides of D5 and some short sequence stretches of D1. The unpaired adenosine in D6 marked by an asterisk (7 or 8 nt away from the 3' splicing site, respectively) is also conserved and plays a central role in the splicing process (see Section VI.A).

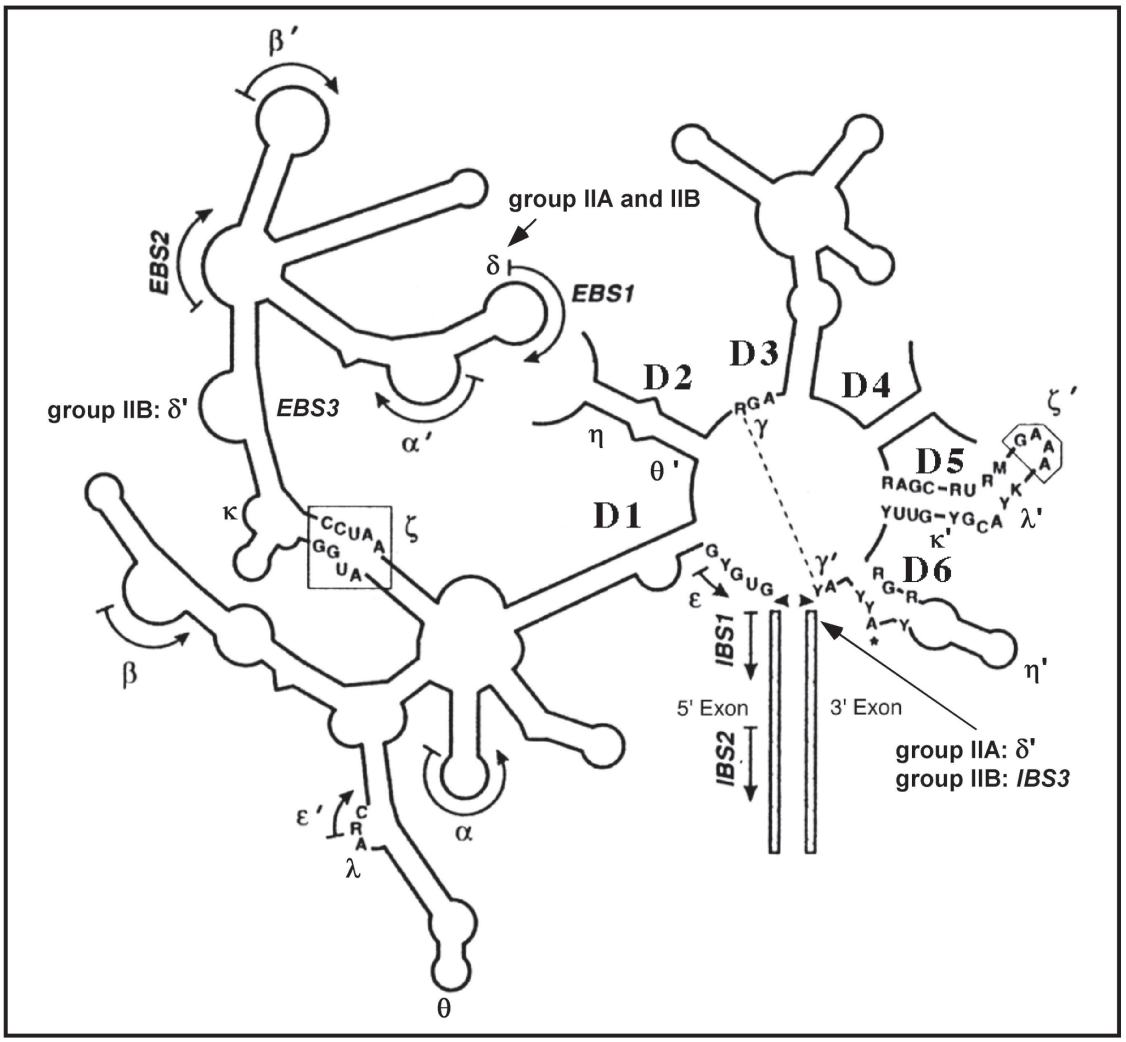

FIGURE 2. Structure and consensus nucleotides of a group II intron. The figure shows the predicted secondary structure model of a conventional group II intron patterned after previously published Figures (Schmidt *et al.*, 1998; Perlman and Podar, 1996). The intron structure is characterized by six major domains (D1 to D6) radiating from a central wheel. The nucleotides shown (R = purine, Y = pyrimidine, M = A or C, K = G or U) are conserved in both subgroups (IIA and IIB). Tertiary interactions important for correct folding or splice site selection are marked by large arrows, broken lines, or boxes and labeled with Greek lettering (IBS = intron binding site, EBS = exon binding site). The branch-point A is marked by an asterisk.

This low sequence conservation in combination with the rare findings of phylogenetic co-variations has impeded the identification of tertiary interactions within the intron and the intron and its upstream and downstream exon sequences. The following interactions have been identified to date (see also Figure 2). The Watson-Crick pairings between the intron and the 5' exon via IBS-EBS interactions 1 and 2 (IBS = intron binding site, EBS = exon binding site). The pairings of the intron and the 3' exon either via the δ-δ' in subgroup IIA or the IBS3-EBS3 interaction in subgroup IIB. Some pairings within the large domain 1 (α-α', ε-ε', β-β') and a few tertiary contacts between different intron domains (γ-γ', ζ-ζ', η-η', θ-θ', κ-κ' and λ-λ') (reviewed in Michel *et al.*, 1989; Michel and Ferat, 1995; Qin and Pyle, 1998) (for κ-κ' and λ-λ' see Boudvillain and Pyle, 1998; Boudvillain *et al.*, 2000).

The relatively low number of genetically or biochemically provable tertiary interactions, the enormous size of the molecules, as well as the lack of a compact planar structure compared with other types

of introns, entail that the overall spatial folding of group II introns is only partially understood, and a complete three-dimensional model is not available. However, excellent biochemical strategies and techniques have been established in the meantime, allowing, for instance, a systematic analysis of the sugar-phosphate backbone or the identification of unusual base-base contacts that have led to the development of first three-dimensional models of the catalytic core of group II introns (for details see Section VII).

V. DEGENERATED AND *TRANS*-SPLICED GROUP II INTRONS

Most of the group II introns known so far show the typical structural features described above. However, some other types of group II intron organization have been found.

In plant organelles degenerated group II introns could be detected, in which some of the common structural elements are completely missing (Michel *et al*., 1989). For instance, lots of the group II introns identified on the plastid DNA of *Euglena gracilis* are quite short (≤ 600 nt), lacking an identifiable domain 1, 2, 3, or 4. The degree of relationship to group II introns for many of those small introns is even so low that they have been arranged in a class of their own: the group III introns. Nevertheless, the degenerated introns are fully functional. As shown for the degenerated *Euglena* group II and III introns, they seem to depend on *trans*-acting factors (RNAs or proteins) for their splicing reaction. Therefore, these introns are considered as an evolutionary link between the conventional *cis*-splicing group II introns and the nuclear introns that depend on a complex *trans*-acting splicing machinery, the spliceosome (Christopher and Hallick, 1989; Copertino and Hallick, 1993; Doetsch *et al*., 2001).

It has been shown that some group II introns from plastid and mitochondrial DNA of higher plants and chloroplast DNA of *Chlamydomonas reinhardtii* reorganize a fully spliceable complete intron structure from two or more pieces *in trans* (reviewed in Butow and Perlman, 1991; Bonen, 1993). The affected discontinuous organized genes consist of exons spreadout over the genome that are flanked by 5' and/or 3' regions of group II introns. It is assumed that such an arrangement of single parts of genes over the genome is due to DNA rearrangements within the group II introns. Most of the *trans*-spliced introns consist of two parts that are joined to functional introns via base pairings in the proximal helical regions of D3 or D4. In a few cases, for example, the *psaA* gene from *C. reinhardtii*, even a tri-molecular intron-complex is formed. The intron is assembled by the exon flanking parts and an independently transcribed middle part encoded by the *tscA* gene. The three parts are joined by base pairings in the proximal helices of D1 and D4 (Kück *et al*., 1987; Goldschmidt-Clermont *et al*., 1991).

VI. CATALYTICAL PROPERTIES OF GROUP II INTRONS

Despite the large number of known group II introns, only a few splice autocatalytically from their pre-mRNA *in vitro*. Although only a few examples of introns have been published that are not reactive under all detected and optimized *in vitro* conditions (for examples see Schäfer *et al*., 1991; Costa *et al*., 1997a), it is generally assumed that most of the group II introns possess only little or no autocatalytic activities.

Because autocatalysis depends on quite unphysiological reaction conditions (see below), it was postulated that even introns with an efficient *in vitro* splicing reaction depend on proteins *in vivo* (Lambowitz and Perlman, 1990). Furthermore, the actual data of different model organisms show clear evidence for a protein-dependent catalysis *in vivo*. Proteins interacting with intron RNAs seem to support RNA folding or to stabilize the active conformation, whereas the catalytic potential is clearly located in the RNA itself (see Sections VIII and IX). Let us, however, first answer the question what kind of reactions are catalyzed by the known self-splicing introns *in vitro*?

A. The Splicing Reaction

The ability of group II introns to splice autocatalytically was first described in 1986 for both the group IIB introns, aI5γ and bI1, from the mitochondrial genome of the baking yeast *Saccharomyces cerevisiae* (Peebles *et al*., 1986; Schmelzer and Schweyen, 1986; van der Veen *et al*., 1986). The intron aI5γ is one of up to three group II introns located in the gene for the first subunit of cytochrome-C-oxidase (*COXI*), whereas the intron bI1 is the single group II intron in the cytochrome b gene (*COB*).

Basically, the *in vitro* splicing reaction of group II introns is characterized by two consecutive reaction steps, the cut at the 5' and the 3' splice site. The first reaction step can be initiated by two different mechanisms competing with each other (Peebles *et al*., 1987; Jarrell *et al*., 1988b; Daniels *et al*., 1996). The nucleophilic attack on the 5' splice site in the first mechanism is started by the 2'-OH group of the unpaired adenosine in domain 6 of the intron (see Figure 2), resulting in the release of the 5' exon and an intron/3' exon lariat (branch-point splicing, Figure 3A). The lariat is created by formation of a 2'-5' phosphodiester bond that links the first residue of the intron to the unpaired adenosine. Alternatively, the first splicing step is initiated by the nucleophilic attack of water or OH^-, resulting in a free 5' exon and a linear intron/3' exon intermediate (hydrolytic splicing, Figure 3B). The released 5' exon remains linked to the intron via base pairings of the intron binding sites (IBS1 and IBS2) to the exon binding sites (EBS1 and EBS2) of the intron (for details see Section VII.A). The second reaction step in both mechanisms is initiated by the nucleophilic attack of the 3'-OH from the terminal 5' exon nucleotide on the 3' splicing site, leading to the ligation of the two exons and the release of the intron in either a lariat or a linear form. Recent data for intron aI5γ show that group II introns can also excise in a true circular form. However, circle formation requires the release of the 3' exon from the precursor by a *trans*-splicing mechanism. The 2'-OH of the terminal intron nucleotide attacks the 5' splice site, and a circle is created by formation of a 2'-5' phosphodiester bond linking the first and the last intron nucleotide (Figure 3C) (Murray *et al*., 2001). Release of the 3' exon could be initiated by a free 5' exon that attacks the 3' splice site. The free exon might be generated by a further well-known intron-dependent reaction: the hydrolytical cleavage of the joined exons at their ligation site, the so-called "spliced-exon reopening" or short SER (Figure 3D). The SER reaction is independent of the conformation of the free intron (lariat or linear molecule), and the intron acts as a true ribozyme leaving the reaction without any modification (Jarrell *et al*., 1988b; Daniels *et al*., 1996).

As described for most ribozymes, group II introns are also strictly dependent on Mg^{2+}-ions to fold into the correct tertiary structure and for their catalytic function (reviewed in Pyle, 1993; Pyle, 2002). Furthermore, an optimum temperature of 45°C is characteristic for the *in vitro* catalytic reaction of all group II introns.

1. Splicing Reaction of the Group IIB Introns

Interestingly, when looking at the efficiency of the *in vitro* splicing reaction, the small introns of group IIB (≤ 1 kb) turned out to be much more reactive than the large introns of group IIA (≈ 2.5 kb). The size difference can be explained primarily by the presence of a long open reading frame (ORF) in domain 4 of the known self-splicing group IIA introns (for details see Section VIII.A). Group IIB introns can be spliced in low-salt buffers with low Mg^{2+} concentrations, even though the reaction is slow and less efficient. A typical reaction mix, already used in the first published autocatalytic experiments, contains 10 m*M* Mg^{2+}, 2 m*M* spermidine, and 40 m*M* Tris at pH 7.6 to 7.8 (Peebles *et al*., 1986; Schmelzer and Schweyen, 1986; van der Veen *et al*., 1986). Spermidine is a polyamine that stabilizes the tertiary structure of RNA, similar to Mg^{2+}. Later on it was shown that the addition of spermidine is not required when the concentration of Mg^{2+} in the reaction mix is raised to at least 100 m*M* (Peebles *et al*., 1987; Jarrell *et al*., 1988b). Using this reaction conditions, all products of the two splicing pathways (transesterification and hydrolysis), the circularization and the hydrolytic SER can be observed. As no accumulation of intron/3' exon intermediates can be detected, the cut at the 5' splice site seems to be the rate-limiting step of the reaction. The reaction efficiency of group IIB introns can be enhanced by high concentrations of a salt with a monovalent cation. Preferred salts are $(NH_4)_2SO_4$, NH_4Cl, and KCl in the range of 0.5 to 1.5 *M* (Kück *et al*., 1990; Schmidt *et al*., 1990). The chosen monovalent salt determines which of the two splicing pathways is preferred: $(NH_4)_2SO_4$ yield the highest rate of transesterification products, KCl the highest rate of hydrolytic products, while in reactions using NH_4Cl all the different products are detectable in similar concentrations. The hydrolytic cleavage of the ligated exons can be observed in all buffer systems as well, but is most strongest with added KCl (Peebles *et al*., 1987; Jarrell *et al*., 1988b; Daniels *et al*., 1996).

2. Splicing Reaction of Group IIA Introns

Using the reaction conditions mentioned above, the large introns of group IIA are either nonreactive or much less reactive when compared with the small group IIB introns. This was first shown for the group IIA intron located in the mitochondrial gene encoding

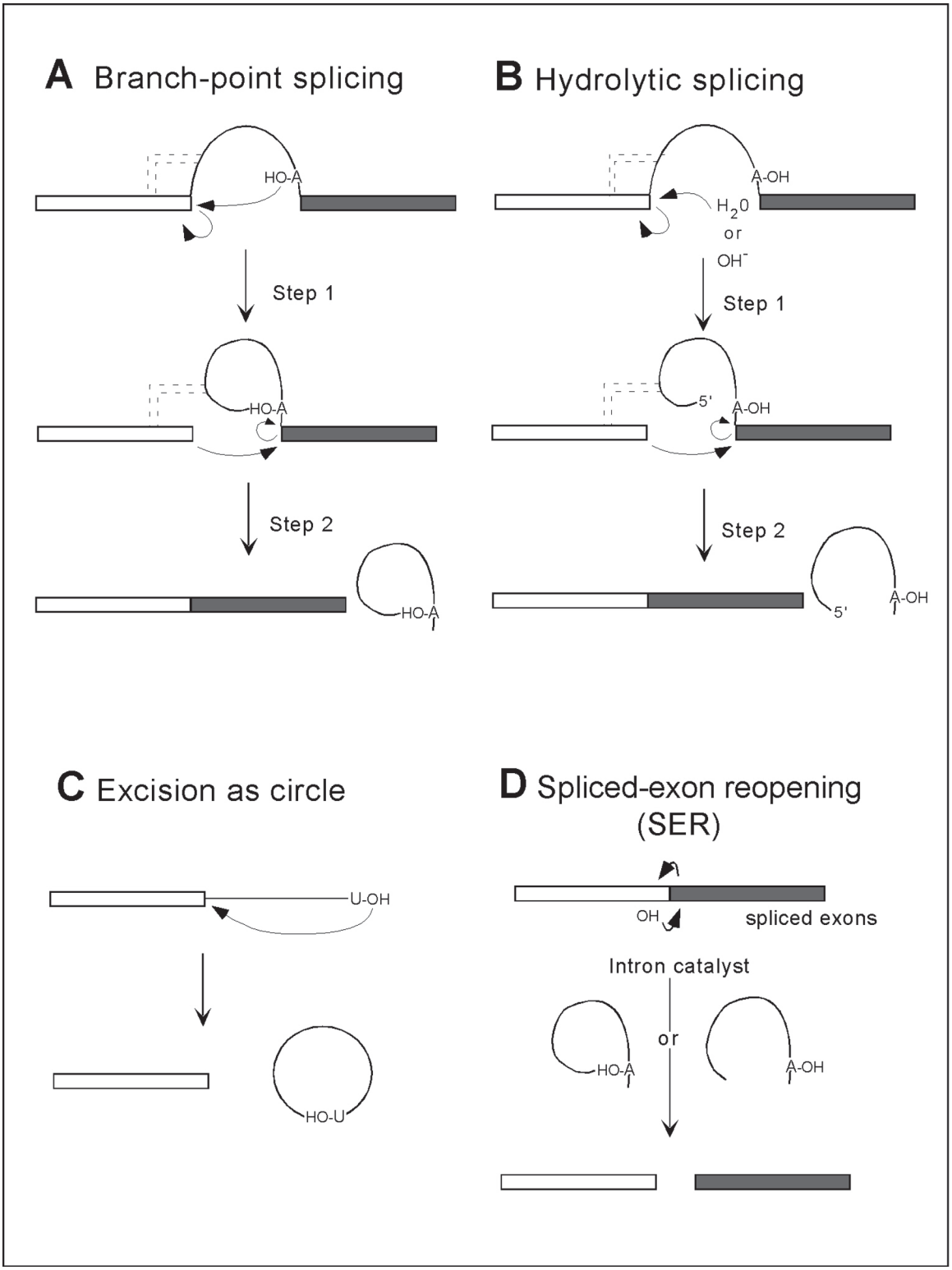

FIGURE 3. The reaction pathways of group II introns during self-splicing *in vitro*. The intron is shown as a solid black line flanked by the 5' (white box) and 3' exon (grey box). The nucleophilic attacks initiating the different reaction steps are marked by arrows. The IBS/EBS interactions important for binding of the 5' exon to the intron after the first reaction step (reactions A and B) are indicated by broken lines. For more details see the text.

the first subunit of the cytochrome-c-oxidase of the filamentous fungus *Podospora anserina* (*COXI* I1) (Schmidt *et al*., 1990). The analysis of intron aI1, one out of two group IIA introns in the mitochondrial *COXI* gene of the yeast *Saccharomyces cerevisiae*, revealed similar results (Hebbar *et al*., 1992). The *in vitro* data for both introns can be summarized as follows.

Apparently, both introns are not reactive under low-salt conditions. The splicing reaction depends on a high concentration of a monovalent salt, for example, 1 *M* NH_4Cl or higher. However, even under optimized conditions, only weak splicing activity can be detected for the *Podospora* intron and the 5' splice site is exclusively cleaved hydrolytically. In comparison, the *Saccharomyces* intron is more reactive and products resulting from hydrolysis and transesterification at the 5' splice site are generated. To facilitate the analysis of ribozyme activity, the introns were shortened in D4 to 1 kb or 1.5 kb, respectively. In the nonconserved D4, comprising 78% (!) of the complete intron sequence of the *Podospora COXI* I1, the major part of a protein-coding open reading frame (ORF) is looped out, a typical feature of many group IIA introns (see also Section VIII.A). In addition, it was already known from deletion experiments of group IIB intron aI5γ that the removal of D4 down to a short helical structure connecting D3 and D5 in the central core has no negative effects on the autocatalytical reaction of the intron (Jarrell *et al*., 1988a). The shortened group IIA introns showed an even better reaction, and furthermore the 5' cut via transesterification and formation of the intron lariat for the *Podospora* intron could be restored with low efficiency. As already observed for group IIB introns, the reaction mechanism shifts in favor of hydrolysis at the 5' splice site when using KCl as the monovalent salt. Under $(NH_4)_2SO_4$ conditions favorable for transesterification of group IIB introns, the introns of group IIA are barely reactive, although they show a spectra of products similar to NH_4Cl reactions.

Moreover, the efficient SER reaction observed for group IIB introns is only barely detectable for both the group IIA introns mentioned above. Even under SER-optimized conditions with high KCl concentrations, a free 3' exon can only barely be found (Schmidt *et al*., 1990; Hebbar *et al*., 1992). The free 5' exon has to be regarded as an intermediate of the hydrolytic pathway, because only the linear intron/3' exon, but not the intronlariat/3' exon, is detectable. This observation indicates that cleavage at the 5' splice site is again the rate-limiting factor of the branch-point dependent reaction, while, in contrast to group IIB introns, cleavage at the 3' splice site seems to be rate-limiting step of the hydrolytic reaction.

3. 5' Hydrolysis In Vivo

When looking at the *in vivo* splicing reaction of group II introns, it was assumed for a long time that the cut at the 5' splice site occurs exclusively via transesterification preventing any hydrolytic reaction.

Two strongly accumulating intron-specific bands for every group II intron were detected in mitochondrial RNA preparations of *Saccharomyces cerevisiae* and *Podospora anserina* when running denaturing agarose gels (Bonitz *et al*., 1980; Halbreich *et al*., 1980; Hensgens *et al*., 1983; Kück *et al*., 1985; Zennaro *et al*., 1985). Both bands appear as circular and linear molecules of the same length under the electron microscope (Arnberg *et al*., 1980; Halbreich *et al*., 1980; Hensgens *et al*., 1983; Schmidt *et al*., 1987b). Later on, the typical branched circular structure (with a 2'–5' phosphodiester bond and a free 3'-OH end) could be determined for the circular form, which seemed to be responsible for the accumulation of the molecules due to resistance against cellular nucleases (van der Veen *et al*., 1986). Transcript mapping and enzymatic analysis (RNaseH digestion, debranching-assay) of the 2.5-kb intron aI1 of *Podospora anserina* leads to the conclusion that the slower-running molecule is the intron lariat, whereas the faster-running form corresponds to the linear intron (Schmidt *et al*., 1987b). At that time it was generally assumed that the linear species was a branched linear molecule originating from unspecific breakage of the lariat form.

Although evidence for a first-step hydrolytic cut in the *in vitro* reaction of group II introns was published at the same time (Peebles *et al*., 1987; Jarrell *et al*., 1988b) it took until 1996 for Pyle and co-workers to present for the first time the idea of a 5' hydrolysis also in the *in vivo* reaction (Daniels *et al*., 1996). This notion was predominantly based on the following observations:

1. All known *in vitro* splicing conditions yield products of the hydrolytic pathway at least to a certain degree.
2. The intracellular salt concentrations in combination with other components that support ca-

talysis (polyamines, cationic peptides) should provide *in vivo* reaction conditions similar to the optimized conditions described for 5' hydrolysis *in vitro*.

3. When digesting the two observed *in vivo* forms of the aI1 intron of *P. anserina* by RNaseH (Schmidt *et al*., 1987b), the resulting single band is not accompanied by a smear that should be present if the linear species contains only statistically broken lariat molecules.
4. In some chloroplast group II introns, the unpaired adenosine at the branch point is missing (Michel *et al*., 1989) suggesting *in vivo* splicing using the linear, hydrolytic pathway.

Two years later Podar *et al*. (1998a) could demonstrate that first-step hydrolysis works also *in vivo*. They used yeast strains carrying a point mutation at the branch site of group IIB intron aI5γ, the terminal intron of the *COXI* gene, leading to a strong reduction or even block of the 5' transesterification reaction. Under optimized conditions a hydrolytical turnover of more than 40% of the *COXI* transcripts could be measured. Recently, Vogel and Börner (2002) presented the first evidence for naturally occurring hydrolytic group II intron splicing *in vivo* using the barley chloroplast *trnV* intron of $tRNA^{Val}$(UAC). The *trnV* intron is strongly conserved in all known plant chloroplast genomes and lacks the typical branch-point nucleotide in domain 6 important for transesterification and lariat formation. Following cirularization of chloroplast total RNA using RNA ligase and a RT-PCR based approach, it could be clearly shown that the *trnV* intron is exclusively released in a true linear form.

Generally, it is accepted that from the evolutionary point of view hydrolysis might be the primordial reaction mechanism for release of a group II intron. The reaction center that prefers the more efficient transesterification was developed later on in evolution. Thus, the known group II introns that show no typical branch point nucleotide might belong to a very ancient form of introns whose splicing reaction is still initiated via first-step hydrolysis. In addition, the hydrolytic pathway might be important for regulation of other group II intron-specific features (Daniels *et al*., 1996; Podar *et al*., 1998a):

1. Many group II introns contain an open reading frame (ORF) encoding a protein that is involved in the splicing reaction and the mobility of the introns (see Sections VIII.B and VIII.C) (surveys in Lambowitz and Belfort, 1993; Lambowitz *et al*., 1999; Zimmerly *et al*., 2001). Although the intron-encoded proteins (IEPs) are thought to be translated from a pre-RNA, the expression from an excised, linear intron RNA might be also possible. However, circular or branched circular molecules are not translatable (Kozak, 1979; Konarska *et al*., 1981), unless they contain special internal ribosome entry site elements (Chen and Sarnow, 1995). Thus, linear translatable intron RNAs resulting from a hydrolytic reaction could be essential molecules for splicing via transesterification that depends on an intron-encoded protein.
2. For the mobility of group II introns both the expression of the intron encoded protein and the lariat structure of the free intronic RNA is essential (Michel and Ferat, 1995; Bonen and Vogel, 2001; Belfort *et al*., 2002). Thus, introns with a high rate of hydrolytical splicing should show a low frequency of mobility.

B. The Reverse Splicing Reaction

The reaction potential of group II introns is not restricted to the findings described above. A number of other *in vitro* activities have been discovered for the more reactive group IIB introns, in particular demonstrating their catalytic versatility.

In principle, when looking at the chemistry of phosphodiester transfers during autocatalysis of group II introns, this reaction should also work in reverse direction. Shortly after detecting a complete reverse *in vitro* splicing reaction for the *Tetrahymena* group I intron (Woodson and Cech, 1989), a reverse splicing was also described for the group IIB intron bI1 of yeast (Augustin *et al*., 1990; Mörl and Schmelzer, 1990b). The reverse reaction is initiated by binding the free intron lariat via the EBS elements (EBS1 and EBS2) located in domain 1 to the IBS elements (IBS1 and IBS2) in the 5' exon region next to the exon ligation site. The 3'-OH of the intron attacks the ligation site of the exons and the intron is integrated via two consecutive transesterifications, resulting in the original pre-RNA. Furthermore, Mörl and Schmelzer (1990b) found a certain dependence of the efficiency of the reaction on 3' exon sequences. Later on, this dependence was confirmed with the detection of a contact of the first nucleotide of the 3' exon with a

nucleotide located in subdomain d of domain 1, the EBS3-IBS3 interaction (see also Section VII.G) (Jacquier and Jacquesson-Breuleux, 1991; Costa *et al*., 2000). Remarkably, the two transesterification reactions can be separated. When the free intron is delivered in linear form, only the reversal of the second step leading to the intron/3' exon intermediate takes place. For reversion of the first step, the energy of the 2'–5' phosphodiester bond not present in the linear molecule is essential. When the free intron lariat or the lariat intron/3' exon intermediate is incubated with the free 5' exon, reversal of the first step and formation of the pre-RNA can also be observed (Augustin *et al*., 1990; Mörl and Schmelzer, 1990b). Moreover, Chin and Pyle (1995) could show that the forward branching reaction is highly reversible in the absence of the second splicing step. Their data indicate that the rapid second step serves as a kinetic trap to complete the forward reaction, and that the efficient reversibility provides a potential proofreading mechanism to control the fidelity of 5' splice site selection.

Alternatively, reversal of the first step is also possible with a linear intron/3' exon carrying a triphosphate at its 5' end. The pyrophosphate is cleaved off and the energy is used to bind the 5' exon. This ribozyme-catalyzed condensation of an activated 5' tri-phosphate with a 3'-OH to form a 3'–5' phosphodiester bond is similar to the reaction of today's protein DNA- or RNA-polymerases and therefore very interesting from an evolutionary point of view. This reaction mechanism could have also been used by a prebiotic RNA-dependent replicase for polymerization of RNA (Mörl *et al*., 1992).

In contrast to the forward reaction, only buffer systems with $(NH_4)_2SO_4$ as the essential salt component are useful for the reverse reaction. A concentration of 500 m*M* $(NH_4)_2SO_4$ turned out to be optimal, whereas the temperature optimum of the reverse reaction is 30°C compared with 45°C for the forward reaction. The use of high concentrations of KCl or NH_4Cl activates the hydrolytic mechanism leading to an exon-reopening of the ligated exons or to an unspecific intron-internal cleavage, respectively (Perlman and Podar, 1996).

C. Other Reactions Catalyzed by Group II Introns *in trans*

It became clear from the analysis of the reverse splicing reaction that the structural requirements of an RNA substrate to be converted by a free group II intron RNA are quite simple. The IBS1, which comprises only a few nucleotides, is completely sufficient as the minimal substrate recognition site for a group II intron-catalyzed reaction (Jacquier and Michel, 1987). Additional sequence elements like IBS2 or the IBS3 (δ') nucleotides are obviously only responsible for an increased reactivity and for the reduction of cleavage at cryptic sites (Mueller *et al*., 1988; Jacquier and Jacquesson-Breuleux, 1991; Wallasch *et al*., 1991; Su *et al*., 2001). These simple structural requirements of a specific substrate/ribozyme interaction suggested a high potential of group II introns to catalyze other RNA-dependent reactions.

In addition to reverse splicing, Mörl and Schmelzer (1990b) also demonstrated the complete integration of the lariat of intron bI1 into a foreign RNA *in vitro*. Integration takes place directly upstream of a sequence motif similar to IBS1. The reaction is reversible, and the intronic RNA can be removed again from the foreign RNA under standard conditions of the forward reaction.

Moreover, some activities have been described in which the intron lariat even acts as a true ribozyme, that is, the intron molecule reacts with a substrate, leaves the reaction without any modification, and repeats the same reaction in multiple cycles ("multiple-turnover" ribozyme):

RNA recombinase: the intron lariat catalyzes recombination of two different IBS1-carrying RNA-molecules (RNA 1 and 2). The recombination sites of both molecules are precisely located downstream of a typical IBS1 sequence. The 5' region of RNA 1 is joined to the 3' region of RNA 2 and the 5' region of RNA 2 to the 3' region of RNA 1(Mörl and Schmelzer, 1990a).

Terminal transferase: when using a substrate consisting of a 5' exon ligated to a mono-phosphate (5'E-p), the 3' terminal phosphate is transferred to the intron via the reverse reaction and can (like a 3' exon) be coupled again to the 3' OH of a free 5' exon using the forward reaction (Mueller *et al*., 1991).

RNA polymerase: when incubating the bI1 intron-lariat with a short 19-mer oligoribonucleotide which corresponds to the 3' end of the 5' exon including the IBS1 sequence (GACAGA), an intron-dependent disproportion of the substrate-RNA at the 3' end takes place. The resulting molecules are either shortened by the 3' terminal adenosine or elongated by one or more adenosine residues. The terminal adenosine of

the oligoribonucleotide is cleaved off in the presence of the intron lariat (18-mer) and coupled in a reverse reaction (like a 3' exon) to the 3' end of the lariat (intron charging). The charged intron can now react in a forward reaction (intron-discharging) with a complete 19-mer oligoribonucleotide to give rise to a molecule elongated by one nucleotide. This cycle of precise charging/discharging of the lariat can be repeated many times also using substrate molecules as donors that have been elongated already by one ore more nucleotides. In this case, both the last adenosine residue of the IBS1 sequence and all other 3'-coupled nucleotides are transferred. Thus, we deal with a kind of intron-catalyzed nucleotidyl-insertion leading to a 3' to 5' RNA polymerization (Mueller *et al*., 1993a).

DNA as a substrate: in contrast to other ribozymes, group II introns are also able to use single-stranded DNA as a substrate. Similar to the reverse coupling of a 5' exon RNA-substrate to the triphosphate activated end of a linear ^{ppp}IVS-3'E RNA, also a 5' DNA substrate can be coupled at the activated molecule; however, with much lower efficiency. Furthermore, the lariat and the activated linear molecule (^{ppp}IVS) are also able to specifically cleave a single-stranded DNA corresponding to the ligated exons in a reverse reaction upstream of the IBS1 and to couple the 3' exon DNA to the 3' end of the intron RNA. However, a complete integration of the intron RNA into the DNA could not be demonstrated (Mörl *et al*., 1992). Applying the standard *in vitro* reaction conditions, the complete reverse reaction only succeeds when using a chimeric substrate molecule with a 5' exon RNA sequence and a 3' exon DNA sequence (Mueller *et al*., 1991).

The fundamental findings described above culminate in the development of a short multiple turnover ribozyme consisting of two separated parts (domain 1 and 5) of group IIB intron aI5γ. The new ribozyme cleaves short single-stranded RNA and DNA substrates with almost the same efficiency, a characteristic not known for any other catalytic RNA. The free domain 1 binds the oligonucleotide substrate and after addition of the free domain 5 as the essential catalytic component (see Section VII.E) the substrate is specifically cleaved downstream of the IBS1 sequence (Griffin *et al*., 1995; Michels and Pyle, 1995). The chemical turnover rate of the ribozyme can be improved by joining domain 3 to domain 1 via a short hairpin loop. Obviously, domain 3 seems to act as a kind of catalytic effector (Xiang *et al*., 1998).

The knowledge that group II introns are able to specifically cut foreign RNA and even DNA substrates downstream of a IBS1-like sequence and to integrate completely at least in RNA *in vitro* gave impetus to the evolutionary and molecular biological important search, if today's group II introns are also able to transpose into ectopic DNA sites *in vivo*. Indeed, some time later *in vivo* transposition could be proven for two mitochondrial and two bacterial group II intron, as described in Section VIII.B.2 (Mueller *et al*., 1993b; Sellem *et al*., 1993; Cousineau *et al*., 2000; Martinez-Abarca and Toro, 2000b; Ichiyanagi *et al*., 2002).

VII. THE STRUCTURE ELEMENTS OF GROUP II INTRONS AND THEIR CONTRIBUTION TO THE CATALYTIC REACTION

One of the most critical features when analyzing the structure/function relationships of group II introns is their enormous size compared with other ribozymes. For instance, lots of important sequence stretches and substructures that are involved in the formation of tertiary interactions have not been recognized *in cis* using the complete intron molecule. This problem is caused by the potentially high degree of redundancy of different tertiary interactions within the intron. The elimination of a tertiary contact by mutation might not have any influence on the overall structure and reaction of the intron, because other contacts that still exist are sufficient for correct folding and efficient catalysis. In addition, when analyzing intron mutations *in cis* that completely block the splicing reaction, it is almost impossible to differentiate if the overall tertiary architecture of the intron, the catalytic reaction itself, or both is affected.

To avoid this problem, partial intron structures based on the six phylogenetically defined intron domains (D1 to D6) were generated and analyzed *in trans* reactions. The six intron domains are transcribed *in vitro* as separated molecules that fold into a structure corresponding to their conformation in the complete active intronic RNA. For instance, the isolated D5 binds *in trans* to the rest of the intron or only to D1 with high affinity (Jarrell *et al*., 1988a; Franzen *et al*., 1993; Pyle and Green, 1994). When incubating D5 with a RNA molecule containing the 5' exon and the intron domains 1 to 3 (exD123), a hydrolytic cut at the 5' splicing site takes place that is mechanically,

kinetically, and stereochemically identical to the first step of the *cis*-reaction of the intact intron (Pyle and Green, 1994; Michels and Pyle, 1995; Podar *et al*., 1995a). Moreover, when applying the *trans*-reaction with mutated intron domains, it can be directly investigated if a mutation affects binding, catalysis, or both (Peebles *et al*., 1995; Abramowitz *et al*., 1996; Konforti *et al*., 1998b). Beside D5, also D1 (Michels and Pyle, 1995; Podar *et al*., 1995b) and D6 coupled to D5 (Dib-Hajj *et al*., 1993; Chin and Pyle, 1995; Boudvillain and Pyle, 1998) were tested successfully in *trans*-reactions for their functional importance. More recently, a tri-partite assay system was developed using a short free 5'exon, a linear intron truncated by 6 nt at the 3' end, and a 3' molecule consisting of the last 6 nt of the intron plus a 6 nt 3' exon to bypass the rate-limiting 5' cut for better studying the exon ligation at the 3' splice site (Bar-Shalom and Moore, 2000).

A. Domain 1

Domain 1 (D1) is the largest of all intron domains, when not considering the optional open reading frame (ORF) looped out in domain 4 (D4) of many group II introns. Due to its large size, the domain has been divided into the four subdomains a, b, c, and d. D1 contains all the sequence elements required for the correct recognition of the 5' splice site *in cis* (see Qin and Pyle, 1998 for a review). Moreover, it is also indispensable for recognition and binding of DNA or RNA substrates *in trans* or in the reverse splicing reaction, respectively (Pyle, 1996; Eskes *et al*., 1997; Guo *et al*., 1997). Deletion studies with yeast intron aI5γ showed that *in vitro* D1 and D5 are sufficient for catalytic activity and the cut at the 5' splice site both *in cis* (Koch *et al*., 1992) and *in trans* (Michels and Pyle, 1995). As substrate binding of D1 is not influenced by the addition of other domains important for catalysis *in trans* (D3, D5), it is assumed that D1 delivers the molecular scaffold of defined architecture that assembles the intron into its catalytic active structure (Qin and Pyle, 1997).

In all functional group II introns two exon binding sites (EBS1 and EBS2), located in different stem-loop regions of subdomain d, could be identified as the crucial elements for precise recognition and binding of the 5' exon (see Figure 2; Jacquier and Rosbash, 1986; Jacquier and Michel, 1987). Both exon binding sites base pair to the intron binding sites (IBS1 and IBS2) at the 3' end of the 5' exon creating as the typical case 10 to 13 Watson-Crick pairings (overview in Michel *et al*., 1989). The important role of EBS1 could be confirmed for group IIA intron *COXI* I1 of the filamentous fungus *Podospora anserina*. A change of the spatial orientation of the EBS1 stem-loop in subdomain d of D1 by deletion of the first nucleotide of the stem leads to a strong reduction of the splicing efficiency at the 5' splicing site and a selection of different cryptic splice sites upstream in the 5' exon. The cryptic cuts are located at positions where the typical EBS1/IBS1 interaction can be restored at least partially (U. Schmidt, R. Sägebarth, and U. Stahl, unpublished data).

The so-called ε-ε' interaction is a further structurally essential D1 contact in which nucleotides 3 and 4 of the intron, in most cases a GY (ε), base-pair with two nucleotides RC (ε') of a bulge or internal loop region of subdomain c1 (see also Figure 2, Y = pyrimidine, R = purine). This interaction is also critical for recognition of the 5' splice site and is directly involved in the correct positioning of the highly conserved first intron-nucleotide (G1) to facilitate the nucleophilic attack at the 5' splice site (Jacquier and Michel, 1990). The molecular nature of the G1 nucleotide seems to be important for the required interactions in the catalytic core of group II introns. Substitution of G1 by one of the other three nucleotides C, U, or A leads to a reduction of the efficiency of both splicing steps *in vitro* and *in vivo* (Chanfreau and Jacquier, 1993; Peebles *et al*., 1993; Holländer and Kück, 1999).

The importance of the ε-ε' interaction for transesterification at the 5' splice site was emphasized by nucleotide analogue interference mapping (NAIM) of yeast intron aI5γ (Boudvillain and Pyle, 1998). Atomic changes in the sugar-phosphate backbone or in specific functional groups of single bases in both the ε and ε' region strongly impair the *trans*-branching reaction between a molecule consisting of 5' exon sequences, intron domains 1, 2, and 3 (exD123 RNA), carrying the modified nucleotides, and a domain 5 and 6 molecule (D56). The ε' region and the unpaired nucleotides of the λ region located immediately upstream (see Boudvillain *et al*., 2000 and Section VII.E for details) were also identified as a strong binding site for a divalent metal ion (like Mg^{2+}) that might contribute to stabilize the intron structure at the catalytically active site (Sigel *et al*., 2000). Using a different approach, Hertweck and Mueller (2001) identi-

fied an internal loop located upstream of the ε'/λ region as one of the most prominent metal binding sites in intron bI1.

Deletion analyses for group IIB intron bI1 from *S. cerevisiae* indicate a certain degree of redundancy between the three interactions EBS1-IBS1, EBS2-IBS2, and ε-ε' in recognition of the 5' splice site at least *in vitro*. Hetzer *et al.* (1997) found that the 5' splice site is also precisely cleaved in the absence of EBS1; however, the efficiency of the reaction is strongly decreased when compared with the wild-type intron.

Footprinting data for the group IIB intron located in the mitochondrial rRNA gene of the brown algae *Pylaiella littoralis* (Pl.LSU/2) indicate that the affinity to bind the 5' exon via the EBS-IBS interactions at D1 is strongest when the intron is completely folded and the catalytic core is correctly formed. The biological sense in preventing premature exon binding might be that it interferes with the assembly of single structural components of the intron, causing delays in or even prevents correct and complete folding into the active state (Costa and Michel, 1999). These data are supported by recent kinetic footprinting studies of a D135 ribozyme derived from intron aI5γ indicating that the formation of the catalytic core directly proceeds to a fully catalytic active intron conformation (Swisher *et al.*, 2002). Moreover, binding of a full-length exon substrate (with IBS1 and IBS2) yield a loss of binding energy that appears to be linked to a conformational rearrangement of the ribozyme (Qin and Pyle, 1999).

Apart from interactions predominantly involved in recognition of the 5' exon some other long range tertiary interactions essential for the correct folding and stability of the intron have been localized within D1. The highly conserved α-α' pairing between the terminal loop of subdomain b and an internal bulge-loop of subdomain d3 has been recognized by phylogenetic analyses (Jacquier and Michel, 1987; Michel *et al.*, 1989) and was demonstrated to be functionally important by genetic studies *in vitro* (Harris-Kerr *et al.*, 1993). In contrast, the β-β' interaction seems to be less important. It represents a pairing between two phylogenetically highly variable loop regions of subdomains c2 and d2 that when first presented could only be identified in a few predominantly self-splicing introns (Michel and Ferat, 1995). More recently, a new D1 internal base-base interaction called δ-δ' was demonstrated by comparative sequence analysis and chemical footprinting. This interaction is restricted to intron members of subgroup IIB and might facilitate base pairing between the 5' exon and the intron. The nucleotide immediately 5' of EBS1 (δ) can form a canonical pair with a nucleotide located on the 5' half of an internal loop in subdomain d (δ') (Costa *et al.*, 2000). Interestingly, in most members of the IIA subgroup the same δ nucleotide base pairs to a completely different site, the first nucleotide of the 3' exon. In subgroup IIB, however, the first 3' exon nucleotides (IBS3) pair to an intron partner (EBS3) located opposite to the δ' nucleotide on the 3' half of the same internal loop of subdomain d (Costa *et al.*, 2000) (see also Figure 2 and Section VII.G).

B. Domain 2

Domain 2 (D2) is a phylogenetically less conserved region, and particularly when comparing the two intron subgroups (IIA and IIB) a high degree of sequence (no conserved nucleotides, high length variability), and structure variation can be observed (Michel *et al.*, 1989). Thus, this domain was assigned as functionally unimportant right from the start of *in vitro* analysis of self-splicing group IIB introns.

In a *cis*-reaction the complete domain can be removed, leaving only a short hairpin structure (16 nt for intron bI1; 18 nt for intron aI5γ) as a joiner between D1 and D3 with minor effects on the autocatalytic activity. In contrast to the wild-type introns, the second splice step is now rate-limiting, as indicated by accumulation of the typical splicing intermediates (free 5' exon and intron/3' exon lariat) (Kwakman *et al.*, 1989; Bachl and Schmelzer, 1990). This observations can be explained by elimination of the so-called η-η' interaction identified later on by Chanfreau and Jacquier (1996). η-η' is an interaction of two consecutive C·G base pairs located on the basis of the second of three helices starting from an internal loop of D2 with the endloop GUAA of Domain (D6) (Figure 2). The disruption of the two C·G base pairs of the receptor helix in D2 of intron aI5γ or base substitutions in the D6 GUAA end loop, which do not correspond to the consensus, yield a strong reduction of the splicing efficiency in the second step *in vitro*, while the rate of the first splicing step is even improved (Chanfreau and Jacquier, 1996). Due to the absence of the interaction more molecules seems to fold into an active state for the 5' cut, but seems to be trapped in this conformation. Interestingly, only a

few of the known group IIB introns can potentially form this interaction. For the majority of the group IIB introns, either a different D2/D6 interaction has to be postulated or the change of the conformation between the first and the second step does not depend on this interaction at least *in vivo*. In this case, for example, protein factors (chaperons) might compensate for the missing interaction to enable the conformational change.

Four nucleotide loop regions with a GNRA consensus sequence (R = purine, N = every base) are frequently involved in the stabilization of tertiary interactions. First indications came from observations of phylogenetic co-variations in group I and group II introns, RNase P RNA as well as ribosomal RNA, in which a nucleotide substitution in the loop region was compensated by a second substitution in the putative receptor helix (Costa and Michel, 1997 and further citation therein). The existence of interactions mediated by a GNRA tetraloop was confirmed by biochemical data (for example Jaeger *et al*., 1994; Murphy and Cech, 1994; Costa and Michel, 1995; Abramovitz and Pyle, 1997), crystallographic studies (Pley *et al*., 1994; Cate *et al*., 1996) as well as *in vitro* selection of tetraloop receptor regions (Costa and Michel, 1997). On the molecular level, the contact is stabilized by a complex network of hydrogen bridges involving 2'-OH groups of the riboses on the tetraloop and receptor site. Positions 3 and 4 of the tetraloop dock in the minor groove of the RNA helix in which the receptor sequences are localized.

On the basis of phylogenetic data, another interaction between D2 and D6 similar to η-η' was postulated for a number of group IIA introns. Interestingly, when compared with group IIB the GNRA tetraloop and the receptor site have exchanged their positions in group IIA (Costa *et al*., 1997b). Many group IIA introns possess a relatively short unbranched D2 with a typical GNRA tetraloop, while the peripheral region of D6 is highly variable in size and length, with only a very few introns carrying a GNRA loop at its end. In contrast, two consecutive C·G base pairs can be frequently found in a correct receptor position for a η-η' interaction. The receptor looks different when the D2 tetraloop is a GAAA. In this case, the receptor site belongs to a family of sequences corresponding to a 11 nt motif (CCUAAG..UAUGG) that is well known to make a highly specific contact only to GAAA tetraloops (Costa and Michel, 1995).

As described in Section VI.A, group IIA introns are much less reactive during autocatalysis *in vitro* when compared to group IIB introns. The 5' splice site is predominantly cleaved by hydrolysis and splicing intermediates accumulate (linear intron/3' exon and free 5' exon) indicative of a less-efficient cut at the 3' splice site. Interestingly, the rate of 5' transesterification for group IIA introns is strongly influenced by the putative η-η' interaction. When using intron aI1 from the *COXI* gene of *S. cerevisiae* Costa *et al*. (1997b) could show that enhancement of the putative η-η' interaction almost completely blocks transesterification, while the elimination of this interaction has the reverse effect: 5' transesterification is much more efficient and hydrolytic products are only barely visible. Similar results were obtained earlier for group IIA intron I1 of the *COXI* gene of the filamentous fungus *P. anserina*, although no indications of an interaction between D2 and D6 existed at that time. An adaptation of the peripheral structure of D6 to the corresponding structure of group IIB introns or the exchange of the domain with the complete domain of intron aI5γ leads to a substantial improvement of the transesterification reaction at the 5' splice site (Schmidt *et al*., 1993). These alterations remove the η' receptor structure in D6, the 11 nt motif (CCUAAG..UAUGG), thereby disrupting the η-η' interaction with the GAAA terminal loop of D2.

Thus, what is the basic role of the η-η' interaction during group II splicing? There are a number of biochemical data pointing to the fact that both splicing steps are catalyzed by a single active site (see Chanfreau and Jacquier, 1994) that requires a structural rearrangement of the intron to catalyze the reaction of both successive substrates (5' and 3' splice site) (Chanfreau and Jacquier, 1993; Steitz and Steitz, 1993). Since both ends of the intron are positioned in close proximity to each other by covalent binding to D6 (see Section VII.F), a minor rearrangement of D6 between the first and the second splicing step should be sufficient. This rearrangement might be mediated by a contact of D6 and D2 via the η-η' interaction. The experimental data presented above indicate that *in vitro* many of the group II intron containing precursor RNAs accidentally form the η-η' interaction already prior to the first splicing step. As a consequence, the very reactive group IIB introns show a reduced transesterification rate, while the less reactive group IIA introns preferentially splices via hydrolysis. *In vivo*, however, the intron molecules could be trapped into a conformation, for example, by binding of *trans*-acting factors (RNAs, proteins) that prevent the formation of the η-η' interaction until the 5' cut has been completed.

A further well known tertiary interaction involving D2 is the so-called θ-θ' interaction between D2 and D1 (Chanfreau and Jacquier, 1996; Costa *et al*., 1997b). The θ-θ' interaction shows similar structural features as η-η'; however, in contrast to η-η', this interaction is a permanent existing contact in correctly folded introns. The presence of θ-θ' has been proven for self-splicing group II introns of both subgroups. It is an interaction of the terminal loop of the c1 stem of domain 1 belonging to the GNRA loop family with an associated receptor sequence in the basal region of domain 2 next to the central core structure (Figure 2). As for η-η', the receptor corresponds to the typical 11 nt or a similar motif in group IIA introns, while the receptor sequence is represented by the second and third C·G base pair of the basal stem in group IIB introns.

For the group IIA intron aI1 from yeast, introduction of a suboptimal θ-θ' interaction leads to a reduction of the overall splicing activity *in cis* with the strongest effect observed when substituting the GUAA loop in c1 by the noncompatible UUCG loop (Costa *et al*., 1997b). In contrast, the existence of θ-θ' in group IIB intron aI5γ was demonstrated in a bimolecular reaction in which an aI5γ Δc1 molecule reacts *in trans* with the c1 subdomain of domain 1. While binding of the c1 subdomain depends on the tetraloop/receptor combination, catalysis is barely impaired. This result shows that the primary role of the θ-θ' interaction is stabilization of the correct folding of the catalytic core of the intron (Costa *et al*., 1997b). Although a θ-θ' interaction is possible for almost all autocatalytic group II introns, there is a large number of group II introns that do not show the typical GNRA tetraloop at the end of the C1 stem. In these cases the contact between C1 and domain 2 is possibly formed by other interactions unknown to date or the active intron structure is stabilized by other *trans*-acting RNAs or RNA binding proteins.

C. Domain 3

Domain 3 (D3) is a relatively short stem-loop structure that is considered as an essential joiner between the upstream and downstream structural elements important for catalysis (D1 and D5). The complete deletion of D3 in group IIB intron aI5γ leads to strong impairment of *in vitro* catalysis *in cis* (Koch *et al*., 1992). In a *trans*-reacting multiple-turnover ribozyme constructed with the same intron, the separated D3 works as a catalytic effector that strongly improves the chemical conversion rate of the short oligonucleotide substrates (Xiang *et al*., 1998). Podar *et al*. (1995b) showed that D3 when delivered *in trans* binds firmly to other parts of the intron and, when added to constructs containing D1/D5 strongly improves the rate of splicing. In addition, as shown by UV-crosslinking studies, D3 folds into its stable active structure independently of other parts of the intron. Even after binding to other intron domains, the stable structure remains unchanged. Two regions of D3 could be identified that are important for the function of the catalytically essential D5 using modification interference assays, DMS-footprinting (DMS = dimethyl sulfate), and NAIM studies (Jestin *et al*., 1997; Boudvillain and Pyle, 1998; Konforti *et al*., 1998a).

> The principle of the modification interference technique is to analyze a radioactive labelled RNA1, which was modified in a way that statistically only one modification is present per molecule. DEPC, hydrazine or phosphorothioates are frequently used for modification. The labelled and modified RNA1 is now incubated with a RNA2, which is known to bind *in trans* to the unmodified form of RNA1. In the next step, all RNA1 molecules in which modification interferes with binding are removed from the reaction mixture selecting only molecules that were still able to bind to RNA2. Following cleavage of the selected RNA1 molecules at there modification sites and separation on a denaturing polyacrylamide-gel, the missing bands in the pattern are indicative for nucleotide positions whose modification led to the loss of binding.

The nucleotides of the phylogenetically conserved stem-loop at the base of D3 and some nucleotides of a nonconserved loop-region on the 5' site of the domain are critical for catalysis. The deletion of single bases in the conserved loop-region leads to a loss in splicing activity *in vivo* (Koll *et al*., 1987; Schmidt *et al*., 1998). In contrast, removing distal localized D3 sequences of group IIB intron bI1 that do not match the structures mentioned above has no effect on the splicing reaction *in vitro* (Bachl and Schmelzer, 1990). Keeping only the proximal stem-loop structure unchanged, while the 5' loop region is missing also results in drastically reduced autocatalytic activity.

Although there is no clear evidence yet that D3 interacts directly with D5, data of Jestin *et al*. (1997) and Swisher *et al*. (2001) imply that the stem-loop region of D3 might form particularly intimate con-

tacts with the active site (see also Section VII.E). The important structural role of D3 is also emphasized by the finding that some elements of the domain also represent strong binding sites for divalent metal ions such as Mg^{2+} (Sigel *et al*., 2000; Hertweck and Mueller, 2001).

All current molecular data indicate that D3 is packed very close within or near the catalytic active site of the intron. Thus, even if D3 is not important for the stabilization of the association of D5 and D1, it seems to play an important role for the catalytic mechanism of group II introns.

D. Domain 4

As regards to length and sequence, domain 4 (D4) is the least conserved region of all structure elements present in group II introns. Complete deletion of D4 leads to a reduced splicing activity *in vitro* (Bachl and Schmelzer, 1990; Koch *et al*., 1992); however, this can be assigned to a lower intron flexibility and an impairment of the correct spatial orientation of regions important for catalysis.

Many of the known group II introns encode a multifunctional protein, whose ORF is located completely or almost completely in D4 (see Section VIII.A), resulting in a domain of enormous size. For instance, D4 makes up more than 75% (!) of the complete intron sequence in the protein-encoding mitochondrial group IIA intron I1 of the *COXI* gene from *P. anserina* (size: 2539 nt) and aI1 of the same gene in *S. cerevisiae* (size: 2448 nt). A reduction of the size of the *Podospora* intron in D4 by more than a half or a substitution of the complete D4 by a 24 nt stem-loop leads to a strongly improved splicing reaction *in vitro* (Schmidt *et al*., 1990; Schmidt *et al*., 1993). Reduction of the size of D4 in the *S. cerevisiae* intron, a better splicer than the *Podospora* intron has no effect on the reaction efficiency *in vitro* (Hebbar *et al*., 1992).

All data indicate that D4 is not required for autocatalysis of group II introns *in vitro*, as long as a short stem-loop remains as a flexible joiner between D3 and D5.

E. Domain 5

Domain 5 (D5) is, together with D1, the only structural element that is absolutely required for the catalytic reaction. All known *cis*- and *trans*-reactions catalyzed by group II introns and group II intron derived ribozymes depend on D5 as an essential component of the active site (Michels and Pyle, 1995; Pyle, 1996; Qin and Pyle, 1998). The central catalytic role of D5 is reflected by its primary sequence and secondary structure as the phylogenetically most conserved part of group II introns (Michel *et al*., 1989). In general, D5 is a short stem-loop structure of mostly 34 nt, characterized by a 2 nt bulge at the 3' site and a GNRA endloop. Until 1998, a CG bulge was predicted at pos. 25 and 26, when an alternative AC bulge was proposed at pos. 24 and 25 by chemical probing and an expanded comparative sequence analysis (Costa *et al*., 1998; Konforti *et al*., 1998b). The AC bulge separates the proximal and the distal stem composed of 9 bp and 5 bp, respectively (Figure 4).

The central role of D5 in group II splicing has initiated numerous efforts to determine the sequence stretches, the nucleotides, and their functional groups important for catalysis and binding. Furthermore, there have also been a growing number of attempts to unravel and to understand the complete spatial conformation of this domain in the catalytic center of the intron. Most investigations are based on a detailed analysis of the effects of extensive molecular and atomic changes within D5 on the splicing reaction *in vitro*, in a few approaches also *in vivo*. More recently, these data have been confirmed and completed by crystallographic studies.

> The D5 changes analyzed include simple substitution of single nucleotides, chemical modification of bases with DEPC, DMS, hydrazine etc. and specific atomic changes in the sugar-phosphate-backbone or in single functional groups of the bases. For instance, when analyzing the sugar-phosphate-backbone, deoxynucleotides are incorporated to clarify the influence of the 2'-OH groups or the two oxygen atoms of the phosphate that are not involved in the formation of diester-bonds, are substituted by sulphur. To investigate functional groups of single bases, base-analogues like inosine, 2-aminopurine, 2,6-Diaminopurine or 7-deaza-guanine are incorporated, leading to modification of single positions in the scaffold of the bases (for examples see Chanfreau and Jacquier, 1994; Boudvillain and Pyle, 1998; Costa *et al*., 1998; Konforti *et al*., 1998b; Konforti *et al*., 1998a; Gordon and Piccirilli, 2001).

The most highly conserved sequence motif consists of an AGC triad located at pos. 2, 3, and 4 in the basal stem of the domain, of which the G is completely invariant (Michel *et al*., 1989; Michel and

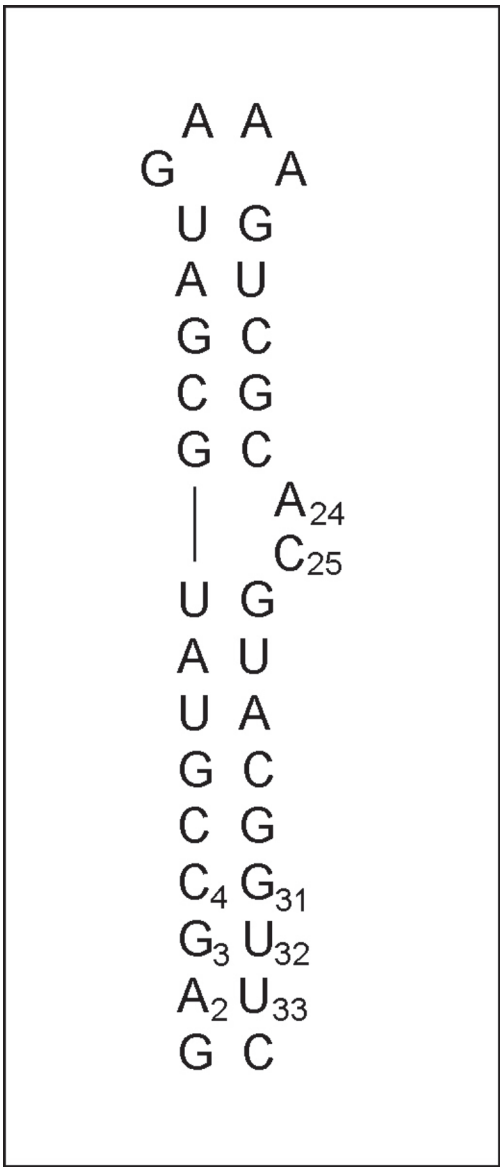

FIGURE 4. Primary sequence and postulated secondary structure of the 34 nt domain 5 of yeast mitochondrial group II intron aI5γ.

Ferat, 1995). Consistent with their high degree of conservation, the three nucleotides are very sensitive to base changes *in vivo* (Boulanger *et al*., 1995). The substitutions tested result in a complete block of the splicing reaction, while substitutions of the pairing nucleotides GUU on the opposite site (position 31-33) is tolerated (Figure 4). The Watson-Crick pairings at pos. 2 and 4 seem to be more important than the identity of the bases themselves. In contrast, the presence of the G at pos. 3 and the geometry of the G·U wobble pair is an absolute requirement for the correct chemical function of D5. In addition, when analyzing those mutations *in vitro*, substitutions at pos. G3 and C4 show the strongest impairments in activity *in cis* (Boulanger *et al*., 1995) as well as *in trans* (Peebles *et al*., 1995). These findings also agree with the data of different chemical modification analyses, showing that some of the functional groups of the three conserved nucleotides are essential for the catalysis of group II introns (Chanfreau and Jacquier, 1994; Abramovitz *et al*., 1996; Boudvillain and Pyle, 1998; Konforti *et al*., 1998b). For instance, the analysis of Konforti *et al*. (1998b) revealed that even minor atomic mutations or modifications at different exo- and endocyclic positions of the G3 nucleotide yield a dramatic loss of the *trans*-activity of D5.

A further important element for D5 function (binding and catalysis) is the 2 nt bulge at the 3' site of the domain (Figure 4). For a long time it was proposed that the conserved but not invariant nucleotides at pos. 25 and 26 (predominantly C25 or U25 and G26) form the bulge structure. However, the good accessibility of the conserved A at pos. 24 to chemical modification under native condition (Costa *et al*., 1998; Konforti *et al*., 1998b), and an expanded comparative sequence analysis of group II introns sequences has led to the conclusion that the bulge is formed by nucleotide 24 and 25. Therefore, the U9·A24 Watson-Crick pair is replaced by the U9·G26 wobble pair (Costa *et al*., 1998). This new bulge structure has been also confirmed by recent cristallographic studies (Zhang and Doudna, 2002). An extensive mutagenic analysis of the bulge region of intron aI5γ, although performed on the assumption that C25 and G26 represent the bulge, revealed that the number of nucleotides is critical for D5 function, while the exact sequence does not seem to be crucial for splicing (Schmidt *et al*., 1996). Bulges are known to induce bends in DNA and RNA helices and the angle of the axial bend was found to increase with the number of nucleotides in the bulge (reviewed by Lilley, 1995). Thus, the change of angle of the axial bend of the complete helix might be responsible for the weak binding of D5 in the bulge mutants. Alternatively, elimination or change of the number of nucleotides in the bulge might influence the structure of the proximal major groove next to the bulge. Bulge- and internal loop-regions tend to open the major groove base pairs (Weeks and Crothers, 1991; Portmann *et al*., 1996), thereby influencing the local positioning of the atomic backbone of single nucleotides in and around those structural elements. Notably, there are a number of findings showing that some 2'-OH groups, phosphates, and exocyclic amines of bases in or at the bulge have important functions for binding and catalysis (Chanfreau and Jacquier, 1994; Abramovitz *et al*., 1996; Boudvillain and Pyle, 1998). Thus, the effect of a single base mutation in the bulge region could also be assigned to the loss of the functionality of these atomic groups.

There is also growing amount of information on the potential contacts of D5 to other intron domains and on its spatial orientation in the catalytic core of group II introns. Although no Watson-Crick interactions between D5 and other intron parts could be detected, it was well known for a long time that the domain strongly binds to an exD123 molecule *in trans* (Franzen *et al*., 1993; Pyle and Green, 1994; Peebles *et al*., 1995). Thus, it was postulated early that the interaction is not mediated via classic base pairings. Costa and Michel (1995) succeeded first in proving an interaction between D5 and its primary binding site, D1. Using intron aI5γ, they were able to show by base-substitution experiments that the GAAA endloop of D5 binds to a typical 11 nt receptor motif (GCUAAG..UAUGC) in subdomain d of domain 1 (Figure 2). This so-called ζ-ζ' interaction corresponds structurally to the η-η' interaction that was postulated for some group IIA introns as a D2/D6 contact (see Section VII.B). The importance of single nucleotides of the two regions and the functionality of the ζ-ζ' interaction was confirmed by different *trans*-analyses with the same intron (Chanfreau and Jacquier, 1994; Abramovitz *et al*., 1996; Jestin *et al*., 1997; Boudvillain and Pyle, 1998; Konforti *et al*., 1998a). The first proof of the ζ-ζ' interaction in a *cis*-molecule, the spliced intron lariat of the group IIB intron Pl.LSU/2 from the mitochondrial pre-rRNA of the brown alga *Pylaiella littoralis*, succeeded applying the differential DMS modification assay (Costa *et al*., 1998).

Dimethysulfate (DMS) methylates adenine at the N1 position and, with lower efficiency, cytosine at the N3 position, provable by a footprinting-assay. Using the differential DMS modification-assay, the chemical accessibility of RNA molecules is comparatively analyzed under native and semidenaturing conditions. While adenines, important for spatial interactions, are protected against methylation under native conditions, a modification is possible under semidenaturing conditions (Costa *et al*., 1998). The RNA to be analyzed in this assay must rapidly form a homogeneous population of molecules with an active conformation, as it is characteristic for the self-splicing intron Pl.LSU/2 from *Pylaiella littoralis* (Costa *et al*., 1997a).

To identify further D5/D1 contacts, Boudvillain and Pyle (1998) developed a strategy that combines the already introduced NAIM technique (see Section VII.A) with the nucleotide analogue interference suppression technique (NAIS).

For nucleotide analogue interference suppression (NAIS) positions in the RNA molecule are chosen which show strong interference in NAIM. The introduction of an additional site-directed mutation or modification at such positions can lead to a loss of interference at positions in other RNA regions, being most likely in spatial contact with the first region (Strobel, 1999). In the case of group II intron D5 it was analyzed, to what extent single mutations or atomic changes at carefully selected sites in the exD123 molecule leads to a suppression of the nucleotide analogue interference at specific positions in the D56 molecule using a *trans*-assay.

In the κ-κ' interaction at least two nucleotides of the basal stem of D5, belonging to two consecutive C·G base pairs (C4·G31, C5·G30) contact a region of subdomain d of D1 next to the 11 nt receptor-motif of the ζ-ζ' interaction (Figure 2). The four nucleotides in D1 proven to be essential for D5 binding are unpaired nucleotides of a three-way junction representing an unusual interrupted GAAA tetraloop, in which the last A is separated from the GAA sequence (Boudvillain and Pyle, 1998; Konforti *et al*., 1998a). Although interrupted by other nucleotides or additional sequence stretches, those GN_nRA loops are well known to interact with the corresponding receptor motifs (Abramovitz and Pyle, 1997; Massire *et al*., 1998). Thus, the κ-κ' interaction represents a typical tetraloop/receptor complex, in which the two contact sites in D1 and D5 have changed position compared to the ζ-ζ' interaction.

More recently, a third D5/D1 interdomain contact was identified applying the same technique. The λ-λ' interaction is a contact of the two consecutive G·C base pairs located at the base of the distal stem next to the AC bulge (G10·C23 C11·G22) of D5 with two unpaired nucleotides (G and A) adjacent to the ε' motif, which brings the two nucleotides into close proximity to each other. Using kinetic data of *trans*-branching and hydrolytic reactions of exD123 mutants Boudvillain *et al*. (2000) conclude, that in contrast to ζ-ζ' and κ-κ' which contributes primarily to binding, the λ-λ' interaction influences D1 and D5 functionalities with specific roles in catalysis. The interaction seems to support the correct orientation of the ε-ε' nucleotides that contributes to the selection of the 5' splice site (see Section VII.A), to regions of D5 essential for catalysis.

Using the coordinates of the crystallographically determined GAAA tetraloop structure of the hammerhead ribozyme (Pley *et al*., 1994) and the data of the current chemical modification studies (Kwakman *et al*., 1990; Chanfreau and Jacquier, 1994), Pyle and co-workers (Abramovitz *et al*., 1996) developed a first three-dimensional computer model of D5 of intron aI5γ. The model was further refined adding new interference and modification data, without substantially changing the very interesting picture (Konforti *et al*., 1998b): strikingly, the functional groups important for catalysis are all clustered on one site around a major groove, defining the "chemical face" of D5, whereas groups important for binding are located at the opposite site of the domain, defining a continuous "binding face". The close spatial arrangement of functional groups of the conserved AGC triad and the AC bulge critical for reaction chemistry has been confirmed in an actual cristallographic study using crystals obtained from a modified D5/D6 construct (Zhang and Doudna, 2002). More recently, two new models describing the complex between D5 and D1 have been developed using the LSU/2 intron of *P. littoralis* (Costa *et al*., 1998) and the aI5γ intron of *S. cerevisiae* (Swisher *et al*., 2001). Although the two models differ from each other, each model provides an unique insight into the architectural organization of the active site of group II introns. As discussed by Swisher *et al*. (2001), the differences might predominantly originate from, (1) differences in the secondary structure of the two introns that makes it difficult to arrange some regions similarly, (2) creating the models from different intron subsections, and (3) differences

in the preferences and constraints used for model construction.

As an example, the three-dimensional model of the aI5γ intron core is presented in Figure 5 (see also cover of issue for a color version). The regions of the two domains used for modeling are arranged in a way that the major elements of D1 (shown in light grey, middle grey, and dark grey in the front; cover: green, blue, and yellow) almost completely surround D5 (shown in dark grey in the back; cover: red). The model combines the findings of the "D5 only" model with the D5/D1 contacts required for the first splicing step known so far. Thus, the elements of D1 contributing to binding (in middle grey; cover: green) are orientated toward the binding face of D5, including the ζ-ζ' and κ-κ' interactions, while the elements participating in the reaction chemistry (in light grey; cover: yellow) are located on the opposite site (D5 major groove), including the λ-λ', ε-ε' and EBS1/IBS1 interactions. This placement of the contacts for binding and catalysis constrains the surrounding helices in a way that the five-way junctions of D1 (joining the D1 basal helix and the helices of subdomain a, b, c, and d, see Figure 2) spans the D5 tetraloop, thereby placing the α-α' and subdomain d (d3) helices along a third interface with D5 (in dark grey; cover: blue). Notably, this model is consistent with the results of a hydroxy radical footprinting study showing a high degree of protection for most of the regions that seems to be tightly packed within the intron core (Swisher *et al*., 2001).

D5 could also be identified as a major binding site for divalent metal ions for both the chemical reaction, which is known to depend on Mg^{2+} as an essential component for both splicing steps (Steitz and Steitz, 1993; Sontheimer *et al*., 1999), and for folding of the intron core. Major sites providing ligands for metal ions are the 2 nt bulge (Sigel *et al*., 2000) and the pro-S_p oxygen of the A of the conserved AGC triad (Gordon and Piccirilli, 2001). In addition, many of the regions that are closely packed to D5 as well as regions that play other important structural roles have also been observed as metal binding sites, like, for instance, the nucleotides in the joiner region between D2 and D3 (Sigel *et al*., 2000). Two of those nucle-

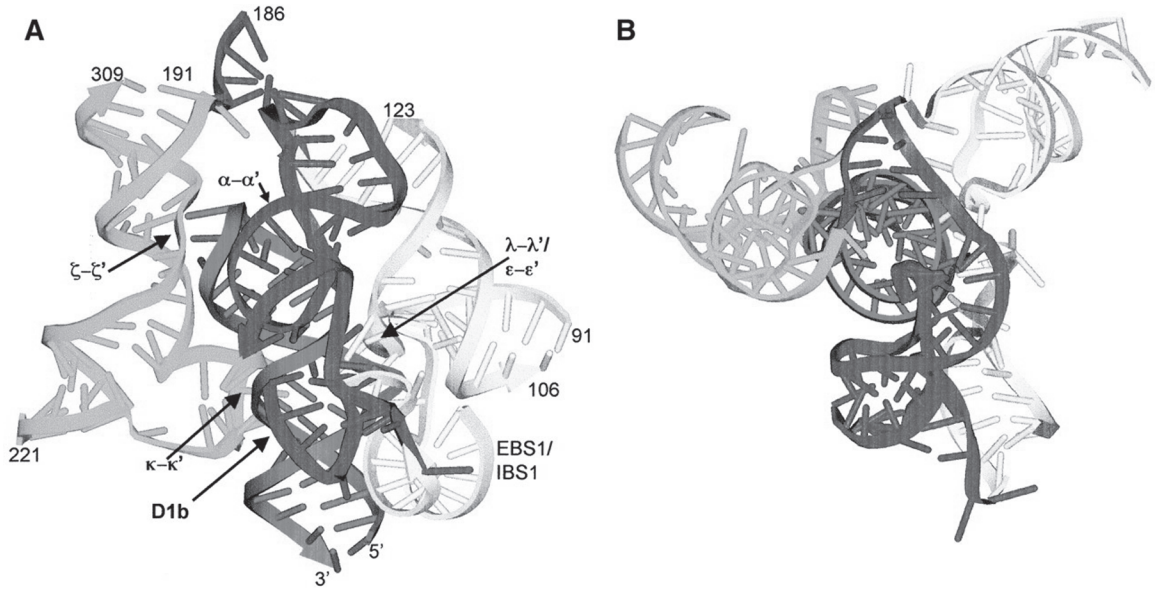

FIGURE 5. Three-dimensional model describing the complex between D5 and D1 of group II intron aI5γ. (**A**) D5 is shown in dark gray (back), D1 elements that contribute to binding are shown in middle grey (including the ζ-ζ' and κ-κ' interactions), and D1 elements that participate in the reaction chemistry are shown in light gray (including the λ-λ', ε-ε', and EBS1/IBS1 interactions). Helices of D1 that span the binding and the chemical face are shown in dark gray (front) (including the α-α' interaction). (**B**) A 90° rotation of the model around the x-axis, looking down from the top of the D5 tetraloop. (For a color version of the Figure: see the cover of this issue; the Figure is also available for download in color from the journal's website.) (Figure reprinted from: Swisher, J., Duarte, C.M., Su, L.J., and Pyle, A.M.: Visualizing the solvent-inaccessible core of a group II intron ribozyme. *EMBO J.* 20 (8): 2051–2061, with permission of Oxford University Press)

otides have been shown to form a crosslink with the uridines at pos. 32 and 33 in D5 (Podar *et al*., 1998b). A third nucleotide of the region forms the γ-γ′ interaction with the last intron nucleotide, which is important for the second splicing step (see Figure 2 and Section VII.G for more details).

F. Domain 6

The last structural domain of group II introns upstream of the 3' splice site, domain 6 (D6), forms a relatively variable hairpin structure, whose peripheral part is larger in many group IIA introns when compared with group IIB introns (see Michel *et al*., 1989). The most conserved feature of almost all functional group II introns is a stretch of several consecutive paired purine residues on the 5' site at the base of D6 predominantly composed of guanosines with a fully conserved G at pos. 4, surrounding an unpaired adenosine on the opposite site (Michel *et al*., 1989; Chu *et al*., 2001). The 2' hydroxyl group of the unpaired adenosine is used as a nucleophile for transesterification at the 5' splice site, resulting in a branched molecule, the typical intron lariat, with a 2' – 5' phosphodiester bond between the A and the first intron nucleotide (G1) (see also Section VI.A). Until recently, the branch site was always proposed as a single nucleotide bulge. However, cristallographic studies of a modified D5/D6 construct combined with the analysis of D6 mutations of intron aI5γ suggest that D6 might contain a two-nucleotide bulge composed of the adenine and the 3' localized uridine (Zhang and Doudna, 2002).

Deletion of the unpaired adenosine at the branch point completely blocks 5' transesterification and lariat formation can no longer be detected for both the two group IIB introns bI1 and aI5γ *in vitro* (Schmelzer and Mueller, 1987; van der Veen *et al*., 1987a) and for aI5γ also *in vivo* (Podar *et al*., 1998a). The reaction exclusively proceeds via 5' hydrolysis releasing linear intron molecules. The exchange of the adenosine by another unpaired nucleotide is also critical for 5' transesterification. Cytosine completely blocks 5' transesterification (Liu *et al*., 1997; Podar *et al*., 1998a), while guanosine or uracil only allows an extremely low rate of lariat formation (Gaur *et al*., 1997; Liu *et al*., 1997; Podar *et al*., 1998a). The same could be observed for intron bI1 when analyzing the A → C mutation (Schmelzer and Mueller, 1987).

Substitution of the adenosine by a number of different modified nucleosides (abasic nucleoside, naphtyl, inosine, 2-aminopurine, purine, 2,6-diaminopurine and N7-deaza adenine) revealed that the atomic structure of the base is essential for the efficiency of the branching reaction. Although two independent analyses led to different, and even contradictory, results when looking at the effects of single atomic changes, the data suggest that the adenine at this site precisely matches the active center of the intron, thereby discriminating the binding site against the functional groups of other bases (Gaur *et al*., 1997; Liu *et al*., 1997).

The local structure flanking the branch point A seems to be important to guarantee a dynamic and flexible structure required for effective and correct branching. Lariat formation is strongly reduced *in vitro* when the adenosine is trapped in a Watson-Crick pairing by adding a uracil on the opposite site (Schmelzer and Mueller, 1987; van der Veen *et al*., 1987b; Chu *et al*., 1998; Chu *et al*., 2001) and cannot be detected at all *in vivo* (Podar *et al*., 1998a). In contrast, when a guanosine is added the self-splicing efficiency is similar to the wild type (Chu *et al*., 1998, 2001) by the formation of a more flexible base pair either a G·A non-Watson-Crick pair with the branch-site nucleotide or formation of a G·U wobble pair with the proposed second bulge nucleotide. *In vivo*, the reaction is also efficient; however, accumulation of the intron/3' exon intermediate indicates a modest defect in the second splicing step (Podar *et al*., 1998a). The importance of the conformational flexibility of this region is also underscored by phylogenetic data showing a preference for a wobble or non-Watson-Crick geometry (predominantly G·U) to flank the branch point adenosine (Michel *et al*., 1989; Chu *et al*., 2001). Substitution of those base pairs by G·C base pairs leads to stronger reduction of the branching reaction as, for instance, the introduction of the flexible G·A pair at the branch point (Chu *et al*., 1998; Podar *et al*., 1998). Moreover, the crystallographic data of Zhang and Doudna (2002) suggest that the two bulge nucleotides of the branch site are stacked creating a local backbone geometry that exposes the 2' hydroxyl of the adenosine to solvent possibly facilitating its role as a nucleophile to attack the 5' splice site.

All available data point to the fact that branching occurs with high precision at the correct site or it does not occur at all. Mutants missing a functional branch point show no lariat formation and splice exclusively via the hydrolytic pathway (Schmelzer and Mueller, 1987; Liu *et al*., 1997; Chu *et al*., 1998; Chu *et al*., 2001). Chu *et al*. (2001) could show by analyzing a

large set of D6 mutants of intron aI5γ that the spatial positioning of the branch site relative to the base of the D6 stem, the G·U wobble base pair above the branch site, and the size of the joiner sequence between D5 and D6 are the major structural determinants for correct branch site selection. These findings are supported by the modification interference studies of Boudvillain and Pyle (1998), indicating that there is not a single substituent within D6 essential for the correct spatial orientation of the domain and the branch point adenosine. Obviously, the correct docking of D5 in the active center of the intron is crucial for the correct positioning of D6 as well (Dib-Hajj *et al*., 1993; Chin and Pyle, 1995). D5 and D6 are separated by a short joiner sequence of mostly 2 nt in group IIA introns and 3 or 4 nt in group IIB introns (Michel *et al*., 1989). If D5 is also responsible for the positioning of D6 changing the length of the joiner should have an effect on catalysis. Indeed, mutation analyses of the joiner of intron aI5γ indicate that the wild-type length of 3 nt is optimal. A shorter or deleted joiner results in the loss of the 5' transesterification, while a longer joiner yield at least a reduced activity (Boulanger *et al*., 1996; Chu *et al*., 2001).

Despite this rather passive role of D6, there are also data showing that the peripheral structural elements of the domain could have an influence on the splicing reaction. For instance, the loop and helix regions, which are involved in the formation of the η-η' interaction between D2 and D6 (see Section VII.B) are important for the transesterification at the 3' splice site. Moreover, there are indications that the loss or change of the stability of peripheral stem-loop structures distal from the branch point adenosine might also affect transesterification at the 5' splice site. The deletion of both peripheral stem-loops of the large D6 of the *Podospora COXI* I1 intron result in a complete loss of 5' transesterification (Schmidt *et al*., 1993), whereas mutation of the single internal loop of the short D6 of intron aI5γ to a thermodynamically very stable helix at least reduces transesterification at the 5' splicing site (Chu *et al*., 1998).

G. The Cut at the 3' Splice Site

Similar to D6, the downstream located sequences at the 3' splice site are also more likely to be aligned in a passive way into the catalytic active site. Therefore, it is not very surprising that D6 plays an important role in this process. When the complete D6 of intron aI5γ is deleted, leaving the distance between the upstream and downstream located sequences unchanged, the second splicing step is rate-limiting and cryptic 3' cuts upstream and downstream of the correct splice site are activated (Jacquier and Jacquesson-Breuleux, 1991; Koch *et al*., 1992). This observation is also supported by mutation analyses of the D5/D6 joiner sequence (Boulanger *et al*., 1996). A longer or shorter joiner affects not only transesterification at the 5' splice site, but also precision of the 3' cut and the activation of cryptic splice sites. In contrast to the alignment of the 5' splice site in the catalytic core that involves three strong classic base-pairings (EBS1-IBS1, EBS2-IBS2, ε-ε', see Section VII.A), only some weak tertiary interactions have been described, involved in the efficiency and precision of the second splicing step.

The γ-γ' interaction is a contact between a nucleotide in the D2/D3 joiner region (γ) and the terminal intron nucleotide (γ') (see Figure 2), which shows phylogenetic co-variation with the most frequent combinations A(γ)·U(γ') or G(γ)·C(γ') (Michel and Jacquier, 1987; Michel *et al*., 1989). While disruption of γ-γ' in group IIB intron aI5γ reduces the reaction efficiency of the second step of splicing, the interaction seems to play a minor role in splice site recognition (Jacquier and Michel, 1990). Similar observations were made for the *in vivo* and *in vitro* splicing reaction of the group IIB intron localized in the gene for the large rRNA of the green alga *Scenedesmus obliquus* (Holländer and Kück, 1999). Cryptic 3' splice sites are activated by the additional deletion of D6 and are located at positions that do not necessarily restore a typical γ-γ' interaction. Single cryptic splice sites are only used at higher frequencies, when a strong γ-γ' interaction at this position is restored by mutation (Jacquier and Jacquesson-Breuleux, 1991). In contrast, group IIA introns seems to be more sensitive. The substitution of the γ' nucleotide in the G(γ)·C(γ') interaction of the *COXI* I1 intron of *P. anserina* by A or U leads to a reduced 3' cut and to the choice of cryptic splice sites downstream in the 3' exon predominantly at positions where a wild-type γ-γ' interaction is restored. A possible explanation might be that *in vitro* a fraction of the catalytic active intron molecules is folded into a spatial structure that allows a precise 3' cut only when a stable γ-γ' interaction is present (Schmidt *et al*., 1993). In a more recent study, the two conserved unpaired nucleotides immediately

downstream of the γ nucleotide were also found to play an important role in the efficiency of the second step of splicing (Mikheeva *et al*., 2000). The same nucleotides were shown to be important metal binding sites (Sigel *et al*., 2000) and to crosslink to residues in the bottom helix of D5 (Podar *et al*., 1998b), indicating a positioning near the active site, at least during the second splicing step. Whether the two nucleotides interact directly with other residues near the 3' splice site remains to be to seen.

The second known base-pairing is represented by an interaction of the first nucleotide of the 3' exon (δ' in group IIA introns or IBS3 in group IIB introns) and a nucleotide within domain 1 of the intron important for guiding exon ligation. The δ' nucleotide in group IIA introns forms a well-conserved canonical base-pair with a nucleotide immediately 5' of the EBS1 sequence (δ), while analysis of group IIB introns fail to reveal any statistical evidence for this interaction (Michel *et al*., 1989; Costa *et al*., 2000). The intron partner of IBS3 in this subgroup is located within an internal loop in subdomain d (EBS3), while the nucleotide downstream of the EBS1 site pairs to a nucleotide on the opposite site of the same loop (see also Figure 2 and Section VII.A). The disruption of this guiding interaction in group IIB intron Pl.LSU/2 from *P. littoralis* yield a strong impairment of exon ligation (Costa *et al*., 2000). However, similar to the γ-γ' interaction, the guide does not appear to be important for splice site selection. The cryptic splice sites generated by disruption of the γ-γ' interaction of the *Podospora COX1* I1 are all located at positions where no strong guide interaction can be formed (Schmidt *et al*., 1993).

In summary, all results point to the fact that the correct positioning of the 3' splice site depends predominantly on D6 and its role in stabilizing the intron structure and the catalytic center for the second splicing step. Other elements like γ-γ' and δ-δ' (EBS3-IBS3) support the correct choice of the 3' splice site only when this region is located in or near to the active center of the intron.

VIII. PROTEIN DEPENDENT REACTIONS OF GROUP II INTRONS

As has already been outlined in Section VI, only some group II introns are known to splice autocatalytically from their pre-RNAs in a well-defined *in vitro* system. In general, this *in vitro* reaction depends on extreme unphysiological conditions (high salt, high temperatures, etc.) and even when optimized turnover rates are slow. For instance, a 10-min reaction time is required for 50% turnover of a pre-mRNA containing group IIB intron aI5γ under optimized *in vitro* conditions (Peebles *et al*., 1987; Jarrell *et al*., 1988b; Daniels *et al*., 1996). Interestingly, before autocatalysis of group II introns was discovered, genetic analyses clearly indicated that proteins are involved in the *in vivo* splicing process (for examples see Carignani *et al*., 1983; Schmelzer *et al*., 1983; Carignani *et al*., 1986). Thus, it is generally accepted that proteins contribute to the folding and catalysis of most, if not all group II introns. The splicing factors known are either encoded by the introns themselves (maturases) or proteins encoded by other genes of the host organism.

Most of the protein-encoding group II introns are not only catalytic active RNAs but also represent a special class of mobile genetic elements. The successful distribution as mobile elements is closely associated to their catalytic potential. The free intron-lariat molecule is able to cleave both cellular RNA as well as DNA site-specifically and to integrate into the cleavage site by reversal of the splicing reaction. In contrast to the RNA integration (see Section VI.B), complete integration into DNA depends on protein factors remaining complexed with the intron lariat in ribonucleoprotein particles (RNPs) after excision from the pre-RNA. The intron either integrates efficiently into the intronless allele of the same gene, a process called *homing*, or it *transposes* with low frequency into ectopic (non-allelic) target sites. The mobility process depends on a reverse transcriptase and on an endonuclease activity, which in general are both associated with the intron-encoded protein (IEP) (Bonen and Vogel, 2001; Belfort *et al*., 2002).

A. Group II Intron-Encoded Proteins (IEPs) and Their Multifunctionality

The typical open reading frame (ORF) of a mobile group II introns is about 2 kb in size and is invariably located within intron domain 4, outside the catalytic core (see Figure 6) (Lambowitz *et al*., 1999; Zimmerly *et al*., 2001). The long ORF sequences of group II intron have been found in mitochondria, chloroplasts, and bacteria (Lambowitz and Belfort, 1993; Michel and Ferat, 1995; Zimmerly *et al*., 2001; Dai and Zimmerly, 2002). As a consequence of the

growing number of bacterial genome projects, the first group II introns including ORF-less and ORF-containing members have also been identified recently in archaebacteria (Galagan *et al*., 2002; Dai and Zimmerly, 2003).

The intronic ORFs is either freely translated from a start codon within the intron, like the ORF of the *Lactococcus lactis* intron located in the ltrB gene, or translated in frame with the upstream exon sequences like the ORFs of the *S. cerevisiae* mitochondrial introns aI1 and aI2 in the mitochondrial *COX1* gene. Freely translated ORFs are found in mitochondrial, chloroplast, and bacterial introns, whereas the other variant is found only among introns located in mitochondria (Zimmerly *et al*., 2001; Singh *et al*., 2002). In the latter case, the primary translation product is a large, chimeric fusion protein encoded by upstream exon and intron sequences. Remarkably, however, only smaller ORF products could be identified for the yeast introns. For instance, a 96-kDa protein should be translated from exon 1 and intron aI1, but the protein identified is only 68 kDa in size (Carignani *et al*., 1983; Moran *et al*., 1994). A similar result was obtained for intron aI2 of the same gene. Here the estimated size of the protein encoded by exon 1, exon 2, and the intronic ORF was 98 kDa, and the protein found was 57 kDa (Bergantino and Carignani, 1990) and 62 kDa in size, respectively (Moran *et al*., 1994; Zimmerly *et al*., 1999). These data have led to the conclusion that the smaller proteins are the result of specific proteolytic cleavage of a pre-protein following translation. In fact, the large proteins seem to be processed by nuclear encoded mitochondrial localized proteases (van Dyck *et al*., 1995; Arlt *et al*., 1998).

In general, all group II intron ORFs are characterized by up to three major conserved domains that are related to different functions that show different levels of phylogenetical and functional conservation (Lambowitz *et al*., 1999; Zimmerly *et al*., 2001; Belfort *et al*., 2002).

The N-terminal part of the ORF (about 250 to 300 amino acids) shows significant homologies to reverse transcriptases (RTs) of retroviruses and other known retroelements (Michel and Lang, 1985). Seven blocks of short amino acid segments common to all retroelements have been identified in this N-terminal domain (no. 1 to 7 in Figure 6) (Xiong and Eickbush, 1990), which correspond to the palm and finger regions in the crystal structure of the HIV-RT (Kohlstaedt *et al*., 1992). Block 5 carries the highly conserved YXDD motif (single letter code, X = every amino acid) that is well known as a part of the active center of the RT of the HIV-1 retrovirus (Steitz *et al*., 1993). Phylogenetic analyses revealed that group II introns are closely related to a class of abundant eukaryotic retrotransposable elements missing long terminal repeats at their ends; this is a characteristic for a different class, including retroviruses. Typical members of those non-long-terminal-repeats (non-LTR) retrotransposable elements are, for instance, the LINE (L1) elements frequently found in human or the R2 elements of insects (Doolittle *et al*., 1989; Xiong and Eickbush, 1990; McClure, 1991). In a more recent comprehensive study of non-LTR elements, the number of conserved areas has been extended from seven to 11 and group II introns are also supposed to contain sequences that can be considered homologous to the four new identified regions (Malik *et al*., 1999). One of the new predicted RT regions termed region "0" localized upstream of the other regions was previously known as domain "Z" (Figure 6) (Doolittle *et al*., 1989; McClure, 1991), however, at that time not clearly considered as an RT-specific element. The first biochemical evidence of a RT activity associated with a group II IEP has been demonstrated for the mitochondrial localized introns aI1 and aI2 of *S. cerevisiae* (Kennell *et al*., 1993) and *COX1* I1 of *P. anserina* (Faßbender *et al*., 1994). RT activity was also subsequently demonstrated for ORFs of bacterial group II introns, such as the intron in the gene encoding a conjugative relaxase (*ltrB*) of *Lactococcus lactis* (Matsuura *et al*., 1997) and the *RmInt1* found to be inserted in a group of IS elements in *Sinorhizobium meliloti* (Martinez-Abarca *et al*., 1999).

Domain X downstream of the RT domain comprises a poorly conserved region of about 100 amino acids (Figure 6), which is supposed to be analogous to the thumb motif of the HIV-RT and other polymerases. In contrast to the RT and the Zn domain (see below), domain X is present in all known IEPs of group II introns, suggesting an essential function as a putative RNA binding domain required for RNA splicing (maturase function, see also Section VIII.C). This assumption is supported by the observation that conserved nucleotides of domain X of different MatK proteins and the mat-r protein encoded in organellar group II introns of higher plants are restored by RNA editing (Thompson *et al*., 1997; Vogel *et al*., 1997).

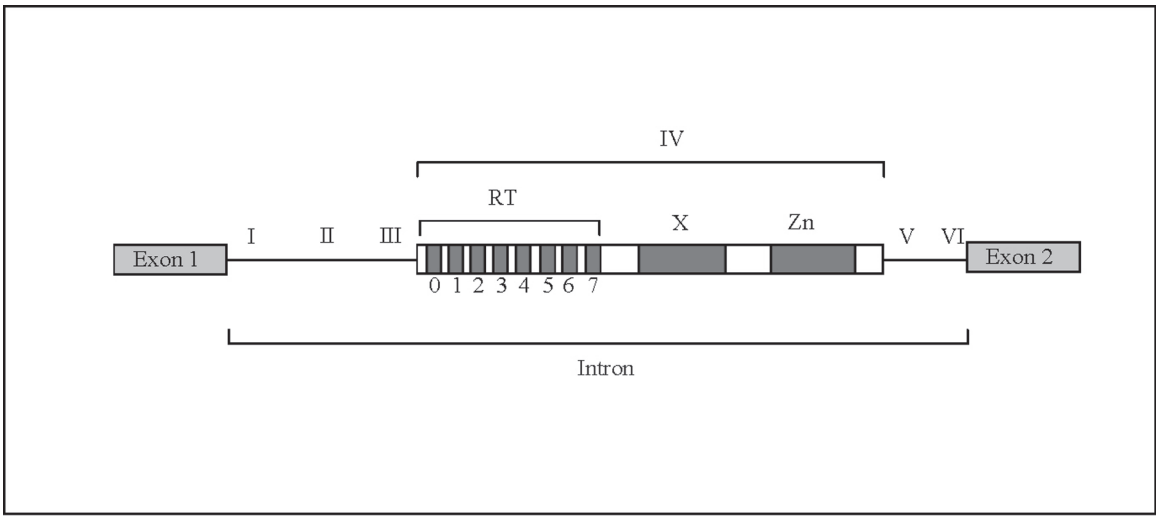

FIGURE 6. General structure and localization of a typical open reading frame (ORF) of a group II intron. The intron is drawn as a solid line flanked by the 5' and 3' exon (light gray boxes). The roman numbers indicate the six intron domains. The ORF (open box) is located within domain 4 and characterized by three different domains (dark gray boxes): a reverse transcriptase (RT), a maturase (X), and an endonuclease (Zn) domain. The RT domain contains some blocks (numbers 0 to 7) with high homology common to all retroelements.

Although it is well known that the IEP interacts directly with its corresponding intronic RNA (Saldanha *et al*., 1999; Wank *et al*., 1999), it is still a matter of speculation whether domain X is the primary RNA binding site. Most domain X sequences show a strikingly high number of basic amino acids characteristic for nucleic acids binding proteins; however, homologies to typical RNA binding motifs are only barely found (Mohr *et al*., 1993).

The C-terminal region of many group II IEPs is characterized by a short domain of 50 to 80 amino acids with two pairs of cysteine residues that fit the consensus of a major class of zinc fingers (Figure 6, Zn domain), a typical DNA binding motif (Klug and Schwabe, 1995). The same region also includes a conserved DNA endonuclease domain with significant similarities to endonucleases of bacteria (bacteriocin DNases) and of a certain type of ORF-containing group I intron. Based on the most conserved amino acids, this family of endonucleases was termed $EX_1HH\text{-}HX_3H$ (Gorbalenya, 1994) or H-N-H (Shub *et al*., 1994). Some of the group II IEPs were indeed found to express a specific DNA endonuclease activity important for intron mobility that could be assigned to the Zn domain (Zimmerly *et al*., 1995b; Guo *et al*., 1997; Singh and Lambowitz, 2001). Moreover, further detailed analyses revealed that the conserved cysteine residue pairs seem to play a critical role in maintaining the structure of the DNA endonuclease region, while a short region upstream of the Zn domain is critical for DNA-binding. The DNA-binding region contains two functionally important segments, a cluster of basic amino acids and an α-helix, conserved in related group II IEPs (San Filippo and Lambowitz, 2002).

B. The Mobility of Group II Introns

1. The Homing Process

The first genetic evidence for the mobility of ORF-containing group II introns were a result of crossing different yeast strains of *Saccharomyces cerevisiae* carrying or lacking group IIA intron aI1 and aI2 in the mitochondrial *COXI* gene. Both introns insert site-specifically into the intronless allele of the gene with almost 100% frequency (Meunier *et al*., 1990). A similar observation was made for intron aI1 of *Kluyveromyces lactis*, which sequence is 96% identical to the *S. cerevisiae* intron aI2 (Skelly *et al*., 1991). The process was called intron homing. In these early studies Meunier *et al*. (1990) already recognized that the homing process depends on the splicing reaction of the intron as well as the intron-encoded protein. Later on, ORF-containing prokaryotic group II introns like the *Lactoccocus lactis* Ll.LtrB (Mills

et al., 1997) and the *Sinorhizobium meliloti* RmInt1 (Martinez-Abarca *et al*., 2000) were shown to insert very efficiently into their homing sites.

In applying an elegant combination of genetic and biochemical analyses, many details of the complex homing process were elucidated for both the *S. cerevisiae* introns as well as for the the bacterial introns mentioned above. The central step for group II intron mobility is a mechanism known as target DNA-primed reverse transcription (TPRT) (Zimmerly *et al*., 1995b; Zimmerly *et al*., 1995a; for reviews see Curcio and Belfort, 1996; Yang *et al*., 1996; Eskes *et al*., 1997; Lambowitz *et al*., 1999; Martinez-Abarca and Toro, 2000a; Belfort *et al*., 2002). The TPRT mechanism resembles retrotransposition of the nuclear non-LTR transposons and the mechanism described for telomer formation of eukaryotic chromosomes (Zimmerly *et al*., 1995a; overview in Grivell, 1996; Eickbush, 1997; Eickbush, 1999).

Actual data define multiple pathways for group II intron homing, all of which have been shown to exist for yeast intron aI2 (see Figure 7, Eskes *et al*., 2000). The process depends on specific cleavage of the recipient target DNA by an endonuclease. The endonucleolytic activity is associated with RNP particles, in which the spliced intron lariat is complexed with the IEP. The complex of intron RNA and protein is very stable, and no additional protein components are needed for endonucleolytic and TPRT activity (Saldanha *et al*., 1999; Zimmerly *et al*., 1999). Remarkably, both the RNA and the protein components of the RNP complex contribute to the precise recognition of the DNA target site. Actual kinetic data suggest that the RNP particle binds the DNA duplex nonspecifically at any accessible site and then searches the bound DNA before undergoing a conformational change that is associated with the identification of its specific binding site (Aizawa *et al*., 2003). The lactococcal endonuclease RNP particle was shown to protect a relatively long sequence stretch on the DNA target extending from position –25 in the 5' exon to position +19 in the 3' exon on the sense (top) strand and position –28 to +16 on the antisense (bottom) strand (Singh and Lambowitz, 2001). The target sites required for efficient homing were found to comprise 31 bp from pos. –21 to +10 for yeast introns aI2 (Guo *et al*., 1997), 35 bp from pos. –26 to +9 for the *L. lactis* Ll.LtrB intron (Mohr *et al*., 2000), and 25 bp from pos. –20 to +5 for the *S. meliloti* RmInt1 (Jiménez-Zurdo *et al*., 2003).

The putative C-terminal DNA-binding region of the protein first recognizes a small number of nucleotides in the distal 5' exon region of the target site in the DNA duplex (San Filippo and Lambowitz, 2002), which is somehow opened, allowing the intron lariat to base pair to the sense strand of the DNA target (pos. –12 to +1 for intron aI2, pos. –12 to +3 for the Ll.LtrB intron and pos. –13 to +1 for the RmInt1) (Eskes *et al*., 2000; Singh and Lambowitz, 2001; Jiménez-Zurdo *et al*., 2003). The same three interactions that are important for the splicing reaction and the reverse splicing into RNA substrates (EBS1-IBS1, EBS2-IBS2 and δ-δ' or EBS3-IBS3) are also essential for binding of the intron to the DNA (Zimmerly *et al*., 1995b; Eskes *et al*., 1997; Guo *et al*., 1997; Jiménez-Zurdo *et al*., 2003).

After binding, the intron lariat cleaves the sense strand of the DNA duplex precisely at the exon target (5'/3'exon junction) and inserts into DNA by a full or a partial reverse splicing reaction (Figure 7). Meanwhile, the intron-encoded protein that interacts with the 3' exon region cleaves the antisense strand downstream of the insertion site in the 3' exon at pos. +10 in the yeast introns (Zimmerly *et al*., 1995a) and pos. +9 in the *Lactococcus* intron (Matsuura *et al*., 1997). The antisense-strand cleavage by the protein was shown to occur independent of the sense-strand cleavage by the RNA. Moreover, the first integration step of the intron lariat is highly reversible, and the trapping of the reversed-spliced product might be promoted by reactions that occur on the antisense strand of the DNA target site (Aizawa *et al.,* 2003).

The 3' end of the cleaved antisense strand serves as a DNA primer for reverse transcription (TPRT) of the partially or fully integrated intron RNA by the IEP (Figure 7, pathway a, b, and c). Strikingly, the two yeast introns aI1 and aI2 show an efficient unidirectional coconversion of upstream exon sequences independent of the chosen template (partially or fully integrated RNA, Figure 7 pathway b and c), that is, in wild-type crosses the recipient integrates the intron and flanking sequences of the 5' exon (Lazowska *et al*., 1994; Moran *et al*., 1995; Eskes *et al*., 1997). However, coconversion of flanking exon sequences is not compatible with an event, in which, after synthesis of a full-length antisense cDNA of the inserted intron RNA, the homing process is completed by a simple DNA repair mechanism. An attractive hypothesis to explain coconversion is that after initiating first strand cDNA synthesis, the retrohoming pro-

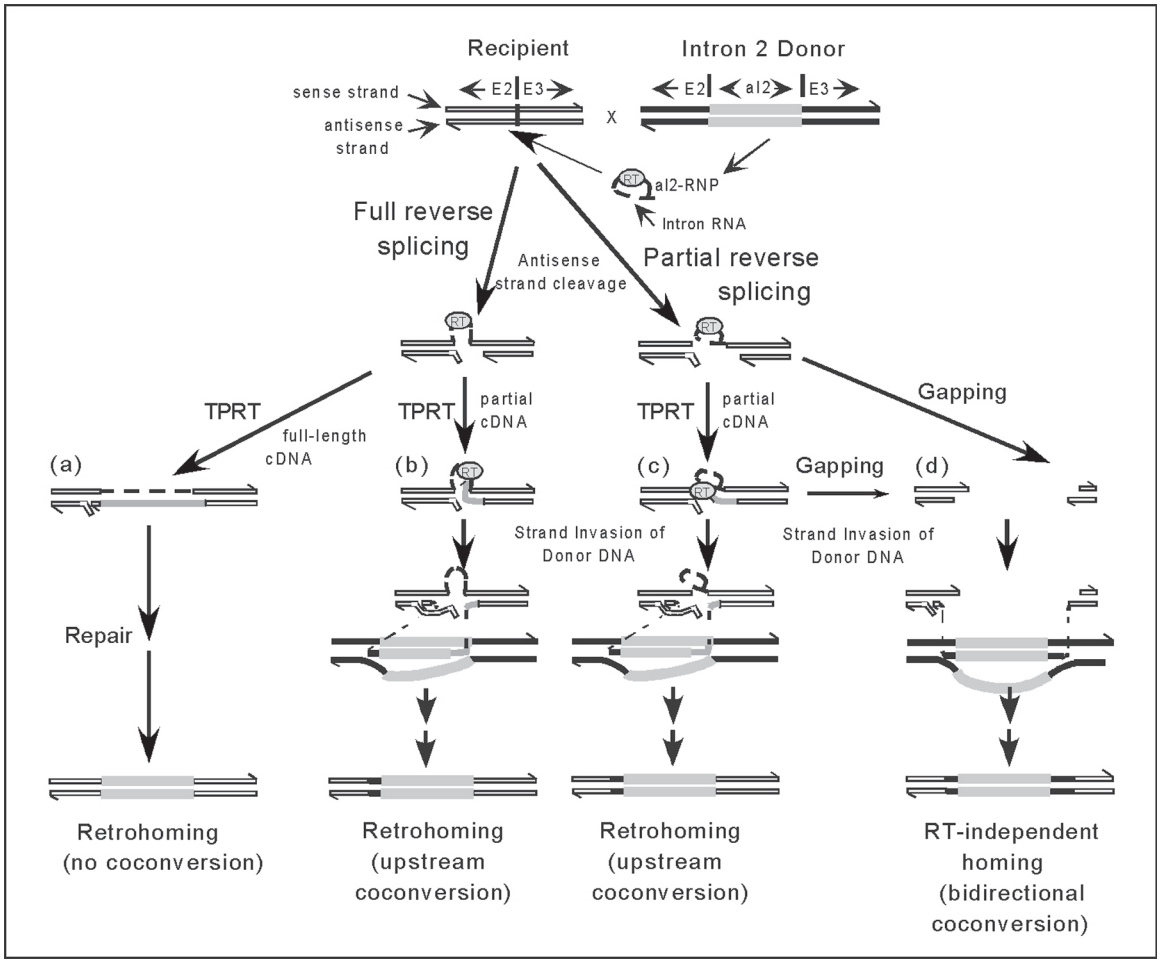

FIGURE 7. Different homing pathways described for the yeast mitochondrial group II intron al2. In the top line a donor and a recipient allele is shown. The sense and antisense strand as well as the strand polarity (half arrowheads at the 3' ends) are indicated. The exon strands (E2, E3) are shown in black (donor) or white (recipient) and the intron strands of the donor are drawn in thick grey lines. The RNP particles generated by the donor strain contain the excised intron lariat (dashed line) and the intron-encoded protein (RT). Homing is initiated by full or partial reverse splicing of the intron lariat into the sense strand of the recipient DNA target site. Following cleavage of the antisense strand by the endonuclease a cDNA is synthesized from the fully or partially integrated intron RNA. In pathway (b) and (c) the homing process is completed by a double-strand brake repair recombination leading to unidirectional coconversion of upstream exon sequences. In pathway (a) the homing process occurs via full-length cDNA synthesis and a repair mechanism without coconversion of flanking exon sequences. In pathway (d) the intron is inserted by gapping and strand invasion of the donor mtDNA in a RT-independent process that gives rise to coconversion of 5' and 3' exon sequences. (Figure reprinted from Eskes, R., Liu, L., Ma, H., Chao, M.Y., Dickson, L., Lambowitz, A.M., and Perlman, P.S. 2000. Multiple homing pathways used by yeast mitochondrial group II introns. *Mol. Cell. Biol.* **20**: 8432-8446. Copyright (2000), with permission from the American Society for Microbiology.)

cess is completed by a double-strand break repair (DSBR) recombination with the donor allele, a recombination system that is known to be highly active in yeast mitochondria. Alternatively, a template switch of cDNA synthesis from the reversed spliced intron RNA of the recipient to the pre-mRNA of the donor followed by DSBR might be responsible for the transfer of upstream exon sequences (Moran *et al*., 1995; Eskes *et al*., 1997). In contrast to the yeast introns, the major retrohoming pathway of the *Lactococcus* intron proceeds without coconversion of flanking exon sequences, suggesting that the pathway depends on synthesis of full-length cDNAs. Since efficient homing takes place in the absence of the host specific *recA* recombination system, the cDNAs seem to integrate by a RecA-independent repair mechanism (Mills *et al*., 1997; Cousineau *et al*., 1998). Efficient homing in the absence of a functional homologous recombination system has also been observed for the *RmInt1* intron of *Sinorhizobium meliloti*. However, it remains to be seen whether the *RmInt1* also uses the same homing pathway. In contrast to the *Lactococcus* and the yeast mobile introns, the protein encoded by *RmInt1* lacks the conserved part of the Zn domain essential for the endonuclease activity (Martinez-Abarca *et al*., 2000).

Surprisingly, when using a recipient strain with a mutated target site a retrohoming pathway could be strongly activated for intron aI2 (about 43% of all retrohoming events) in which insertion of the intron occurs without coconversion resembling the pathway of the *Lactococcus* intron (Figure 7, pathway a). The mutant has been shown to increase both the extent of reverse splicing and the proportion of full reverse splicing *in vitro* (Eskes *et al*., 2000).

Apart from RT-dependent homing pathways, a further homing pathway has been described for both yeast introns (Figure 7, pathway d). When using mutant donor strains lacking the intron-encoded RT activity but still having the endonuclease activity, intron mobility is only partially inhibited clearly indicating the existence of an efficient RT-independent homing (Moran *et al*., 1995). The pathway depends on the intron-encoded endonuclease activity that remains active in the RT-deficient mutants (Zimmerly *et al*., 1995b). Following cleavage of the target site, a copy of the intron is inserted by strand invasion of the donor DNA using the mitochondrial DSBR recombination system. RT-independent homing gives rise to bidirectional coconversion of upstream and downstream flanking sequences, a pattern characteristic for conventional DSBR recombination. The RT-independent pathway is one of the two main pathways for yeast intron mobility in standard crosses with wild-type donor strains showing ~ 20% (for aI1) and ~ 40% (for aI2) bidirectional coconversion among the recombinant progeny (Moran *et al*., 1995; Eskes *et al*., 1997; Eskes *et al*., 2000).

Taken together, the data presented show that mobile group II introns were able to use different homing pathways, indicating that homing of group II introns is a remarkably flexible process. Both the choice and the relative levels of the different pathways are strongly influenced by mutations of the intron-encoded protein and the nucleotides found at the DNA target site (Eskes *et al*., 2000).

The potential to use different homing pathways might also be a reason for the successful spreading of group II introns during evolution. The alternative use of different recombination systems or recombination-independent repair systems for integration facilitates adaptation to different host systems (like bacteria, fungi, algae, and plants). Moreover, the use of a pathway avoiding coconversion enables the transposition of the introns into critical ectopic sites such as functional genes of the same (see below) or other genomes.

2. The Transposition Process

Mobile group II introns are not only able to integrate efficiently into an intronless allele of the same gene, they are also able to transpose with low frequency to nonallelic (ectopic) sites of the host DNA, which resembles the normal homing sites. The evolutionary and molecular biologically important search for the corresponding mechanism was initiated by some phylogenetic data and, in particular by the following two observations:

1. The intron lariat was shown to completely or partially integrate into a foreign RNA- and DNA-substrate downstream of a short IBS1-like sequence *in vitro* (see also Section VI.C) (Mörl and Schmelzer, 1990b; Mörl *et al*., 1992).
2. Some deletion events in mitochondrial genomes of the yeasts *S. cerevisiae* and *S. pombe*, as well as in the filamentous fungus *P. anserina* exactly encompass the region from the 3' splice site of a mobile group II intron to a site showing typi-

cal IBS1- and IBS2-like sequences on the wild-type mtDNA (Ahne *et al*., 1988; Belcour *et al*., 1991; Mueller *et al*., 1993b). An attractive hypothesis to explain these deletions was to propose ectopic integration of a group II intron downstream of an IBS-like sequence followed by homologous recombination between the two identical intron copies. Recombination leads to formation of DNA subcircles and loss of parts of the mtDNA. Loss of mtDNA in the facultative anaerobic yeasts is accompanied by a respiratory-deficient phenotype (Mueller *et al*., 1993b; Schmidt *et al*., 1994), whereas for the strictly aerobic filamentous fungus the so-called premature death syndrome can be observed (Belcour *et al*., 1991).

Experiments with three different mitochondrial group IIA introns provided the first direct proof of the existence of ectopic transposition sites. Applying PCR-based analyses, six transposition sites were found for the *COX1* intron aI1 of *S. cerevisiae* and seven for the *COB* intron bI1 of *S. pombe,* with all sites localized in the corresponding original gene (Mueller *et al*., 1993b; Schmidt *et al*., 1994), while for *COX1* intron aI1 of *P. anserina* a single transposition site at a different location outside of the *COX1* gene was detected (Sellem *et al*., 1993). Some of the transposition sites correspond exactly to the mtDNA deletion sites described above. It was also subsequently shown that the bacterial group II intron Ll.LtrB of *Lactococcus lactis* and RmInt1 of *Sinorhizobium meliloti* invade different ectopic sites of the chromosome and natural occurring host plasmids, respectively (Cousineau *et al*., 2000; Martinez-Abarca and Toro, 2000b; Munoz *et al*., 2001; Ichiyanagi *et al*., 2002). Like the homing process, transposition also depends on an efficient splicing reaction and integration of the intron takes place at sites resembling the intron binding stretch for homing (IBS1, IBS2, δ'/IBS3). The frequency of conserved nucleotides is relatively high in the short stretch of IBS1- and δ'-like sequences at the integration site, while IBS2 and the 5' and 3' flanks recognized by the IEP show only a weak conservation (Yang *et al*., 1998; Cousineau *et al*., 2000; Ichiyanagi *et al*., 2002).

Now, when looking at the different pathways of intron homing (see above), what is the general pathway in which transposition occurs?

Initially, when the first ectopic integration events of introns were characterized in fungal mtDNA, an insertion of the complete reverse spliced intron into RNA was favored (Mueller *et al*., 1993b). The transcript containing the intron is reverse transcribed and the new produced chimeric cDNA is integrated into mtDNA by homologous recombination. This idea was supported by the finding that all identified transposition sites show no coconversion of flanking exon sequences. At that time, coconversion was diagnostic for all known retrohoming pathways of yeast introns and for DNA as a target. In addition, it was well known that autocatalytic group II introns are able to reverse-splice into RNA *in vitro* (Augustin *et al*., 1990; Mörl and Schmelzer, 1990b).

However, using an *in vitro* system with purified intron-specific RNP particles Yang *et al*. (1998) could show that yeast intron aI1 is able to integrate directly with low frequency in some of the ectopic insertion sites described by Mueller *et al*. (1993b), indicating a retrotransposition pathway initiated by reverse splicing of the intron RNA directly into DNA target sites.

In vivo analysis of the same intron clearly revealed that transposition to ectopic sites depends on the RT activity of the IEP. Moreover, mutations of the endonuclease domain, which is required for target DNA-primed reverse transcription (TPRT), strongly inhibits retrotransposition. Notably, the use of DNA as a target for retrotransposition was supported by the finding that ectopic insertion is not restricted to the sense strand orientation, as expected for targeting of an expressed RNA (Dickson *et al*., 2001).

The DNA targeting pathway was also proposed for the two analyzed mobile bacterial group II introns (Martinez-Abarca and Toro, 2000b; Ichiyanagi *et al*., 2002), although a frequent use of RNA as a target in a first study of the *Lactococcus* intron was suggested. This interpretation was largely caused by an experimental system that selects for ectopic insertion events arguing in favor of integration into RNA (Cousineau *et al*., 2000). As it is indicative for the use of a DNA target, mutations of the recA recombinase, which disrupt homologous DNA recombination, does not significantly reduce the frequency of retrotransposition (Martinez-Abarca and Toro, 2000b; Ichiyanagi *et al*., 2002).

In contrast to yeast intron aI1, the mutation of the endonuclease activity of the *Lactococcus* intron has little effect on retrotransposition, suggesting bottom strand cleavage of the double-stranded DNA substrate by other nonspecific host endonucleases. A different very attractive hypothesis might be that a

substantial portion of retrotransposition events is initiated by reverse splicing of the intron into single-stranded DNA during DNA replication using the 3' ends of the newly synthesized Okazaki fragments of the lagging strand as primers or a *de novo* system for reverse transcription (Dickson *et al*., 2001; Ichiyanagi *et al*., 2002). Strikingly, the *Lactococcus* intron inserts more efficiently into single-stranded DNA than double-stranded DNA substrates *in vitro*. In addition, when looking at the orientation of the *in vivo* inserted intron relative to the replication forks insertion occurs preferentially on the lagging strand template (Ichiyanagi *et al*., 2002).

Taken together, all available *in vivo* data indicate that retrotransposition into DNA is the most prominent pathway used for transposition of mobile group II introns in fungal mitochondria and bacteria (Martinez-Abarca and Toro, 2000b; Dickson *et al*., 2001; Ichiyanagi *et al*., 2002).

Since the presence of IBS-like sequences immediately upstream of a potential insertion site is, in principal, sufficient for ectopic transposition of functional mobile group II introns, the introns should also be able to integrate into the canonical 5' exon/intron junction on the genome with the original IBS1-IBS2 sequences at the 3' end of the exon. Strikingly, both mobile group II introns of *S. cerevisiae* were shown to integrate efficiently at their corresponding exon/intron junction *in vitro* (Zimmerly *et al*., 1995b; Yang *et al*., 1998). Integration at the canonical site leads to *in tandem* organized or even to multimeric intron molecules. The application of specific PCR-based strategies revealed the existence of at least duplicated group II intron molecules on the mtDNA of different fungi *in vivo* (Sägebarth *et al*., 1994; Sainsard-Chanet *et al*., 1994; Schmidt *et al*., 1994). Consequently, as described for ectopic transposition (see above), duplication or multimerization of introns should lead to homologous recombination between identical intron copies and to the formation of circular extra-genomic intron molecules. Circular extra copies with a mono-, di-, or multimeric structure for the mobile group II intron aI1 of the filamentous fungus *P. anserina* have been well known for a long time. In this unique case, the free circular intron molecules termed as plasmid-like (pl) DNA or α senDNA accumulates during growth in the mitochondria of *P. anserina* and is found in high concentrations in senescent mycelia of the fungus (Stahl *et al*., 1978; Cummings *et al*., 1979; Stahl *et al*., 1980). The mechanism of amplification of this intron is still questionable, although the most plausible hypothesis might be replication via a RT-based mechanism dependent on the IEP (Sellem *et al*., 2000).

3. Loss of Introns at the DNA Level

The RT encoded by mobile group II introns seems to be involved in a further process that leads to a precise loss of introns at the DNA level (overview in Lambowitz and Belfort, 1993). The process, also known as "DNA splicing", was discovered during genetic studies of intron mutants in the mitochondria of different fungi impeding the splicing reaction. Reversion by deletion of the corresponding intron was observed at a suprisingly high frequency that could not be explained by the random loss of intron sequences (Gargouri *et al*., 1983; Hill *et al*., 1985; Perea and Jacq, 1985; Merlos-Lange *et al*., 1987; Séraphin *et al*., 1988; Sainsard-Chanet *et al*., 1993). The deleted introns belong to group I and group II, and the loss of mutated introns is frequently accompanied by the loss of adjacent functional wild-type introns. Interestingly, deletion occurs regardless of whether the splicing defect is caused by a disruption of the intron structure itself or a mutation of either the intron-encoded maturase (see below) or another intron specific splicing factor. It was shown for the *COXI* gene of *S. cerevisiae* that all deletions of group I and group II introns are associated with the ORF-containing group IIA introns aI1 and aI2. The deletion of any other *COXI* intron is suppressed when both introns are not present (Levra-Juillet *et al*., 1989). Similar observations have been made by Schäfer *et al*. (1991) for introns localized on the mtDNA of *S. pombe*. They could show that alterations of a maturase ORF located in a group I intron were responsible for the loss of mitochondrial introns.

Current data indicate that the group II intron-encoded RTs are able to switch the RNA template and to use noncognate RNA molecules (Sellem *et al*., 2000; Morozova *et al*., 2002). Notably, as was shown for yeast intron aI2, the disruption of the intron structure by mutation leads to a loss of specificity of the wild-type RT in mitochondrial RNP particles and induces cDNA synthesis of noncognate templates through an alternative priming mechanism. The same events have been also observed at low levels for the wild-type intron (Morozova *et al*., 2002). These findings indicate that the intron-encoded RT might be

able to use different partially or completely spliced RNAs of the corresponding intron-containing gene as a template. Subsequent cDNA synthesis and integration of the newly synthesized DNA via homologous recombination should lead to loss of introns at the DNA level (Levra-Juillet *et al*., 1989).

C. The Maturase Function of the Group II IEPs

The proteins encoded by the ORF-containing group II introns play a role in both mobility and the splicing process of the introns. Genetic analyses of intron mutants of *COXI* intron aI1 of *S. cerevisiae* with a splicing defect that turned out to be *trans*-recessive in complementation assays provided first proof of a maturase function of the IEP. In crosses of these intron mutants with tester strains that carry a functional aI1 intron but were respiratory deficient by an exon mutation in the *COXI* gene, respiratory competent zygotes were produced, although no mitochondrial recombination event occurs. The simple and only explanation for the restoration of the respiratory competent phenotype is that the aI1-encoded protein of the tester enables splicing of the homologous intron of the splice-deficient mutant *in trans* (Carignani *et al*., 1983; Moran *et al*., 1994). The splicing defect of the *trans*-recessive intron mutants is the result of a nonsense mutation of the intronic ORF leading to expression of a shortened protein, whereas the RNA structure essential for splicing remains unaffected. A similar complementation analysis revealed a maturase function also for the second ORF-containing group II intron in the *COXI* gene, intron aI2 (Moran *et al*., 1994). The dependence of the *in vivo* splicing reaction from the intron-encoded LtrA protein for the bacterial group II intron Ll.LtrB of *Lactococcus* was initially demonstrated in the heterologous *E. coli* system. The deletions of a large part of the LtrA ORF within intron domain 4 or different missense mutations that do not affect the catalytic activity of the intron yield a complete block of the splicing reaction. Moreover, maturase dependence of the splicing reaction was biochemically demonstrated for the first time using an *in vitro* system with highly purified LtrA-containing RNP particles. The reaction supported by the LtrA protein is independent of ATP, and the maturase is sufficient to promote splicing under *in vitro* conditions at near physiological Mg^{2+} concentrations, where the intron cannot by itself fold into the catalytically active structure (Matsuura *et al*., 1997; Saldanha *et al*., 1999). Moreover, it could be shown that the maturase first binds to a partially folded intron via a high-affinity binding site in a subdomain of D4, an idiosyncratic structure near the beginning of its own coding region, and then makes weaker contacts to secondary binding sites in conserved regions of the catalytic core. The binding of the maturase induces an RNA conformational change, resulting in the formation of key tertiary interactions required for catalysis. Interestingly, introns deleted of the binding site in D4 show a residual maturase-dependent splicing *in vitro* and *in vivo*, indicating that the contacts to other regions are sufficient to promote splicing even in the absence of the primary binding site (Wank *et al*., 1999; Matsuura *et al*., 2001).

A further important question is which of the conserved regions of the protein are essential for splicing? As mentioned earlier (see Section VIII.A), it is assumed that the splice function is very likely associated with domain X as the only conserved part of the ORF found in all known protein-encoding group II introns (Mohr *et al*., 1993). This assumption is impressively supported by the genetic data mentioned above:

1. The aI1 nonsense mutation of the *trans*-recessive mutant analyzed by Moran *et al*. (1994) is located within domain X.
2. The size of all truncated aI1 proteins analyzed by Carignani *et al*. (1983) suggest a mutation of the complete or at least a part of domain X.
3. The missense mutations of both the yeast aI2 protein and the *Lactococcus* LtrA protein that cause a splicing defect are all localized in domain X (Moran *et al*., 1994; Matsuura *et al*., 1997).

In the case of aI2, the substitution of an invariant serine by leucine might be responsible for the loss of function of the protein.

In contrast, the protein retains RNA splicing activity when the conserved C-terminal DNA endonuclease domain required for intron mobility is deleted (Zimmerly *et al*., 1995b; Matsuura *et al*., 1997). Further truncation, which additionally deletes the variable DNA binding region between domain X and the endonuclease (Zn) domain, leaves the splicing activity unaffected (Zimmerly *et al*., 1995b). It remains to be seen to what extent the RT domain, as the second

potential RNA binding region, contributes to the formation of the catalytic active intron structure and the splicing process. For instance, a mutation of the highly conserved YADD motif (to YAHH) in block V of the RT domain eliminates RT activity but not the splicing activity of the aI2 IEP (Moran *et al*., 1995). However, a missense mutation in block IV reveals at least a partial splicing defect and a decreased stability of the spliced intron lariat (Moran *et al*., 1994).

Moreover, to date it is also difficult to obtain an exact and complete picture regarding intron specificity of the maturase function. In principle it is accepted that most IEPs recognize predominantly their own intron as a substrate. Even the proteins encoded by the yeast introns aI1 and aI2, which are characterized by a similar primary sequence show a high degree of intron specificity. The aI1 protein is able to splice the closely related aI2 intron with low efficiency only when overexpressed. This observation has been made for *cis*-dominant splicing deficient mutants of aI1 that accumulate an active aI1 maturase (Carignani *et al*., 1983). Since the aI2 maturase is not synthesized due to the blocked splicing reaction of intron aI1 (see Section VIII.A), the aI1 maturase seems to be involved in aI2 splicing. However, using a similar *cis*-dominant aI1 mutant Moran *et al*. (1994) suggested that due to barely detectable aI1 splicing activity trace amounts of the aI2 protein are produced, sufficient to weakly splice the aI2 intron.

The analysis of two *COX I* deletion mutants in which the 5' part of intron aI1 is fused to the 3' part of intron aI2 revealed that the two yeast maturases might have at least some cross-reactivity (Anziano *et al*., 1990). The splicing defect hybrid introns express a functional, chimeric protein with the N-terminal part of the aI1 protein and the C-terminal, part including domain X of the aI2 protein. The overexpression of this hybrid protein weakly restores the splicing defect of a *trans*-recessive point mutant of intron aI1 (Anziano and Butow, 1991).

The intron-specific role of the two yeast maturases is also emphasized by the observation that the splicing reaction of the two ORF-less group IIB introns (bI1 and aI5γ) on the yeast mitochondrial genome is not supported by the IEPs. Northern analysis showed that the released lariat RNA of intron aI5γ also accumulates in maturase-deficient mutants (Moran *et al*., 1994), and both group IIB introns are known to splice efficiently even when the mitochondrial protein biosynthesis is completely blocked (Hensgens *et al*., 1983).

The lactococcal LtrA maturase was also shown to function specifically in splicing the Ll.ltrB intron *in vitro*. When incubating the protein with transcripts containing the *Lactoccocus* intron or other self-splicing group II introns (yeast introns aI2 and aI5γ, *E. coli* IntB, and *Calothrix* X1) under reaction conditions where the introns are incapable of self-splicing, only splicing of the cognate intron Ll.LtrB is promoted (Saldanha *et al*., 1999).

In contrast to the yeast mitochondrial and the *Lactococcus* system, it has been suggested that other potential maturases in different plastid systems might have evolved from an intron-specific splicing factor to a general group II intron splicing factor. Potential candidates are, for instance, the MatK protein in the intron of the tnrK genes that represents the only known putative maturase in chloroplasts of higher plants, and the maturase-like proteins in the chloroplast introns of *Euglena gracilis*. Sequence comparisons indicate that the consensus of domain X of the plastid ORFs differs widely from the mitochondrial ORFs (Mohr *et al*., 1993; Zimmerly *et al*., 2001). A further indication is the extremely high number of functional group II introns compared with the small number of maturase ORFs found on the *Euglena* chloroplast genome (Zhang *et al*., 1995). Moreover, it is well known that the splicing reaction of some ORF-less group II introns localized on the plastid genomes of barley and maize depends on an active protein biosynthesis of the chloroplasts (Hess *et al*., 1994; Hübschmann *et al*., 1996; Jenkins *et al*., 1997; Vogel *et al*., 1999). Although other plastid localized proteins are certainly involved in the splicing process of those introns, the MatK maturase might play an essential role as a *trans*-acting splicing factor. There are similar considerations regarding the putative MatK maturase of the parasitic flowering plant *Epifagus virginiana*, which, in contrast to the other known MatK proteins, is encoded by a freestanding ORF. Thus, if the encoded protein is an active maturase, it has to act *in trans* on the splicing reaction of one or more of the six existing group II introns (Ems *et al*., 1995). Recently, Mohr and Lambowitz (2003) made the interesting observation that several ORFs related to group II intron-encoded proteins are present in the nuclear genomes of *Arabidopsis thaliana* and *Oryza sativa*. The proteins appear to contain a conserved maturase domain and may be transported into organelles to function in the splicing of group II introns. The *Lactococcus* maturase was found to promote

splicing by interactions with an idiosyncratic region and with conserved domains of the intron (see above). The binding to conserved intron domains might have also facilitate evolution of maturases to general group II intron splicing factors in some systems (Matsuura *et al*., 2001).

IX. DEPENDENCE OF GROUP II SPLICING ON OTHER HOST-ENCODED FACTORS

Genetic analyses of eukaryotic model organisms like the yeast *Saccharomyces cerevisiae* or the alga *Chlamydomonas reinhardtii* have revealed a number of nuclear genes that are required for the *in vivo* splicing reaction of organellar group II introns (see also Lambowitz and Perlman, 1990; Grivell, 1995; Lambowitz *et al*., 1999; Barkan and Goldschmidt-Clermont, 2000).

Many of the known nuclear-encoded factors have been described for the yeast mitochondrial splicing system. More than 15 years ago, Séraphin *et al*. (1987) screened a collection of 300 nuclear *pet* mutations belonging to about 180 different complementation groups for their effects in strains with and without the complete set of mitochondrial introns in order to estimate the number of nuclear genes that support splicing. They demonstrated that about 18 nuclear genes seem to be essential for the splicing reaction of mitochondrial introns. A similar number of yeast mitochondrial splicing factors were calculated on the basis of more than 340 nuclear genes from the *Saccharomyces* databases known to code for mitochondrial localized proteins (Grivell *et al*., 1999).

> Mutations in nuclear genes of the facultative anaerobic yeast *S. cerevisiae* that yield to a respiratory deficient phenotype and to smaller colonies when grown on glucose containing media have been termed *pet* (for nuclear petite mutations). Many *PET* gene products are directly involved in the oxidative metabolism of mitochondria, but some are also found to be important for splicing and translation of mitochondrial genes (Tzagoloff and Dieckmann, 1990).

In the last couple of years some of the nuclear genes that influence group II splicing in chloroplasts have been identified. For instance, at least 14 nuclear in *Chlamydomonas* are required for the splicing reaction of two group II introns localized in the chloroplast psaA gene (Goldschmidt-Clermont *et al*., 1990), and a few of them have been cloned and initially characterized.

Although most of the factors identified seem to support splicing indirectly, some of them are supposed to interacted directly with the intronic RNA. In general, group II intron splicing factors represent proteins with additional cellular functions, in which the splicing function might be adapted from a pre-existing protein during evolution (Lambowitz and Perlman, 1999).

A. Yeast *PET* Gene Products and Group II Splicing

For some of the *pet* mutations that cause a splicing defect of mitochondrial introns, the corresponding genes could be isolated and characterized. It has been shown that three of those genes influence the splicing reaction of group II introns.

The nuclear *MSS51* gene (MSS = mitochondrial splicing system) is required for the maturation of the intron containing pre-mRNA of the mitochondrial *COXI* gene. The nonsense mutation of the isolated *pet* mutant or the disruption of the gene led to a complete block of the splicing reaction of group II intron aI1 and aI2 and to a strong impairment of the excision of group I intron aI4 and group II intron aI5γ, while group I intron aI3 is still spliced efficiently (Simon and Faye, 1984). When the same *pet* mutant carries an intron-less *COXI* gene, a mature mRNA is produced. However, no CoxI protein can be detected, and the strain remains respiratory deficient (Decoster *et al*., 1990). Thus, it is presumed that the Mss51 protein act as a factor important for translation activation of the *COXI* mRNA, while its role in splicing seems to be secondary. As a consequence of the loss of function of *COXI* translation, the expression of the aI1 and aI2 maturase, which essentially are for splicing of the corresponding introns is also blocked (see Section VIII.A). Nevertheless, a direct effect is still possible because the excision of intron aI5γ that splices independently of the mitochondrial protein biosynthesis and the maturases is also markedly reduced in the *pet* mutant.

Similar effects have been observed for a mutation in the *MRF1* gene. The gene encodes a factor with high homology to prokaryotic translation terminators and was indeed shown to be required for the specific termination of translation of single mitochondrial proteins (MRF = mitochondrial peptide

chain release factor) (Pel *et al*., 1992b). *MRF1* was initially characterized as 1 out of 18 nuclear genes essential for expression of the *COXI* gene. The mutation of *MRF1* results in a strong splicing defect of *COXI* introns aI1 and aI2, while all other introns splices to wild-type levels (Pel *et al*., 1992a). However, the expression of an intron-less *COXI* gene copy was also strongly impaired as well as the expression of the always intron-less *COXII* gene. The defects in mRNA processing of the intron-containing *COXI* gene have been also interpreted as a consequence of the markedly reduced or missing maturase expression of the first two introns (Pel *et al*., 1992b).

The most interesting *PET* isolate affecting splicing of group II introns is the *MSS116* gene. In contrast to both the genes mentioned above, strains carrying a mutant allele of the *MSS116* gene are only respiratory deficient when the mitochondrial DNA contains a certain spectrum of group I and group II introns in the *COXI* and the *COB* gene (Séraphin *et al*., 1987). As analyzed so far, the excision of introns aI1, aI5α and/or aI5β of the *COXI* pre-mRNA, as well as the introns bI1, bI2, and/or bI3 of the *COB* pre-mRNA depends of an active Mss116 protein. The Mss116p contains all typical sequence elements of a RNA helicase of the DEAD box subfamily (Séraphin *et al*., 1989; for a review see Lüking *et al*., 1998; de la Cruz *et al*., 1999). The exact function of the Mss116p in the mitochondria of *S. cerevisiae* with regard to the great variety of biological functions of RNA helicases and missing biochemical data is still a matter of speculation. To make it more difficult, the disruption of the *MSS116* gene also influences mitochondrial functions other than splicing. Thus, the protein might play a role in efficient translation of the polycistronic *COXI-COXIII-ATP6* mRNA. Similar to Mss51 and Mrf1, the role of Mss116 in splicing could be restricted to enable synthesis of some intron-encoded maturases. However, even a direct function in splicing cannot be ruled out, since the excision of group IIB intron bI1, which is supported by Mss116 takes place independently of any mitochondrial-encoded protein (Séraphin *et al*., 1989). This idea is supported by the findings of Niemer *et al*. (1995), who showed that the overexpression of the Mss116p stimulates bI1 splicing in mitochondrial lysates when ATP is present. In accordance to the data of Séraphin *et al*. (1989), the reaction of the second maturase independent group IIB intron in the *COXI* gene, intron aI5γ, is not assisted under the same reaction conditions.

Recently, a novel DEAD box helicase has been localized in yeast mitochondria that might be involved in group II splicing. The Mrh4 protein (Mrh = mitochondrial RNA helicase) weakly suppresses *cis*-acting mutants of intron aI5γ when additional copies of the corresponding gene are present on a low-copy number plasmid. However, the disruption of the *MRH4* gene gives rise to a complete loss of mitochondrial DNA, indicating an additional important function of the encoded helicase not related to splicing (Schmidt *et al*., 2002).

B. Multicopy Suppressors of Splicing-Deficient Group II Introns

Using a strategy different from screening *pet* mutants that block splicing of mitochondrial introns, a number of other nuclear genes involved in group II splicing have been detected and characterized. The strategy is based on the overexpression of genes on multicopy plasmids to suppress splicing defects caused by a point mutation in a mitochondrial intron.

1. The Proteins of the NAM *Gene Series*

Starting with a *cis*-dominant mit$^-$ mutation in group I intron bI2 three multicopy suppressors, the genes *NAM1*, *NAM7*, and *NAM8* (NAM = nuclear accommodation of mitochondria) could be isolated from different yeast gene libraries. A first analysis of the three genes made it obvious that the spectrum of suppression is much broader and *cis*-dominant mutations of other mitochondrial introns, for instance, group IIA intron aI1, are also restored to some degree (Ben Asher *et al*., 1989).

The disruption of the *NAM1* gene in a strain with multiple introns in the mitochondrial genome led to a respiratory-deficient phenotype. The phenotype is accompanied by a drastic reduction of the *COB* mRNA, while the mature mRNA of the *COXI* gene and the transcripts of the genes *ATP6-RF3/ENS2* that are co-transcribed with *COXI* and *ATP8* can no longer be detected (Groudinsky *et al*., 1993). In addition, neither the pre-mRNAs of *COB* and *COXI* nor the spliced IIA introns aI1 and aI2 are present, whereas the spliced IIB introns aI5γ and bI1 accumulates in the mitochondria, indicating that transcription and splicing per se is not affected. Similar to the observation made for *MSS116* respiration in a delta *nam1* strain devoid of all mitochondrial introns is not re-

stored. In this strain wild-type levels of *COB* and *COXI* mRNAs are produced, but ATP6 mRNAs can, once more, not be detected. Groudinsky *et al*. (1993) concluded from these data that the Nam1p is required for correct 5'/3' end processing or stability of the intron-containing *COB* and *COXI* transcripts, as well as the *ATP6* transcript. For instance, the Nam1p might specifically interact with certain RNA stem-loop structures to block nucleolytic degradation of intron-containing pre-mRNAs, spliced introns, and the *ATP6* transcript. Another argument in favor of a direct RNA/protein interaction is provided by the localization of the protein. Nam1p is a mitochondrial matrix protein that might be involved in transportation of the mRNAs for translation that occurs in close proximity to the inner mitochondrial membrane (Wallis *et al*., 1994).

The role of the other two isolated multicopy suppressors (*NAM7* and *NAM8*) in the mitochondrial splicing machinery is even less understood and very difficult to define. The elimination of the *NAM7* function in a strain with or without mitochondrial introns leads to partial impairment of the respiratory phenotype that depends on the nonfermentable carbon source used for growth. The gene encodes a very large protein of 971 amino acids (≈109 kDa) that was initially found to contain two putative Zn fingers in the N-terminal part indicative for DNA binding and some helicase-like motifs in the central part of the protein. Suprisingly, no typical targeting sequences for the import of the Nam7p into mitochondria could be detected (Altamura *et al*., 1992), and the protein was indeed found to be associated with poly-ribosomes in the cytosol (Atkin *et al*., 1995). *NAM7* is identical to *UPF1* (UPF = up frameshift), which gene product, in a complex with other proteins, has been implicated in the modulation of efficient translation termination and in the turnover of mRNAs containing premature stop codons, a phenomenon called nonsense-mediated mRNA decay (Leeds *et al*., 1991; Leeds *et al*., 1992). A detailed biochemical analysis of the Nam7/Upf1 protein revealed a nucleic-acid-dependent ATPase and an ATP-dependent helicase activity with the potential to use DNA or RNA as a substrate (Czaplinski *et al*., 1995). Now, how can a protein clearly localized in the cytoplasm work as a mitochondrial suppressor? A possible explanation might be that the overexpressed protein is mislocated at low concentrations into the organelles and splicing is restored by direct binding and correct folding of the RNA. Notably, the Nam7/Upf1 protein is significantly similar to the yeast Sen1p, a component of a splicing complex involved in the endonucleolytic cleavage of tRNA introns in the nucleus (DeMarini *et al*., 1992). However, the cytosolic localization suggests more likely an indirect function of the protein. For instance, the Nam7/Upf1 protein might have some modulatory effects on the expression of one or more nuclear gene-encoding products that are involved in the mitochondrial splicing machinery.

The mitochondrial function of the *NAM8* encoded protein is unclear as well, because its actual site of function was also found outside the mitochondria. Remarkably, the Nam8p has been found stably associated with the U1 snRNA of the corresponding small ribonucleoprotein particle (snRNP) (Gottschalk *et al*., 1998), which is essential for recognition of the 5' splice site and the subsequent assembly of the spliceosome in the nuclear splicing process (for a review see Burge *et al*., 1999). In an independent screen the protein was also identified as an important factor for the splicing reaction of introns with nucleotides at the 5' splice site that differ from the standard consensus sequence, like the intron in the pre-mRNA of the meiosis-specific factor Mer2 (Ogawa *et al*., 1995). Later on the Nam8p was demonstrated to be generally necessary for the 5' splice site recognition under conditions that compromise U1 snRNA recruitment to the 5' splice site (weak 5' splice site, lack of pre-mRNA cap etc.) (Puig *et al*., 1999). The protein is characterized by some putative RNP domains (Ekwall *et al*., 1992; Gottschalk *et al*., 1998) that represent classic RNA binding motifs (for details see Varani and Nagai, 1998). Nevertheless, the mitochondrial function of the Nam8p remains obscure, and both direct RNA binding and indirect effects might be responsible for the multicopy suppression of mitochondrial group I and group II splicing defects (Ekwall *et al*., 1992).

2. The Proteins of the MRS Gene Series

A second group of yeast nuclear genes that were found to suppress mitochondrial splicing defects of group II introns when delivered on a high-copy-number plasmid are the three genes *MRS2*, *MRS3,* and *MRS4* (MRS= mitochondrial RNA splicing). The overexpression of these genes efficiently restored the respiratory phenotype of a strain blocked in the splicing reaction of the group IIB intron bI1 in the *COB* gene (Koll *et al*., 1987). The intron carries a point

mutation (-1 deletion in a sequence stretch of five adenosine residues) in the phylogenetically conserved proximal stem-loop of domain 3 (Schmelzer *et al*., 1983) that probably interacts with the catalytically important domain 5 (see Sections VII.C and VII.E). *Cis*-dominant point mutations in group I introns bI2, bI3, and bI4 of the COB gene are only weakly suppressed if at all (Koll *et al*., 1987).

When disrupting the *MRS2* gene, the splicing reaction of all four group II introns localized on the yeast mitochondrial genome is completely blocked, while the reaction of the group I introns present is only barely affected. Therefore, *MRS2* is the only yeast nuclear gene described so far whose product is specifically important for the splicing reaction of group II introns. However, the disruption of the gene also causes a petite phenotype in a strain devoid of all mitochondrial introns. The intronless delta *mrs2* strain shows reduced levels of cytochrome complexes, indicating an essential role of the encoded 54-kDa protein in the biogenesis of mitochondria (Wiesenberger *et al*., 1992).

Bui *et al*. (1999) could show that the Mg^{2+} concentration in mitochondria of a delta *mrs2* strain is lowered by about 50% of the wild-type level, while in strains overexpressing the *MRS2* gene the concentration is increased by about 60%. Those findings indicate that the protein plays an essential role in mitochondrial Mg^{2+} homeostasis. Moreover, the observed *in vivo* effects on group II splicing might be explained by alterations of the level of Mg^{2+} ions in mitochondria, when looking at the prominent role of Mg^{2+} ions for both correct folding and catalysis of the introns *in vitro*. In the same study, the Mrs2p was shown to be an integral protein of the mitochondrial inner membrane. The protein is anchored in the membrane by two adjacent hydrophobic transmembrane domains so that a long N-terminal part and a small C-terminal part faces the matrix space. Similar features were found for the Lpe10p, a yeast homologue whose sequence is 32% identical to the Mrs2p. However, the two proteins cannot functionally substitute each other (Gregan *et al*., 2001b). Structure analysis of both proteins revealed a weak but significant similarity to Mg^{2+} transporters of the bacterial CorA protein family (Kehres *et al*., 1998). The characteristic features are restricted to the predicted transmembrane domains (near the C-terminus) and a short sequence motif (F/Y)GMN at the end of the first domain. Notably, although the overexpressed CorA protein when targeted to the mitochondria of a delta *mrs2* or delta *lpe10* strain compensates for the absence of both the proteins by restoring wild-type levels of Mg^{2+}, while the respiratory-deficient phenotype is only partly suppressed (Bui *et al*., 1999; Gregan *et al*., 2001b). When searching the sequence databases some homologues to Mrs2p and Lpe10p could be identified in other organisms such as *Schizosaccharomyces pombe*, *Arabidopsis thaliana,* and human. Although the overall sequence identity among those proteins is rather low, the plant and human protein are also able to complement the defect of the delta *mrs2* strain comparable to the bacterial CorA protein (Schock *et al*., 2000; Zsurka *et al*., 2001). In two independent studies, several mutant alleles of the *MRS2* gene changing single amino acids have been selected that suppress different splicing defect mit$^-$ mutations located in the group IIB aI5γ and bI1 when delivered on a low copy number plasmid (Schmidt *et al*., 1998; Gregan *et al*., 2001a). Strikingly, all missense mutations were clustered in a single region of the 450 amino acids protein between amino acids 174 and 260 upstream of the membrane spanning domains. This region is supposed to form a helix-turn-helix structure that can participate in homo- or hetero-dimer formation (Schmidt *et al*., 1998). Thus, an attractive but by no means exclusive explanation for the observation that the Mrs2p and Lpe10p cannot substitute for each other might be the loss of a functional hetero-dimer structure of the two proteins. Although the level of the mutant proteins is slightly increased when compared with the wild-type protein, the available data clearly indicate that it is not the elevated protein level, but rather a more active protein that is responsible for the strong influence on splicing (Schmidt *et al*., 1998; Gregan *et al*., 2001a). Remarkably, some of the mutant proteins were shown to cause a 40% increase of the intramitochondrial Mg^{2+} concentration, suggesting that Mg^{2+} is also one of the key players in group II intron splicing *in vivo* (Gregan *et al*., 2001a). However, a more direct role of the protein cannot be excluded. The overexpressed wild-type allele restores splicing much less efficiently than the mutant alleles, although the Mg^{2+} concentration is raised to the same level (Gregan *et al*., 2001a). In addition, the functional homologues described above give rise to nearly wild-type levels of mitochondrial Mg^{2+}, while splicing of group II introns is only very weakly supported (Bui *et al*., 1999; Schock *et al*., 2000; Zsurka *et al*., 2001).

Two regions of the Mrs2p have been postulated as candidates for a potential RNA binding: the relatively hydrophilic "mutational hot spot" in the middle part described above and a strong hydrophilic domain with a central KRRRK motif in the C-terminal part of the protein (Schmidt *et al*., 1998). Although no classic RNA binding motifs are present, actual data clearly show that both parts of the protein specifically bind to different domains of the aI5γ intron (K. Lehmann and U. Schmidt, unpublished results).

In contrast to *MRS2*, it is known that the genes *MRS3* and *MRS4* are unimportant for mitochondrial biogenesis or other cellular functions as their single or double disruption does not negatively effect the phenotype of the cells (Schmidt *et al*., 1987a; Wiesenberger *et al*., 1991). *MRS3* and *MRS4* are homologues, the products of which have identical characteristics identifying them as members of the mitochondrial carrier protein family localized in the inner membrane. Although it cannot be ruled out that membrane transport and support of splicing are independent functions of the Mrs3p and Mrs4p, localization and structure indicate transport as the sole direct function of both proteins. Moreover, most parts of the proteins are enclosed by the inner-mitochondrial membrane, leaving only short stretches exposed to the matrix space where intron splicing takes place (Wiesenberger *et al*., 1991). Remarkably, the overexpression of Mrs3 and Mrs4 also complements the respiratory deficient (pet^-) phenotype caused by the delta *mrs2* mutation in strains with and without mitochondrial introns (Waldherr *et al*., 1993), and the mitochondrial Mg^{2+} concentrations were shown to be elevated to levels close to those found in the wild type (Gregan *et al*., 2001a).

Using a yeast genomic library carrying a *MRS2*, *3,* and *4* triple disruption, Waldherr *et al*. (1993) could isolate the suprisingly high number of 10 further multicopy suppressors that compensate for the loss of the respiratory competence caused by the elimination of the *MRS2* gene copy. The 10 suppressors fall into two different groups: five suppressors (*MRS5*, *MRS7*, *MRS11*, *MRS12,* and *MRS13*) are able to restore the phenotype of the delta *mrs2* strain carrying the complete set of mitochondrial introns, while five (*MRS6*, *MRS14*, *MRS15*, *MRS16,* and *MRS17*) only restore the phenotype of the intronless strain. To date, three of the splicing suppressors, *MRS5*, *MRS11,* and *MRS12*, have been analyzed in detail.

The Mrs12p shows all structural characteristics of a mitochondrial carrier protein. However, in contrast to *MRS3* and *MRS4*, disruption of the *MRS12* gene interferes with cell growth and gives rise to a complete loss of mitochondrial DNA. The gene was independently isolated as *RIM2* that suppresses a temperature-sensitive growth defect caused by the absence of a DNA helicase. The helicase is described to be involved in recombination, repair, and stability of the mtDNA. Although it is very difficult to predict the exact function of the Rim2/Mrs12 protein due to its divers genetic effects, it is supposed that the carrier protein suppresses the defects caused by the delta *mrs2* allele in a way similar to Mrs3 and Mrs4 (van Dyck *et al*., 1995).

It has been demonstrated that the gene products encoded by *MRS5* and *MRS11* are essential components of the import machinery of mitochondrial carrier proteins. Both the proteins are localized in the mitochondrial intermembrane space in a complex with the inner membrane protein Tim22 (TIM = transport inner membrane) (Koehler *et al*., 1998). Mrs5 and Mrs11 are two small, related proteins of 12 kDa with a sequence identity of about 35%, but they were not able to complement each other (Jarosch *et al*., 1996; Jarosch *et al*., 1997). In fact, it has been shown that the transport functions of the two proteins are different. The Mrs11p dissociates faster from its complex and is involved in the transport of carrier preproteins from the outer to the inner membrane, while the Mrs5p binds stronger to the complex and mediates the insertion of the carriers into the inner membrane. According to the nomenclature introduced for protein components involved in transport across the inner membrane, the two proteins were renamed as Tim10 (Mrs11) and Tim12 (Mrs5) (Koehler *et al*., 1998). The molecular events leading to high-copy-number suppression of mitochondrial splicing defects are still a matter of speculation, although the localization of the proteins definitely argues against a direct function. Overexpression might change, for instance, the membrane topology, thereby affecting the processing of RNA molecules, which is suggested to take place next to the inner mitochondrial membrane (for a review see Fox, 1996). Alternatively, splicing and thus respiratory defects might be overcome to some degree by a more general enhancement of mitochondrial processes, such as transcription and RNA processing caused by the overexpressed genes (Jarosch *et al*., 1996; Jarosch *et al*., 1997).

C. Other Nuclear-Encoded Splicing Factors of Group II Introns

In the last couple of years a number of nuclear-encoded protein factors required for the *in vivo* splicing reaction of group II introns localized on the chloroplast genome of higher land plants and algae have been identified by genetic screens, and a few of them have also been characterized further by initial biochemical studies.

Mutations in the *crs1* gene (crs = chloroplast RNA splicing) of maize specifically blocks the splicing reaction of a single intron, the group IIA intron in the plastid *atpF* gene that encodes a subunit of the ATP synthase (Jenkins *et al*., 1997; Till *et al*., 2001). The 80-kDa CRS1 protein is not related to any other known group II intron splicing factor and harbors three copies of a 10-kDa highly basic domain related to a conserved freestanding ORF of unknown function found in prokaryotic genomes. The basic characteristic of the protein as well as its localization in a chloroplast ribonucleoprotein (RNP) complex that cofractionates with the *atpF* intron RNA suggest a direct interaction of the protein with the intronic RNA. However, similar to all known group II splicing factors, the protein seems to be at least bifunctional, with additional functions in plastid translation or biogenesis of the translation machinery (Till *et al*., 2001).

In contrast to CRS1, it has been demonstrated that the product of the *crs2* gene is essential for the splicing of many group II introns in different maize chloroplast genes. Interestingly, the *crs2* mutants analyzed exclusively affect the splicing reaction of group IIB introns (Jenkins *et al*., 1997). Cloning and molecular characterization revealed that the CRS2 protein is localized in the stromal compartment of the chloroplasts that very likely also contains its target introns, suggesting a direct protein/RNA interaction. The protein was shown to be closely related to bacterial peptidyl-tRNA hydrolases (PTH) that are important to prevent the accumulation of peptidyl-tRNAs prematurely released from the ribosome. However, the CRS2 protein has some nonconservative changes of amino acid residues critical for the bacterial PTH function. CRS2 failed to complement an *E. coli* pth mutant, and therefore it is not clear whether the protein maintains PTH activity. Nevertheless, a possible scenario might be that the protein has lost its ancestral function during evolution, thereby acquiring a new function by supporting the splicing of group IIB introns. This hypothesis is supported by the following observations: (1) A second type of PTH homologue has been identified from the maize genomic sequence that may function as a bona fide PTH in chloroplasts and other cellular compartments. (2) Notably, the CRS2 protein carries a C-terminal extension not present in the bacterial PTH and the second maize protein resembling the six amino acid motif RYRYKF found as RNA binding site in the Cbp2 protein, a factor that promotes splicing of yeast mitochondrial group I introns (Jenkins and Barkan, 2001).

The *psaA* gene in the unicellular alga *Chlamydomonas reinhardtii* that encodes one of two proteins of the reaction center of photosystem I is composed of three exons scattered around the chloroplast genome. The exons are flanked by two group II introns that are removed from the different pre-mRNAs in a *trans*-splicing process in which the first intron is assembled from three and the second intron from two separated parts (see Section V for details). The splicing reaction of the two introns was demonstrated to be affected by a variety of different nuclear mutations that could be arranged in 14 different complementation groups (Goldschmidt-Clermont *et al*., 1990). The 14 corresponding genes fall into three classes: most of the genes function in the splicing reaction of either intron 1 (class C) or intron 2 (class A), while a few are involved in splicing of both the introns (class B). Two of the genes have been cloned and analyzed in detail to date.

The *Maa2* gene (Maa = maturation of psaA) encodes a class A factor required for splicing of the second *PsaA* intron. The protein has been detected in the low-density chloroplast membrane fraction that also contains RNA-binding proteins involved in translation, and it seems to be associated with a membrane component by ionic interactions. Database searches revealed significant similarities to conserved domains of pseudouridine synthases that posttranscriptionally converts uridine into pseudouridine in RNAs such as tRNA, rRNA, and snRNA molecules. Interestingly, mutations of conserved residues necessary for the synthase function did not affect the splicing activity of the protein. Thus, the two functions can be separated, although neither pseudouridine synthase activity nor a direct binding of the intron RNA could be demonstrated to date. The Maa2 protein might represent a further case of a RNA binding protein with dual function, in which the binding activity of an ancestral protein was recruited to assist group II intron splicing (Perron *et al*., 1999).

The second gene cloned is *Raa3* (Raa = RNA maturation of psaA), which codes for a class C factor involved in splicing of the first *PsaA* intron. The *Raa3* gene was found to encode a large protein of 176 kDa with no significant similarity to any other known protein, except for a small domain of about 50 amino acids of unknown function that is present in pyridoxamine 5'-phosphate oxidases. In contrast to Maa2, the protein was localized in the soluble fraction of the chloroplasts as part of a high-molecular-weight 1700-kDa complex with other unknown proteins that contains the *tscA* RNA, the middle part of the tripartic intron, and the *psaA* exon1 transcript. This chloroplast RNP particle is discussed as a kind of primitive counterpart of the eukaryotic snRNPs involved in the nuclear splicing machinery (Rivier *et al*., 2001).

Alternatively to the genetic analyses described above, Bunse *et al*. (2001) chose a biochemical strategy developed by Richard and Dujon (1997) to search for group II splicing factors in chloroplasts of *C. reinhardtii*. In this work intron-specific RNA binding proteins have been detected by comparison of the sedimentation patterns of protein-free pre-RNAs, splicing intermediates and released introns to the pattern of the same molecules in native cellular extracts. The mitochondrial group II intron rI1 of the green alga *Scenedesmus obliquus* integrated into the chloroplast *tscA* gene of *C. reinhardtii* was used as a target to screen for general group II splicing factors. It is well known that the mitochondrial *Scenedesmus* intron splices efficiently from the *Chlamydomonas* chloroplast gene. Applying this method a 31-kDa and a 61-kDa protein could be identified that specifically bind to a rI1 intron-transcript containing domain 4, 5, and 6 (Bunse *et al*., 2001). More work still needs to be done to further characterize these proteins.

X. SUMMARY AND FUTURE PROSPECTS

This review gives an actual and comprehensive survey of all aspects of the structure/function relationships and the catalytic potential of group II introns.

The primary role of group II introns is to mediate splicing of transcript from the genes in which they are located. The splicing process is characterized by two consecutive transesterification reactions, leading to the precise excision of the intron in a typical lariat form and the correct ligation of the upstream and downstream exon.

The free intron lariat can act as a true ribozyme able to perform a variety of different chemical transformation reactions on foreign RNA substrates *in vitro*. Moreover, a "multiple turnover" ribozyme has been engineered that cleaves both single-stranded RNA and DNA substrates, a feature not known for any other catalytic RNA.

The size and complexity of group II introns as well as conformational changes during splicing makes it very difficult to determine their overall catalytic structure. Nevertheless, a lot of progress has been made over the last few years that provides first insights into the three-dimensional architecture of the catalytic center of group II intron ribozymes.

Some group II introns that contain an open reading frame (ORF) have been shown to act as mobile genetic elements. The mobility process is initiated by a ribonucleoprotein (RNP) complex consisting of the multifunctional intron-encoded protein (endonuclease, reverse transcriptase) and the excised lariat, leading to integration into double-stranded DNA at the homologous sites (homing) or with lower frequency at ectopic sites (transposition).

The finding that the DNA target site is recognized primarily by base pairing of a short sequence (14–16 nt) of the intron RNA gave rise to the idea that mobile group II introns can be retargeted to insert into virtually any desired DNA. Thus, first targeted group II introns ("targetrons") have been developed demonstrating their enormous potential in genetic engineering, functional genomics, and human gene therapy (Guo *et al*., 2000; Karberg *et al*., 2001; Frazier *et al*., 2003; Zhong *et al*., 2003). However, when thinking about applications in gene therapy, it should be kept in mind that even in well-tested systems with high specificity of target DNA insertion, the risk for mistargeting is always present and should not be underestimated, as has been done before when applying retroviral gene therapy vectors.

The *in vivo* splicing reaction of many, if not of all, group II introns depends on proteins either encoded by the introns themselves (maturases) or by other host-encoded factors. While maturases are well known to promote splicing directly by binding to the intronic RNA, there are only preliminary data available on the specific binding of other proteins. Thus, an important aspect of future work will be to elucidate how host-encoded proteins affect the catalysis of group II introns. These data might be important for the following two reasons:

1. Almost all of the known splicing factors have or had other primary cellular functions, and the splice function appears to be adapted later on in evolution. How has the function in catalysis developed from the original one?
2. Ribozymes engineered from group II introns are able to cleave or to integrate into foreign RNA or DNA substrates. In the long run, data on proteins that interact with group II introns might lead to the development of protein-regulated RNA catalytic systems.

ACKNOWLEDGMENTS

The authors thank Oliver Goldenberg for help with the first version of the manuscript and Roslin Bensmann for carefully checking the English.

REFERENCES

Abelson, J., Trotta, C.R., and Li, H. 1998. tRNA splicing. *J. Biol. Chem.* **273:** 12685–12688.

Abramovitz, D.L. and Pyle, A.M. 1997. Remarkable morphological variability of a common RNA folding motif: the GNRA tetraloop-receptor interaction. *J. Mol. Biol.* **266:** 493–506.

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.

Ahne, A., Müller-Derlich, J., Merlos-Lange, A.M., Kanbay, F., Wolf, K., and Lang, B.F. 1988. Two distinct mechanisms for deletion in mitochondrial DNA of *Schizosaccharomyces pombe* mutator strains. Slipped mispairing mediated by direct repeats and erroneous intron splicing. *J. Mol. Biol.* **202:** 725–734.

Aizawa, Y., Xiang, Q., Lambowitz, A.M., and Pyle, A.M. 2003. The pathway of DNA recognition and RNA integration by a group II intron retrotransposon. *Mol. Cell* **11:** 795–805.

Altamura, N., Groudinsky, O., Dujardin, G., and Slonimski, P.P. 1992. NAM7 nuclear gene encodes a novel member of a family of helicases with a Zn-ligand motif and is involved in mitochondrial functions in *Saccharomyces cerevisiae*. *J. Mol. Biol.* **224:** 575–587.

Altmann, S. and Kirsebom, L. 1999. Ribonuclease P. **In:** *The RNA World, 2nd edition*. pp. 351–380. Gesteland, R.F., Cech, T.R., and Atkins, J.F., Eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Anziano, P.Q. and Butow, R.A. 1991. Splicing-defective mutants of the yeast mitochondrial *COXI* gene can be corrected by transformation with a hybrid maturase gene. *Proc. Natl. Acad. Sci. USA* **88:** 5592–5596.

Anziano, P.Q., Moran, J.V., Gerber, D., and Perlman, P.S. 1990. Novel hybrid maturases in unstable pseudorevertants of maturaseless mutants of yeast mitochondrial DNA. *Nucleic Acids Res.* **18:** 3233–3239.

Arlt, H., Steglich, G., Perryman, R., Guiard, B., Neupert, W., and Langer, T. 1998. The formation of respiratory chain complexes in mitochondria is under the proteolytic control of the m-AAA protease. *EMBO J.* **17:** 4837–4847.

Arnberg, A.C., Van Ommen, G.J., Grivell, L.A., Van Bruggen, E.F., and Borst, P. 1980. Some yeast mitochondrial RNAs are circular. *Cell* **19:** 313–319.

Atkin, A.L., Altamura, N., Leeds, P., and Culbertson, M.R. 1995. The majority of yeast UPF1 co-localizes with polyribosomes in the cytoplasm. *Mol. Biol. Cell.* **6:** 611–625.

Augustin, S., Mueller, M.W., and Schweyen, R.J. 1990. Reverse self-splicing of group II intron RNAs *in vitro*. *Nature* **343:** 383–386.

Bachl, J. and Schmelzer, C. 1990. Effect of deletions at structural domains of group II intron bI1 on self-splicing *in vitro*. *J. Mol. Biol.* **212:** 113–125.

Barkan, A. and Goldschmidt-Clermont, M. 2000. Participation of nuclear genes in chloroplast gene expression. *Biochimie* **82:** 559–572.

Bar-Shalom, A. and Moore, M.J. 2000. Tri-partite assay for studying exon ligation by the ai5γ group II intron. *Biochemistry* **39:** 10207–10218.

Been, M.D. and Wickham, G.S. 1997. Self-cleaving ribozymes of hepatitis delta virus RNA. *Eur. J. Biochem.* **247:** 741–753.

Belcour, L., Begel, O., and Picard, M. 1991. A site-specific deletion in mitochondrial DNA of *Podospora* is under the control of nuclear genes. *Proc. Natl. Acad. Sci. USA* **88:** 3579–3583.

Belfort, M., Derbyshire, V., Parker, M.M., Cousineau, B., and Lambowitz, A.M. 2002. Mobile introns: pathways and proteins. **In:** *Mobile DNA II*. pp. 761–783. Craig, N.L., Craigie, R., Gellert, M., and Lambowitz, A.M., Eds., ASM Press Publishers, Washington, DC.

Ben Asher, E., Groudinsky, O., Dujardin, G., Altamura, N., Kermorgant, M., and Slonimski, P.P. 1989. Novel class of nuclear genes involved in both mRNA splicing and protein synthesis in *Saccharomyces cerevisiae* mitochondria. *Mol. Gen. Genet.* **215:** 517–528.

Bergantino, E. and Carignani, G. 1990. Antibodies against a fused gene product identify the protein

encoded by a group II yeast mitochondrial intron. *Mol. Gen. Genet.* **223:** 249–257.

Birikh, K.R., Heaton, P.A., and Eckstein, F. 1997. The structure, function and application of the hammerhead ribozyme. *Eur. J. Biochem.* **245:** 1–16.

Bonen, L. 1993. *Trans*-splicing of pre-mRNA in plants, animals, and protists. *FASEB J.* **7:** 40–46.

Bonen, L. and Vogel, J. 2001. The ins and outs of group II introns. *Trends Genet.* **17:** 322–331.

Bonitz, S.G., Coruzzi, G., Thalenfeld, B.E., Tzagoloff, A., and Macino, G. 1980. Assembly of the mitochondrial membrane system. Structure and nucleotide sequence of the gene coding for subunit 1 of yeast cytochrome oxidase. *J. Biol. Chem.* **255:** 11927–11941.

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 tertiary interaction that links active-site domains to the 5' splice site of a group II intron. *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.

Boulanger, S.C., Faix, P.H., Yang, H., Zhuo, J., Franzen, J.S., Peebles, C.L., and Perlman, P.S. 1996. Length changes in the joining segment between domains 5 and 6 of a group II intron inhibit self-splicing and alter 3' splice site selection. *Mol. Cell. Biol.* **16:** 5896–5904.

Bui, D.M., Gregan, J., Jarosch, E., Ragnini, A., and Schweyen, R.J. 1999. The bacterial magnesium transporter CorA can functionally substitute for its putative homologue Mrs2p in the yeast inner mitochondrial membrane. *J. Biol. Chem.* **274:** 20438–20443.

Bunse, A.A., Nickelsen, J., and Kück, U. 2001. Intron-specific RNA binding proteins in the chloroplast of the green alga *Chlamydomonas reinhardtii*. *Biochim. Biophys. Acta* **1519:** 46–54.

Burge, C.B., Tuschl, T., and Sharp, P.A. 1999. Splicing of precursors to mRNAs by the spliceosomes. **In:** *The RNA World, 2nd edition*. pp. 525–560. Gesteland, R.F., Cech, T.R., and Atkins, J.F., Eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Butow, R.A. and Perlman, P.S. 1991. Introns in pieces. *Curr. Biol.* **1:** 331–333.

Carignani, G., Netter, P., Bergantino, E., and Robineau, S. 1986. Expression of the mitochondrial split gene coding for cytochrome oxidase subunit I in *S. cerevisiae*: RNA splicing pathway. *Curr. Genet.* **11:** 55–63.

Carignani, G., Groudinsky, O., Frezza, D., Schiavon, E., Bergantino, E., and Slonimski, P.P. 1983. An mRNA maturase is encoded by the first intron of the mitochondrial gene for the subunit I of cytochrome oxidase in *S. cerevisiae*. *Cell* **35:** 733–742.

Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Kundrot, C.E., Cech, T.R., and Doudna, J.A. 1996. Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science* **273:** 1678–1685.

Cech, T.R. 1990. Self-splicing of group I introns. *Annu. Rev. Biochem.* **59:** 543–568.

Cech, T.R. 1993. Structure and mechanism of the large catalytic RNAs: group I and group II introns and ribonuclease P. **In:** *The RNA World.* pp. 239–269. Gesteland, R.F. and Atkins, J.F., Eds., Cold Spring Harbor Press.

Chanfreau, G. and Jacquier, A. 1993. Interaction of intronic boundaries is required for the second splicing step efficiency of a group II intron. *EMBO J.* **12:** 5173–5180.

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.

Chen, C.Y. and Sarnow, P. 1995. Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science* **268:** 415–417.

Chin, K. and Pyle, A.M. 1995. Branch-point attack in group II introns is a highly reversible transesterification, providing a potential proofreading mechanism for 5'-splice site selection. *RNA* **1:** 391–406.

Christopher, D.A. and Hallick, R.B. 1989. *Euglena gracilis* chloroplast ribosomal protein operon: a new chloroplast gene for ribosomal protein L5 and description of a novel organelle intron category designated group III. *Nucleic Acids Res.* **17:** 7591–7608.

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.M. 2001. Control of branch-site choice by a group II intron. *EMBO J.* **20:** 6866–6876.

Collins, R. A. 2002. The *Neurospora* Varkud satellite ribozyme. *Biochem. Soc. Trans.* **30:** 1122–1126.

Copertino, D.W. and Hallick, R.B. 1993. Group II and group III introns of twintrons: potential relationship with nuclear pre-mRNA introns. *Trends Biochem. Sci.* **18:** 467–471.

Costa, M. and Michel, F. 1995. Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J.* **14:** 1276–1285.

Costa, M. and Michel, F. 1997. Rules for RNA recognition of GNRA tetraloops deduced by *in vitro* selection: comparison with *in vivo* evolution. *EMBO J.* **16:** 3289–3302.

Costa, M. and Michel, F. 1999. Tight binding of the 5' exon to domain I of a group II self-splicing intron requires completion of the intron active site. *EMBO J.* **18:** 1025–1037.

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., Fontaine, J.M., Loiseaux-de Goer, S., and Michel, F. 1997a. A group II self-splicing intron from the brown alga *Pylaiella littoralis* is active at unusually low magnesium concentrations and forms populations of molecules with a uniform conformation. *J. Mol. Biol.* **274:** 353–364.

Costa, M., Deme, E., Jacquier, A., and Michel, F. 1997b. Multiple tertiary interactions involving domain II of group II self-splicing introns. *J. Mol. Biol.* **267:** 520–536.

Cousineau, B., Lawrence, S., Smith, D., and Belfort, M. 2000. Retrotransposition of a bacterial group II intron. *Nature* **404:** 1018–1021.

Cousineau, B., Smith, D., Lawrence-Cavanagh, S., Mueller, J.E., Yang, J., Mills, D., Manias, D., Dunny, G.M., Lambowitz, A.M., and Belfort, M. 1998. Retrohoming of a bacterial group II intron: mobility via complete reverse splicing, independent of homologous DNA recombination. *Cell* **94:** 451–462.

Cummings, D.J., Belcour, L., and Grandchamp, C. 1979. Mitochondrial DNA from *Podospora anserina*. II. Properties of mutant DNA and multimeric circular DNA from senescent cultures. *Mol. Gen. Genet.* **171:** 239–250.

Curcio, M.J. and Belfort, M. 1996. Retrohoming: cDNA-mediated mobility of group II introns requires a catalytic RNA. *Cell* **84:** 9–12.

Czaplinski, K., Weng, Y., Hagan, K.W., and Peltz, S.W. 1995. Purification and characterization of the Upf1 protein: a factor involved in translation and mRNA degradation. *RNA* **1:** 610–623.

Dai, L. and Zimmerly, S. 2002. Compilation and analysis of group II intron insertions in bacterial genomes: evidence for retroelement behavior. *Nucleic Acids Res.* **30:** 1091–1102.

Dai, L. and Zimmerly, S. 2003. ORF-less and reverse-transcriptase-encoding group II introns in archaebacteria, with a pattern of homing into related group II intron ORFs. *RNA* **9:** 14–19.

Daniels, D.L., Michels, W.J., Jr., 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 la Cruz, J., Kressler, D., and Linder, P. 1999. Unwinding RNA in *Saccharomyces cerevisiae*: DEAD-box proteins and related families. *Trend. Biochem. Sci.* **24:** 192–198.

Decoster, E., Simon, M., Hatat, D., and Faye, G. 1990. The *MSS51* gene product is required for the translation of the *COX1* mRNA in yeast mitochondria. *Mol. Gen. Genet.* **224:** 111–118.

DeMarini, D.J., Winey, M., Ursic, D., Webb, F., and Culbertson, M.R. 1992. *SEN1*, a positive effector of tRNA-splicing endonuclease in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **12:** 2154–2164.

Dib-Hajj, S.D., Boulanger, S.C., Hebbar, S.K., Peebles, C.L., Franzen, J.S., and Perlman, P.S. 1993. Domain 5 interacts with domain 6 and influences the second transesterification reaction of group II intron self-splicing. *Nucleic Acids Res.* **21:** 1797–1804.

Dickson, L., Huang, H.R., Liu, L., Matsuura, M., Lambowitz, A.M., and Perlman, P.S. 2001. Retrotransposition of a yeast group II intron occurs by reverse splicing directly into ectopic DNA sites. *Proc. Natl. Acad. Sci. USA* **98:** 13207–13212.

Doetsch, N.A., Thompson, M.D., and Hallick, R.B. 1998. A maturase-encoding group III twintron is conserved in deeply rooted euglenoid species: are group III introns the chicken or the egg? *Mol. Biol. Evol.* **15:** 76–86.

Doetsch, N.A., Thompson, M.D., Favreau, M.R., and Hallick, R.B. 2001. Comparison of psbK operon organization and group III intron content in chloroplast genomes of 12 Euglenoid species. *Mol. Gen. Genet.* **264:** 682–690.

Doolittle, R.F., Feng, D.F., Johnson, M.S., and McClure, M.A. 1989. Origins and evolutionary relationships of retroviruses. *Q. Rev. Biol.* **64:** 1–30.

Doudna, J.A. and Cech, T.R. 2002. The chemical repertoire of natural ribozymes. *Nature* **418:** 222–228.

Eickbush, T.H. 1997. Telomerase and retrotransposons: which came first? *Science* **277:** 911–912.

Eickbush, T.H. 1999. Mobile introns: retrohoming by complete reverse splicing. *Curr. Biol.* **9:** R11–14.

Ekwall, K., Kermorgant, M., Dujardin, G., Groudinsky, O., and Slonimski, P.P. 1992. The *NAM8* gene in *Saccharomyces cerevisiae* encodes a protein with putative RNA binding motifs and acts as a suppressor of mitochondrial splicing deficiencies when overexpressed. *Mol. Gen. Genet.* **233:** 136–144.

Ems, S.C., Morden, C.W., Dixon, C.K., Wolfe, K.H., dePamphilis, C.W., and Palmer, J.D. 1995. Transcription, splicing and editing of plastid RNAs in the nonphotosynthetic plant *Epifagus virginiana*. *Plant Mol. Biol.* **29:** 721–733.

Eskes, R., Yang, J., Lambowitz, A.M., and Perlman, P.S. 1997. Mobility of yeast mitochondrial group II introns: engineering a new site specificity and retrohoming via full reverse splicing. *Cell* **88:** 865–874.

Eskes, R., Liu, L., Ma, H., Chao, M.Y., Dickson, L., Lambowitz, A.M., and Perlman, P.S. 2000. Multiple homing pathways used by yeast mitochondrial group II introns. *Mol. Cell. Biol.* **20:** 8432–8446.

Faßbender, S., Brühl, K.H., Ciriacy, M., and Kück, U. 1994. Reverse transcriptase activity of an intron encoded polypeptide. *EMBO J.* **13:** 2075–2083.

Fedor, M.J. 2000. Structure and function of the hairpin ribozyme. *J. Mol. Biol.* **297:** 269–291.

Feig, A.L. and Uhlenbeck, O.C. 1999. The role of metal ions in RNA biochemistry. **In:** *The RNA World, 2nd edition*. pp. 287–319. Gesteland, R.F., Cech, T.R., and Atkins, J.F., Eds., Cold Spring Harbor Laboratory Press.

Ferat, J.-L. and Michel, F. 1993. Group II self-splicing introns in bacteria. *Nature* **364:** 358–361.

Ferat, J.-L., Le Gouar, M., and Michel, F. 1994. Multiple group II self-splicing introns in mobile DNA from *Escherichia coli*. *C. R. Acad. Sci. III, Sci. Vie* **317:** 141–148.

Fox, T.D. 1996. Translational control of endogenous and recoded nuclear genes in yeast mitochondria: regulation and membrane targeting. *Experientia* **52:** 1130–1135.

Frank, D.N. and Pace, N.R. 1998. Ribonuclease P: unity and diversity in a tRNA processing ribozyme. *Annu. Rev. Biochem.* **67:** 153–180.

Franzen, J.S., Zhang, M., and Peebles, C.L. 1993. Kinetic analysis of the 5' splice junction hydrolysis of a group II intron promoted by domain 5. *Nucleic Acids Res.* **21:** 627–634.

Frazier, C.L., San Filippo, J., Lambowitz, A.M., and Mills, D.A. 2003. Genetic manipulation of *Lactococcus lactis* by using targeted group II introns: generation of stable insertions without selection. *Appl. Environ. Microbiol.* **69:** 1121–1128.

Galagan, J.E., Nusbaum, C., Roy, A., Endrizzi, M.G., Macdonald, P., FitzHugh, W., Calvo, S., Engels, R., Smirnov, S., Atnoor, D., Brown, A., Allen, N., Naylor, J., Stange-Thomann, N., DeArellano, K., Johnson, R., Linton, L., McEwan, P., McKernan, K., Talamas, J., Tirrell, A., Ye, W., Zimmer, A., Barber, R.D., Cann, I., Graham, D.E., Grahame, D.A., Guss, A.M., Hedderich, R., Ingram-Smith, C., Kuettner, H.C., Krzycki, J.A., Leigh, J.A., Li, W., Liu, J., Mukhopadhyay, B., Reeve, J.N., Smith, K., Springer, T.A., Umayam, L.A., White, O., White, R.H., Conway de Macario, E., Ferry, J.G., Jarrell, K.F., Jing, H., Macario, A.J., Paulsen, I., Pritchett, M., Sowers, K.R., Swanson, R.V., Zinder, S.H., Lander, E., Metcalf, W.W., and Birren, B. 2002. The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. *Genome Res.* **12:** 532–542.

Gargouri, A., Lazowska, J., and Slonimski, P.P. 1983. DNA-splicing of introns in the gene: a general way of reverting intron mutations. **In:** *Mitochondria 1983: Nucleo-Mitochondrial Interactions*. pp. 259–268. Schweyen, R.J., Wolf, K., and Kaudewitz, F., Eds., de Gruyter, Berlin.

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.

Gerin, J.L. 2001. Animal models of hepatitis delta virus infection and disease. *ILAR J.* **42:** 103–106.

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.

Goldschmidt-Clermont, M., Choquet, Y., Girard-Bascou, J., Michel, F., Schirmer-Rahire, M., and Rochaix, J.D. 1991. A small chloroplast RNA may be required for trans-splicing in *Chlamydomonas reinhardtii*. *Cell* **65:** 135–143.

Gopalan, V., Vioque, A., and Altman, S. 2002. RNase P: variations and uses. *J. Biol. Chem.* **277:** 6759–6762.

Gorbalenya, A.E. 1994. Self-splicing group I and group II introns encode homologous (putative) DNA endonucleases of a new family. *Protein Sci.* **3:** 1117–1120.

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.

Gottschalk, A., Tang, J., Puig, O., Salgado, J., Neubauer, G., Colot, H.V., Mann, M., Séraphin, B., Rosbash, M., Lührmann, R., and Fabrizio, P. 1998. A comprehensive biochemical and genetic analysis of the yeast U1 snRNP reveals five novel proteins. *RNA* **4:** 374–393.

Gregan, J., Kolisek, M., and Schweyen, R.J. 2001a. Mitochondrial Mg^{2+} homeostasis is critical for group II intron splicing *in vivo*. *Genes Dev.* **15:** 2229–2237.

Gregan, J., Bui, D.M., Pillich, R., Fink, M., Zsurka, G., and Schweyen, R.J. 2001b. The mitochondrial inner membrane protein Lpe10p, a homologue of Mrs2p, is essential for magnesium homeostasis and group II intron splicing in yeast. *Mol. Genet. Genomics* **264:** 773–781.

Griffin, E.A., Jr., Qin, Z., Michels, W.J., Jr., 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.

Grivell, L.A. 1995. Nucleo-mitochondrial interactions in mitochondrial gene expression. *Crit. Rev. Biochem. Mol. Biol.* **30:** 121–164.

Grivell, L.A. 1996. Transposition: mobile introns get into line. *Curr. Biol.* **6:** 48–51.

Grivell, L.A., Artal-Sanz, M., Hakkaart, G., de Jong, L., Nijtmans, L.G., van Oosterum, K., Siep, M., and van der Spek, H. 1999. Mitochondrial assembly in yeast. *FEBS Lett.* **452:** 57–60.

Groudinsky, O., Bousquet, I., Wallis, M.G., Slonimski, P.P., and Dujardin, G. 1993. The NAM1/MTF2 nuclear gene product is selectively required for the stability and/or processing of mitochondrial transcripts of the atp6 and of the mosaic, cox1 and cytb genes in *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* **240:** 419–427.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. 1983. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35:** 849–857.

Guo, H., Karberg, M., Long, M., Jones III, J.P., Sullenger, B., and Lambowitz, A.M. 2000. Group II introns designed to insert into therapeutically relevant DNA target sites in human cells. *Science* **289:** 452–457.

Guo, H., Zimmerly, S., Perlman, P.S., and Lambowitz, A.M. 1997. Group II intron endonucleases use both RNA and protein subunits for recognition of specific sequences in double-stranded DNA. *EMBO J.* **16:** 6835–6848.

Guo, H.C., De Abreu, D.M., Tillier, E.R., Saville, B.J., Olive, J.E., and Collins, R.A. 1993. Nucleotide sequence requirements for self-cleavage of *Neurospora* VS RNA. *J. Mol. Biol.* **232:** 351–361.

Halbreich, A., Pajot, P., Foucher, M., Grandchamp, C., and Slonimski, P. 1980. A pathway of cytochrome b mRNA processing in yeast mitochondria: specific splicing steps and an intron-derived circular DNA. *Cell* **19:** 321–329.

Hammann, C. and Lilley, D.M.J. 2002. Folding and activity of the hammerhead ribozyme. *Chembiochemistry* **3:** 690–700.

Hanna, R. and Doudna, J.A. 2000. Metal ions in ribozyme folding and catalysis. *Curr. Opin. Chem. Biol.* **4:** 166–170.

Harris-Kerr, C.L., Zhang, M., and Peebles, C.L. 1993. The phylogenetically predicted base-pairing interaction between α and α' is required for group II splicing *in vitro*. *Proc. Natl. Acad. Sci. USA* **90:** 10658–10662.

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.

Hensgens, L.A., Arnberg, A.C., Roosendaal, E., van der Horst, G., van der Veen, R., van Ommen, G.J., and Grivell, L.A. 1983. Variation, transcription and circular RNAs of the mitochondrial gene for subunit I of cytochrome c oxidase. *J. Mol. Biol.* **164:** 35–58.

Hertweck, M. and Mueller, M.W. 2001. Mapping divalent metal ion binding sites in a group II intron by Mn^{2+}- and Zn^{2+}-induced site-specific RNA cleavage. *Eur. J. Biochem.* **268:** 4610–4620.

Hess, W.R., Hoch, B., Zeltz, P., Hübschmann, T., Kössel, H., and Börner, T. 1994. Inefficient rpl2 splicing in barley mutants with ribosome-deficient plastids. *Plant Cell* **6:** 1455–1465.

Hetzer, M., Wurzer, G., Schweyen, R.J., and Mueller, M.W. 1997. Trans-activation of group II intron splicing by nuclear U5 snRNA. *Nature* **386:** 417–420.

Hill, J., McGraw, P., and Tzagoloff, A. 1985. A mutation in yeast mitochondrial DNA results in a precise excision of the terminal intron of the cytochrome b gene. *J. Biol. Chem.* **260:** 3235–3238.

Holländer, V. and Kück, U. 1999. Group II intron splicing in chloroplasts: identificationof mutations determining intron stability and fate of exon RNA. *Nucleic Acids Res.* **27:** 2345–2353.

Hübschmann, T., Hess, W.R., and Börner, T. 1996. Impaired splicing of the rps12 transcript in ribosome-deficient plastids. *Plant Mol. Biol.* **30:** 109–123.

Ichiyanagi, K., Beauregard, A., Lawrence, S., Smith, D., Cousineau, B., and Belfort, M. 2002. Retrotransposition

of the Ll.LtrB group II intron proceeds predominantly via reverse splicing into DNA targets. *Mol. Microbiol.* **46:** 1259–1272.

Jacquier, A. 1996. Group II introns: elaborate ribozymes. *Biochimie* **78:** 474–487.

Jacquier, A. and Rosbash, M. 1986. Efficient trans-splicing of a yeast mitochondrial RNA group II intron implicates a strong 5' exon-intron interaction. *Science* **234:** 1099–1104.

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.

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.

Jaeger, L., Michel, F., and Westhof, E. 1994. Involvement of a GNRA tetraloop in long-range RNA tertiary interactions. *J. Mol. Biol.* **236:** 1271–1276.

Jarosch, E., Rödel, G., and Schweyen, R.J. 1997. A soluble 12–kDa protein of the mitochondrial intermembrane space, Mrs11p, is essential for mitochondrial biogenesis and viability of yeast cells. *Mol. Gen. Genet.* **255:** 157–165.

Jarosch, E., Tuller, G., Daum, G., Waldherr, M., Voskova, A., and Schweyen, R.J. 1996. Mrs5p, an essential protein of the mitochondrial intermembrane space, affects protein import into yeast mitochondria. *J. Biol. Chem.* **271:** 17219–17225.

Jarrell, K.A., Dietrich, R.C., and Perlman, P.S. 1988a. Group II intron domain 5 facilitates a *trans*-splicing reaction. *Mol. Cell. Biol.* **8:** 2361–2366.

Jarrell, K.A., Peebles, C.L., Dietrich, R.C., Romiti, S.L., and Perlman, P.S. 1988b. 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.

Jenkins, B.D., Kulhanek, D.J., and Barkan, A. 1997. Nuclear mutations that block group II RNA splicing in maize chloroplasts reveal several intron classes with distinct requirements for splicing factors. *Plant Cell* **9:** 283–296.

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.

Jiménez-Zurdo, J.I., Garcia-Rodriguez, F.M., Barrientos-Duran, A., and Toro, N. 2003. DNA target site requirements for homing *in vivo* of a bacterial group II intron encoding a protein lacking the DNA endonuclease domain. *J. Mol. Biol.* **326:** 413–423.

Karberg, M., Guo, H., Zhong, J., Coon, R., Perutka, J., and Lambowitz, A.M. 2001. Group II introns as controllable gene targeting vectors for genetic manipulation of bacteria. *Nat. Biotechnol.* **19:** 1162–1167.

Kehres, D.G., Lawyer, C.H., and Maguire, M.E. 1998. The CorA magnesium transporter gene family. *Microb. Comp. Genomics* **3:** 151–169.

Kennell, J.C., Moran, J.V., Perlman, P.S., Butow, R.A., and Lambowitz, A.M. 1993. Reverse transcriptase activity associated with maturase-encoding group II introns in yeast mitochondria. *Cell* **73:** 133–146.

Kennell, J.C., Saville, B.J., Mohr, S., Kuiper, M.T., Sabourin, J.R., Collins, R.A., and Lambowitz, A.M. 1995. The VS catalytic RNA replicates by reverse transcription as a satellite of a retroplasmid *Genes Dev.* **9:** 294–303.

Klug, A. and Schwabe, J.W. 1995. Protein motifs 5. Zinc fingers *FASEB J.* **9:** 597–604.

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.

Koehler, C.M., Jarosch, E., Tokatlidis, K., Schmid, K., Schweyen, R.J., and Schatz, G. 1998. Import of mitochondrial carriers mediated by essential proteins of the intermembrane space. *Science* **279:** 369–373.

Kohlstaedt, L.A., Wang, J., Friedman, J.M., Rice, P.A., and Steitz, T.A. 1992. Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256:** 1783–1790.

Koll, H., Schmidt, C., Wiesenberger, G., and Schmelzer, C. 1987. Three nuclear genes suppress a yeast mitochondrial splice defect when present in high copy number. *Curr. Genet.* **12:** 503–509.

Konarska, M., Filipowicz, W., Domdey, H., and Gross, H.J. 1981. Binding of ribosomes to linear and circular forms of the 5'-terminal leader fragment of tobacco-mosaic-virus RNA. *Eur. J. Biochem.* **114:** 221–227.

Konforti, B.B., Liu, Q., and Pyle, A.M. 1998a. A map of the binding site for catalytic domain 5 in the core of a group II intron ribozyme. *EMBO J.* **17:** 7105–7117.

Konforti, B.B., Abramovitz, D.L., Duarte, C.M., Karpeisky, A., Beigelman, L., and Pyle, A.M. 1998b. Ribozyme catalysis from the major groove of group II intron domain 5. *Mol. Cell* **1:** 433–441.

Kozak, M. 1979. Inability of circular mRNA to attach to eukaryotic ribosomes. *Nature* **280:** 82–85.

Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E., and Cech, T.R. 1982. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* **31:** 147–157.

Kück, U., Godehardt, I., and Schmidt, U. 1990. A self-splicing group II intron in the mitochondrial large subunit rRNA (LSUrRNA) gene of the eukaryotic alga *Scenedesmus obliquus*. *Nucleic Acids Res.* **18:** 2691–2697.

Kück, U., Chochet, Y., Schneider, M., Dron, M., and Bennoun, P. 1987. Structural and transcription analysis of two homologous genes for the P700 chlorophyll *a*-apoproteins in *Chlamydomonas reinhardtii*: evidence for *in vivo trans*-splicing. *EMBO J.* **6:** 2185–2195.

Kück, U., Osiewacz, H.D., Schmidt, U., Kappelhoff, B., Schulte, E., Stahl, U., and Esser, K. 1985. The onset of senescence is affected by DNA rearrangements of a discontinuous mitochondrial gene in *Podospora anserina*. *Curr. Genet.* **9:** 373–382.

Kurz, J.C. and Fierke, C.A. 2000. Ribonuclease P: a ribonucleoprotein enzyme. *Curr. Opin. Chem. Biol.* **4:** 553–538.

Kwakman, J.H., Konings, D., Pel, H.J., and Grivell, L.A. 1989. Structure-function relationships in a self-splicing group II intron: a large part of domain II of the mitochondrial intron aI5 is not essential for self-splicing. *Nucleic Acids Res.* **17:** 4205–4216.

Kwakman, J.H., Konings, D.A., Hogeweg, P., Pel, H.J., and Grivell, L.A. 1990. Structural analysis of a group II intron by chemical modifications and minimal energy calculations. *J. Biomol. Struct. Dyn.* **8:** 413–430.

Lafontaine, D.A., Norman, D.G., and Lilley, D.M.J. 2002. Folding and catalysis by the VS ribozyme. *Biochimie* **84:** 889–896.

Lai, M.M. 1995. The molecular biology of hepatitis delta virus. *Annu. Rev. Biochem.* **64:** 259–286.

Lambowitz, A.M. and Perlman, P.S. 1990. Involvement of aminoacyl-tRNA synthetases and other proteins in group I and group II intron splicing. *Trends Biochem. Sci.* **15:** 440–444.

Lambowitz, A.M. and Belfort, M. 1993. Introns as mobile genetic elements. *Annu. Rev. Biochem.* **62:** 587–622.

Lambowitz, A.M. Caprara, M.G., Zimmerly, S., and Perlman, P.S. 1999. Group I and group II ribozymes as RNPs: clues to the past and guides to the future. **In:** *The RNA World, 2nd edition*. pp. 451–485. Gesteland, R.F., Cech, T.R., and Atkins, J.F., Eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Lazowska, J., Meunier, B., and Macadre, C. 1994. Homing of a group II intron in yeast mitochondrial DNA is accompanied by unidirectional co-conversion of upstream-located markers. *EMBO J.* **13:** 4963–4972.

Leeds, P., Peltz, S.W., Jacobson, A., and Culbertson, M.R. 1991. The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon. *Genes Dev.* **5:** 2303–2314.

Leeds, P., Wood, J.M., Lee, B.S., and Culbertson, M.R. 1992. Gene products that promote mRNA turnover in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **12:** 2165–2177.

Levra-Juillet, E., Boulet, A., Séraphin, B., Simon, M., and Faye, G. 1989. Mitochondrial introns aI1 and/or aI2 are needed for the *in vivo* deletion of intervening sequences. *Mol. Gen. Genet.* **217:** 168–171.

Lilley, D.M.J. 1995. Kinking of DNA and RNA by base bulges. *Proc. Natl. Acad. Sci. USA* **92:** 7140–7142.

Lilley, D.M.J. 1999a. Folding and catalysis by the hairpin ribozyme. *FEBS Lett.* **452:** 26–30.

Lilley, D.M.J. 1999b. Structure, folding and catalysis of the small nucleolytic ribozymes. *Curr. Opin. Struct. Biol.* **9:** 330–338.

Lilley, D.M.J. 2001. Origins of RNA catalysis in the hairpin ribozyme. *Chembiochem.* **2:** 729–733.

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.

Lüking, A., Stahl, U., and Schmidt, U. 1998. The protein family of RNA helicases. *Crit. Rev. Biochem. Mol. Biol.* **33:** 259–296.

Lykke-Andersen, J., Aagaard, C., Semionenkov, M., and Garrett, R.A. 1997. Archaeal introns: splicing, intercellular mobility and evolution. *Trends Biochem. Sci.* **22:** 326–331.

Malik, H.S., Burke, W.D., and Eickbush, T.H. 1999. The age and evolution of non-LTR retrotransposable elements *Mol. Biol. Evol.* **16:** 793–805.

Martinez-Abarca, F. and Toro, N. 2000a. Group II introns in the bacterial world. *Mol. Microbiol.* **38:** 917–926.

Martinez-Abarca, F. and Toro, N. 2000b. RecA-independent ectopic transposition in vivo of a bacterial group II intron. *Nucleic Acids Res.* **28:** 4397–4402.

Martinez-Abarca, F., Garcia-Rodriguez, F.M., and Toro, N. 2000. Homing of a bacterial group II intron with an intron-encoded protein lacking a recognizable endonuclease domain. *Mol. Microbiol.* **35:** 1405–1412.

Martinez-Abarca, F., Garcia-Rodriguez, F.M., Munoz, E., and Toro, N. 1999. Biochemical demonstration

of reverse transcriptase activity and splicing efficiency of a bacterial group II intron from *Sinorhizobium meliloti*. *Nucleic Acids Symp. Ser.* **41:** 117–119.

Massire, C., Jaeger, L., and Westhof, E. 1998. Derivation of the three-dimensional architecture of bacterial ribonuclease P RNAs from comparative sequence analysis. *J. Mol. Biol.* **279:** 773–793.

Matsuura, M., Noah, J.W., and Lambowitz, A.M. 2001. Mechanism of maturase-promoted group II intron splicing. *EMBO J.* **20:** 7259–7270.

Matsuura, M., Saldanha, R., Ma, H., Wank, H., Yang, J., Mohr, G., Cavanagh, S., Dunny, G.M., Belfort, M., and Lambowitz, A.M. 1997. A bacterial group II intron encoding reverse transcriptase, maturase, and DNA endonuclease activities: biochemical demonstration of maturase activity and insertion of new genetic information within the intron. *Genes Dev.* **11:** 2910–2924.

McClure, M.A. 1991. Evolution of retroposons by acquisition or deletion of retrovirus-like genes. *Mol. Biol. Evol.* **8:** 835–856.

McKay, D.B. and Wedekind, J.E. 1999. Small ribozymes. **In:** *The RNA World, 2nd edition*. pp. 265–286. Gesteland, R.F., Cech, T.R., and Atkins, J.F., Eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Merlos-Lange, A.M., Kanbay, F., Zimmer, M., and Wolf, K. 1987. DNA splicing of mitochondrial group I and II introns in *Schizosaccharomyces pombe*. *Mol. Gen. Genet.* **206:** 273–278.

Meunier, B., Tian, G.-L., Macadre, C., Slonimski, P.P., and Lazowska, J. 1990. Group II introns transpose in yeast mitochondria. **In:** *Structure, Function, and Biogenesis of Energy Transfer Systems.* pp. 169–174. Quagliariello, E., Papa, S., Palmieri, F., and Saccone, C., Eds., Elsevier, Amsterdam.

Michel, F. and Dujon, B. 1983. Conservation of RNA secondary structure in two intron families including mitochondrial-, chloroplast-, and nuclear-encoded members. *EMBO J.* **2:** 33–38.

Michel, F. and Lang, B.F. 1985. Mitochondrial class II introns encode proteins related to the reverse transcriptases of retroviruses. *Nature* **316:** 641–643.

Michel, F. and Jacquier, A. 1987. Long-range intron-exon and intron-intron pairings involved in self-splicing of class II catalytic introns. *Cold Spring Harb. Symp. Quant. Biol.* **52:** 201–212.

Michel, F. and Ferat, J.-L. 1995. Structure and activities of group II introns. *Annu. Rev. Biochem.* **64:** 435–461.

Michel, F., Jacquier, A., and Dujon, B. 1982. Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure. *Biochimie* **64:** 867–861.

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., Jr. 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.

Mills, D.A., McKay, L.L., and Dunny, G.M. 1996. Splicing of a group II intron involved in the conjugative transfer of pRS01 in Lactococci. *J. Bacteriol.* **178:** 3531–3538.

Mills, D.A., Manias, D.A., McKay, L.L., and Dunny, G.M. 1997. Homing of a group II intron from *Lactococcus lactis* subsp. lactis ML3. *J. Bacteriol.* **179:** 6107–6111.

Mohr, G. and Lambowitz, A.M. 2003. Putative proteins related to group II 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, G., Smith, D., Belfort, M., and Lambowitz, A.M. 2000. Rules for DNA target-site recognition by a lactococcal group II intron enable retargeting of the intron to specific DNA sequences *Genes Dev.* **14:** 559–573.

Moran, J.V., Mecklenburg, K.L., Sass, P., Belcher, S.M., Mahnke, D., Lewin, A., and Perlman, P. 1994. Splicing defective mutants of the *COXI* gene of yeast mitochondrial DNA: initial definition of the maturase domain of the group II intron aI2. *Nucleic Acids Res.* **22:** 2057–2064.

Moran, J.V., Zimmerly, S., Eskes, R., Kennell, J.C., Lambowitz, A.M., Butow, R.A., and Perlman, P.S. 1995. Mobile group II introns of yeast mitochondrial DNA are novel site-specific retroelements. *Mol. Cell. Biol.* **15:** 2828–2838.

Mörl, M. and Schmelzer, C. 1990a. Group II intron RNA-catalyzed recombination of RNA *in vitro*. *Nucleic Acids Res.* **18:** 6545–6551.

Mörl, M. and Schmelzer, C. 1990b. Integration of group II intron bI1 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.

Morozova, T., Seo, W., and Zimmerly, S. 2002. Non-cognate template usage and alternative priming by a group II intron-encoded reverse transcriptase. *J. Mol. Biol.* **315:** 951–963.

Mueller, M.W., Schweyen, R.J., and Schmelzer, C. 1988. Selection of cryptic 5' splice sites by group II intron RNAs *in vitro*. *Nucleic Acids Res.* **16:** 7383–7395.

Mueller, M.W., Hetzer, M., and Schweyen, R.J. 1993a. Group II intron RNA catalysis of progressive nucleotide insertion: a model for RNA editing. *Science* **261:** 1035–1038.

Mueller, M.W., Stocker, P., Hetzer, M., and Schweyen, R.J. 1991. Fate of the junction phosphate in alternating forward and reverse self-splicing reactions of group II intron RNA. *J. Mol. Biol.* **222:** 145–154.

Mueller, M.W., Allmaier, M., Eskes, R., and Schweyen, R.J. 1993b. Transposition of group II intron aI1 in yeast and invasion of mitochondrial genes at new locations. *Nature* **366:** 174–176.

Munoz, E., Villadas, P.J., and Toro, N. 2001. Ectopic transposition of a group II intron in natural bacterial populations. *Mol. Microbiol.* **41:** 645–652.

Murphy, F.L. and Cech, T.R. 1994. GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. *J. Mol. Biol.* **236:** 49–63.

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.

Niemer, I., Schmelzer, C., and Börner, G.V. 1995. Overexpression of DEAD box protein pMSS116 promotes ATP-dependent splicing of a yeast group II intron *in vitro*. *Nucleic Acids Res.* **23:** 2966–2972.

Ogawa, H., Johzuka, K., Nakagawa, T., Leem, S.H., and Hagihara, A.H. 1995. Functions of the yeast meiotic recombination genes, *MRE11* and *MRE2*. *Adv. Biophys.* **31:** 67–76.

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 Harb. Symp. Quant. Biol.* **52:** 223–232.

Peebles, C.L., Zhang, M., Perlman, P.S., and Franzen, J.S. 1995. Catalytically critical nucleotide in domain 5 of a group II intron. *Proc. Natl. Acad. Sci. USA* **92:** 4422–4426.

Peebles, C.L., Belcher, S.M., Zhang, M., Dietrich, R.C., and Perlman, P.S. 1993. Mutation of the conserved first nucleotide of a group II intron from yeast mitochondrial DNA reduces the rate but allows accurate splicing. *J. Biol. Chem.* **268:** 11929–11938.

Peebles, C.L., Perlman, P.S., Mecklenburg, K.L., Petrillo, M.L., Tabor, J.H., Jarrell, K.A., and Cheng, H.L. 1986. A self-splicing RNA excises an intron lariat. *Cell* **44:** 213–223.

Pel, H.J., Tzagoloff, A., and Grivell, L.A. 1992a. The identification of 18 nuclear genes required for the expression of the yeast mitochondrial gene encoding cytochrome c oxidase subunit 1. *Curr. Genet.* **21:** 139–146.

Pel, H.J., Maat, C., Rep, M., and Grivell, L.A. 1992b. The yeast nuclear gene *MRF1* encodes a mitochondrial peptide chain release factor and cures several mitochondrial RNA splicing defects. *Nucleic Acids Res.* **20:** 6339–6346.

Perea, J. and Jacq, C. 1985. Role of the 5' hairpin structure in the splicing accuracy of the fourth intron of the yeast cob-box gene. *EMBO J.* **4:** 3281–3288.

Perlman, P.S. and Podar, M. 1996. Reactions catalyzed by group II introns *in vitro*. *Methods Enzymol.* **264:** 66–86.

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.

Pley, H.W., Flaherty, K.M., and McKay, D.B. 1994. Model for an RNA tertiary interaction from the structure of an intermolecular complex between a GAAA tetraloop and an RNA helix. *Nature* **372:** 111–113.

Podar, M., Perlman, P.S., and Padgett, R.A. 1995a. Stereochemical selectivity of group II intron splicing, reverse splicing, and hydrolysis reactions. *Mol. Cell. Biol.* **15:** 4466–4478.

Podar, M., Dib-Hajj, S.D., and Perlman, P.S. 1995b. A UV-induced, Mg^{2+}-dependent crosslink traps an active form of domain 3 of a self-splicing group II intron. *RNA* **1:** 828–840.

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., Zhuo, 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.

Portmann, S., Grimm, S., Workman, C., Usman, N., and Egli, M. 1996. Crystal structures of an A-form duplex with single-adenosine bulges and a conformational

basis for site-specific RNA self-cleavage. *Chem. Biol.* **3:** 173–184.

Puig, O., Gottschalk, A., Fabrizio, P., and Seraphin, B. 1999. Interaction of the U1 snRNP with nonconserved intronic sequences affects 5' splice site selection. *Genes Dev.* **13:** 569–580.

Pyle, A.M. 1993. Ribozymes: a distinct class of metalloenzymes. *Science* **261:** 709–714.

Pyle, A.M. 1996. Catalytic reaction mechanisms and structural features of group II intron ribozymes. **In:** *Nucleic Acids and Molecular Biology*, Vol. 10. pp. 75–107. Eckstein, F. and Lilley, D.M.J., Eds., Springer-Verlag, Berlin.

Pyle, A.M. 2002. Metal ions in the structure and function of RNA. *J. Biol. Inorg. Chem.* **7:** 679–690.

Pyle, A.M. and Green, J.B. 1994. Building a kinetic framework for group II intron ribozyme activity: quantitation of interdomain binding and reaction rate. *Biochemistry* **33:** 2716–2725.

Qin, P.Z. and Pyle, A.M. 1997. Stopped-flow fluorescence spectroscopy of a group II intron ribozyme reveals that domain 1 is an independent folding unit with a requirement for specific Mg^{2+} ions in the tertiary structure. *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.* **291:** 15–27.

Reed, R. 2000. Mechanisms of fidelity in pre-mRNA splicing. *Curr. Opin. Cell Biol.* **12:** 340–345.

Richard, G.F. and Dujon, B. 1997. Association of transcripts from a group-I intron-containing gene with high sedimentation coefficient particles. *Curr. Genet.* **32:** 175–181.

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.

Sägebarth, R., Wedde, M., Schmidt, U., and Stahl, U. 1994. Tandem duplication of group II introns on the mitochondrial genome of the ascomycete *Podospora anserina*, *2nd European Conference on Fungal Genetics*, Lunteren, The Netherlands.

Sainsard-Chanet, A., Begel, O., and Belcour, L. 1993. DNA deletion of mitochondrial introns is correlated with the process of senescence in *Podospora anserina*. *J. Mol. Biol.* **234:** 1–7.

Sainsard-Chanet, A., Begel, O., and Belcour, L. 1994. DNA double-strand break in vivo at the 3' extremity of exons located upstream of group II introns. Senescence and circular DNA introns in *Podospora* mitochondria. *J. Mol. Biol.* **242:** 630–643.

Saldanha, R., Mohr, G., Belfort, M., and Lambowitz, A.M. 1993. Group I and group II introns. *FASEB J.* **7:** 15–24.

Saldanha, R., Chen, B., Wank, H., Matsuura, M., Edwards, J., and Lambowitz, A.M. 1999. RNA and protein catalysis in group II intron splicing and mobility reactions using purified components. *Biochemistry* **38:** 9069–9083.

San Filippo, J. and Lambowitz, A.M. 2002. Characterization of the C-terminal DNA-binding/DNA endonuclease region of a group II intron-encoded protein. *J. Mol. Biol.* **324:** 933–951.

Saville, B.J. and Collins, R.A. 1990. A site-specific self-cleavage reaction performed by a novel RNA in *Neurospora* mitochondria. *Cell* **61:** 685–696.

Schäfer, B., Merlos-Lange, A.M., Anderl, C., Welser, F., Zimmer, M., and Wolf, K. 1991. The mitochondrial genome of fission yeast: inability of all introns to splice autocatalytically, and construction and characterization of an intronless genome. *Mol. Gen. Genet.* **225:** 158–167.

Schmelzer, C. and Schweyen, R.J. 1986. Self-splicing of group II introns *in vitro*: mapping of the branch point and mutational inhibition of lariat formation. *Cell* **46:** 557–565.

Schmelzer, C. and Mueller, M.W. 1987. Self-splicing of group II introns *in vitro*: lariat formation and 3' splice site selection in mutant RNAs. *Cell* **51:** 753–762.

Schmelzer, C., Schmidt, C., and Schweyen, R.J. 1982. Identification of splicing signals in introns of yeast mitochondrial split genes: mutational alterations in intron bI1 and secondary structures in related introns. *Nucleic Acids Res.* **10:** 6797–6808.

Schmelzer, C., Schmidt, C., May, K., and Schweyen, R.J. 1983. Determination of functional domains in intron bI1 of yeast mitochondrial RNA by studies of mitochondrial mutations and a nuclear suppressor. *EMBO J.* **2:** 2047–2052.

Schmidt, C., Söllner, T., and Schweyen, R.J. 1987a. Nuclear suppression of a mitochondrial RNA splice defect: nucleotide sequence and disruption of the *MRS3* gene. *Mol. Gen. Genet.* **210:** 145–152.

Schmidt, U., Kosack, M., and Stahl, U. 1987b. Lariat RNA of a group II intron in a filamentous fungus. *Curr. Genet.* **12:** 291–295.

Schmidt, U., Budde, E., and Stahl, U. 1992. Self-splicing of a mitochondrial group I intron from the cytochrome b gene of the ascomycete *Podospora anserina*. *Mol. Gen. Genet.* **233:** 71–80.

Schmidt, U., Lehmann, K., and Stahl, U. 2002. A novel mitochondrial DEAD box protein (Mrh4) required for maintenance of mtDNA in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **2:** 267–276.

Schmidt, U., Sägebarth, R., Schmelzer, C., and Stahl, U. 1993. Self-splicing of a *Podospora anserina* group IIA intron in vitro. Effects of 3'-terminal intron alterations on cleavage at the 5' and 3' splice site. *J. Mol. Biol.* **231:** 559–568.

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.

Schmidt, U., Riederer, B., Mörl, M., Schmelzer, C., and Stahl, U. 1990. Self-splicing of the mobile group II intron of the filamentous fungus *Podospora anserina* (COI I1) *in vitro*. *EMBO J.* **9:** 2289–2298.

Schmidt, U., Maue, I., Lehmann, K., Belcher, S.M., Stahl, U., and Perlman, P.S. 1998. Mutant alleles of the MRS2 gene of yeast nuclear DNA suppress mutations in the catalytic core of a mitochondrial group II intron. *J. Mol. Biol.* **282:** 525–541.

Schmidt, W.M., Schweyen, R.J., Wolf, K., and Mueller, M.W. 1994. Transposable group II introns in fission and budding yeast. Site-specific genomic instabilities and formation of group II IVS plDNAs. *J. Mol. Biol.* **243:** 157–166.

Schock, I., Gregan, J., Steinhauser, S., Schweyen, R.J., Brennicke, A., and Knoop, V. 2000. A member of a novel *Arabidopsis thaliana* gene family of candidate Mg^{2+} ion transporters complements a yeast mitochondrial group II intron-splicing mutant. *Plant J.* **24:** 489–501.

Schön, A. 1999. Ribonuclease P: the diversity of a ubiquitous RNA processing enzyme. *FEMS Microbiol. Rev.* **23:** 391–406.

Scott, W.G. and Klug, A. 1996. Ribozymes: structure and mechanism in RNA catalysis. *Trends Biochem. Sci.* **21:** 220–224.

Sellem, C.H., Lecellier, G., and Belcour, L. 1993. Transposition of a group II intron. *Nature* **366:** 176–178.

Sellem, C.H., Begel, O., and Sainsard-Chanet, A. 2000. Recombinant mitochondrial DNA molecules suggest a template switching ability for group-II-intron reverse transcriptase. *Curr. Genet.* **37:** 24–28.

Séraphin, B., Simon, M., and Faye, G. 1988. *MSS18*, a yeast nuclear gene involved in the splicing of intron aI5b of the mitochondrial COX1 transcript. *EMBO J.* **7:** 1455–1464.

Séraphin, B., Boulet, A., Simon, M., and Faye, G. 1987. Construction of a yeast strain devoid of mitochondrial introns and its use to screen nuclear genes involved in mitochondrial splicing. *Proc. Natl. Acad. Sci. USA* **84:** 6810–6814.

Séraphin, B., Simon, M., Boulet, A., and Faye, G. 1989. Mitochondrial splicing requires a protein from a novel helicase family. *Nature* **337:** 84–87.

Shearman, C., Godon, J.J., and Gasson, M. 1996. Splicing of a group II intron in a functional transfer gene of *Lactococcus lactis*. *Mol. Microbiol.* **21:** 45–53.

Shih, I.H. and Been, M.D. 2002. Catalytic strategies of the hepatitis delta virus ribozymes. *Annu. Rev. Biochem.* **71:** 887–917.

Shub, D.A., Goodrich-Blair, H., and Eddy, S.R. 1994. Amino acid sequence motif of group I intron endonucleases is conserved in open reading frames of group II introns. *Trends Biochem. Sci.* **19:** 402–404.

Sigel, R.K., Vaidya, A., and Pyle, A.M. 2000. Metal ion binding sites in a group II intron core. *Nat. Struct. Biol.* **7:** 1111–1116.

Simon, M. and Faye, G. 1984. Steps in processing of the mitochondrial cytochrome oxidase subunit I premRNA affected by a nuclear mutation in yeast. *Proc. Natl. Acad. Sci. USA* **81:** 8–12.

Singh, N.N. and Lambowitz, A.M. 2001. Interaction of a group II intron ribonucleoprotein endonuclease with its DNA target site investigated by DNA footprinting and modification interference. *J. Mol. Biol.* **309:** 361–386.

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.

Skelly, P.J., Hardy, C.M., and Clark-Walker, G.D. 1991. A mobile group II intron of a naturally occurring rearranged mitochondrial genome in *Kluyveromyces lactis*. *Curr. Genet.* **20:** 115–120.

Sontheimer, E.J., Gordon, P.M., and Piccirilli, J.A. 1999. Metal ion catalysis during group II intron self-splicing: parallels with the spliceosome. *Genes Dev.* **13:** 1729–1741.

Stahl, U., Kück, U., Tudzynski, P., and Esser, K. 1980. Characterization and cloning of plasmid like DNA of the ascomycete *Podospora anserina*. *Mol. Gen. Genet.* **178:** 639–646.

Stahl, U., Lemke, P.A., Tudzynski, P., Kück, U., and Esser, K. 1978. Evidence for plasmid like DNA in a

filamentous fungus, the ascomycete *Podospora anserina*. *Mol. Gen. Genet.* **162:** 341–343.

Steitz, T.A. and Steitz, J.A. 1993. A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl. Acad. Sci. USA* **90:** 6498–6502.

Steitz, T.A., Smerdon, S., Jager, J., Wang, J., Kohlstaedt, L.A., Friedman, J.M., Beese, L.S., and Rice, P.A. 1993. Two DNA polymerases: HIV reverse transcriptase and the Klenow fragment of *Escherichia coli* DNA polymerase I. *Cold Spring Harb. Symp. Quant. Biol.* **58:** 495–504.

Strobel, S.A. 1999. A chemogenetic approach to RNA function/structure analysis. *Curr. Opin. Struct. Biol.* **9:** 346–352.

Su, L.J., Qin, P.Z., Michels, W.J., and Pyle, A.M. 2001. Guiding ribozyme cleavage through motif recognition: the mechanism of cleavage site selection by a group II intron ribozyme. *J. Mol. Biol.* **306:** 655–668.

Swisher, J., Duarte, C.M., Su, L.J., and Pyle, A.M. 2001. Visualizing the solvent-inaccessible core of a group II intron ribozyme. *EMBO J.* **20:** 2051–2061.

Swisher, J.F., Su, L.J., Brenowitz, M., Anderson, V.E., and Pyle, A.M. 2002. Productive folding to the native state by a group II intron ribozyme. *J. Mol. Biol.* **315:** 297–310.

Symons, R.H. 1997. Plant pathogenic RNAs and RNA catalysis. *Nucleic Acids Res.* **25:** 2683–2689.

Tanner, N.K. 1999. Ribozymes: the characteristics and properties of catalytic RNAs. *FEMS Microbiol. Rev.* **23:** 257–275.

Thomas, B.C., Li, X., Gegenheimer, P. 2000. Chloroplast ribonuclease P does not utilize the ribozyme-type pre-tRNA cleavage mechanism. *RNA* **6:** 545–553.

Thompson, M.D., Zhang, L., Hong, L., and Hallick, R.B. 1997. Extensive structural conservation exists among several homologs of two *Euglena* chloroplast group II introns. *Mol. Gen. Genet.* **257:** 45–54.

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.

Tzagoloff, A. and Dieckmann, C.L. 1990. *PET* genes of *Saccharomyces cerevisiae*. *Microbiol. Rev.* **54:** 211–225.

van der Veen, R., Kwakman, J.H., and Grivell, L.A. 1987a. Mutations at the lariat acceptor site allow self-splicing of a group II intron without lariat formation. *EMBO J.* **6:** 3827–3831.

van der Veen, R., Arnberg, A.C., and Grivell, L.A. 1987b. Self-splicing of a group II intron in yeast mitochondria: dependence on 5' exon sequences. *EMBO J.* **6:** 1079–1084.

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 Dyck, E., Jank, B., Ragnini, A., Schweyen, R.J., Duyckaerts, C., Sluse, F., and Foury, F. 1995. Overexpression of a novel member of the mitochondrial carrier family rescues defects in both DNA and RNA metabolism in yeast mitochondria. *Mol. Gen. Genet.* **246:** 426–436.

Varani, G. and Nagai, K. 1998. RNA recognition by RNP proteins during RNA processing. *Annu. Rev. Biophys. Biomol. Struct.* **27:** 407–445.

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.R. 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.

Vogel, J., Hübschmann, T., Börner, T., and Hess, W.R. 1997. Splicing and intron-internal RNA editing of trnK-matK transcripts in barley plastids: support for MatK as an essential splice factor. *J. Mol. Biol.* **270:** 179–187.

Waldherr, M., Ragnini, A., Jank, B., Teply, R., Wiesenberger, G., and Schweyen, R.J. 1993. A multitude of suppressors of group II intron-splicing defects in yeast. *Curr. Genet.* **24:** 301–306.

Wallasch, C., Mörl, M., Niemer, I., and Schmelzer, C. 1991. Structural requirements for selection of 5'- and 3' splice sites of group II introns. *Nucleic Acids Res.* **19:** 3307–3314.

Wallis, M.G., Groudinsky, O., Slonimski, P.P., and Dujardin, G. 1994. The NAM1 protein (NAM1p), which is selectively required for cox1, cytb and atp6 transcript processing/stabilisation, is located in the yeast mitochondrial matrix. *Eur. J. Biochem.* **222:** 27–32.

Walter, N.G. and Burke, J.M. 1998. The hairpin ribozyme: structure, assembly and catalysis. *Curr. Opin. Chem. Biol.* **2:** 24–30.

Wank, H., SanFilippo, J., Singh, R.N., Matsuura, M., and Lambowitz, A.M. 1999. A reverse transcriptase/maturase promotes splicing by binding at its own

coding segment in a group II intron RNA. *Mol. Cell* **4:** 239–250.

Weeks, K.M. and Crothers, D.M. 1991. RNA recognition by Tat-derived peptides: interaction in the major groove? *Cell* **66:** 577–588.

Wiesenberger, G., Waldherr, M., and Schweyen, R.J. 1992. The nuclear gene *MRS2* is essential for the excision of group II introns from yeast mitochondrial transcripts *in vivo*. *J. Biol. Chem.* **267:** 6963–6969.

Wiesenberger, G., Link, T.A., von Ahsen, U., Waldherr, M., and Schweyen, R.J. 1991. *MRS3* and *MRS4*, two suppressors of mtRNA splicing defects in yeast, are new members of the mitochondrial carrier family. *J. Mol. Biol.* **217:** 23–37.

Woodson, S.A. and Cech, T.R. 1989. Reverse self-splicing of the *Tetrahymena* group I intron: implication for the directionality of splicing and for intron transposition. *Cell* **57:** 335–345.

Xiang, Q., Qin, P.Z., Michels, W.J., Freeland, K., and Pyle, A.M. 1998. Sequence specificity of a group II intron ribozyme: multiple mechanisms for promoting unusually high discrimination against mismatched targets. *Biochemistry* **37:** 3839–3849.

Xiao, S., Scott, F., Fierke, C.A., and Engelke, D.R. 2002. Eukaryotic ribonuclease P: a plurality of ribonucleoprotein enzymes. *Annu. Rev. Biochem.* **71:** 165–189.

Xiong, Y. and Eickbush, T.H. 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J.* **9:** 3353–3362.

Yang, J., Zimmerly, S., Perlman, P.S., and Lambowitz, A.M. 1996. Efficient integration of an intron RNA into double-stranded DNA by reverse splicing. *Nature* **381:** 332–335.

Yang, J., Mohr, G., Perlman, P.S., and Lambowitz, A.M. 1998. Group II intron mobility in yeast mitochondria: target DNA-primed reverse transcription activity of aI1 and reverse splicing into DNA transposition sites *in vitro*. *J. Mol. Biol.* **282:** 505–523.

Zennaro, E., Grimaldi, L., Baldacci, G., and Frontali, L. 1985. Mitochondrial transcription and processing of transcripts during release from glucose repression in 'resting cells' of *Saccharomyces cerevisiae*. *Eur. J. Biochem.* **147:** 191–196.

Zhang, L. and Doudna, J.A. 2002. Structural insights into group II intron catalysis and branch-site selection. *Science* **295:** 2084–2088.

Zhang, L., Jenkins, K.P., Stutz, E., and Hallick, R.B. 1995. The *Euglena gracilis* intron-encoded mat2 locus is interrupted by three additional group II introns *RNA* **1:** 1079–1088.

Zhong, J., Karberg, M., and Lambowitz, A.M. 2003. Targeted and random bacterial gene disruption using a group II intron (targetron) vector containing a retrotransposition-activated selectable marker. *Nucleic Acids Res.* **31:** 1656–1664.

Zimmerly, S., Hausner, G., and Wu, X. 2001. Phylogenetic relationships among group II intron ORFs *Nucleic Acids Res.* **29:** 1238–1250.

Zimmerly, S., Guo, H., Perlman, P.S., and Lambowitz, A.M. 1995a. Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell* **82:** 545–554.

Zimmerly, S., Moran, J.V., Perlman, P.S., and Lambowitz, A.M. 1999. Group II intron reverse transcriptase in yeast mitochondria. Stabilization and regulation of reverse transcriptase activity by the intron RNA. *J. Mol. Biol.* **289:** 473–490.

Zimmerly, S., Guo, H., Eskes, R., Yang, J., Perlman, P.S., and Lambowitz, A.M. 1995b. A group II intron RNA is a catalytic component of a DNA endonuclease involved in intron mobility. *Cell* **83:** 529–538.

Zsurka, G., Gregan, J., and Schweyen, R.J. 2001. The human mitochondrial Mrs2 protein functionally substitutes for its yeast homologue, a candidate magnesium transporter. *Genomics* **72:** 158–168.