Nucleic Acids Research, Vol. 20, No. 7 1747–1754

A maturase-encoding group IIA intron of yeast mitochondria self-splices *in vitro*

Sharda K.Hebbar, Scott M.Belcher[1] and Philip S.Perlman[1]
Department of Molecular Genetics, The Ohio State University, 484 W. 12th Avenue, Columbus, OH 43210 and [1]Department of Biochemistry, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75345, USA

Received November 4, 1991; Revised and Accepted February 20, 1992 EMBL accession no. Z11698

ABSTRACT

Intron 1 of the *coxI* gene of yeast mitochondrial DNA (aI1) is a group IIA intron that encodes a maturase function required for its splicing *in vivo*. It is shown here to self-splice *in vitro* under some reaction conditions reported earlier to yield efficient self-splicing of group IIB introns of yeast mtDNA that do not encode maturase functions. Unlike the group IIB introns, aI1 is inactive in 10 mM Mg^{2+} (including spermidine) and requires much higher levels of Mg^{2+} and added salts (1M NH_4Cl or KCl or 2M $(NH_4)_2SO_4$) for ready detection of splicing activity. In KCl-stimulated reactions, splicing occurs with little normal branch formation; a post-splicing reaction of linear excised intron RNA that forms shorter lariat RNAs with branches at cryptic sites was evident in those samples. At low levels of added NH_4Cl or KCl, the precursor RNA carries out the first reaction step but appears blocked in the splicing step. AI1 RNA is most reactive at 37–42°C, as compared with 45°C for the group IIB introns; and it lacks the KCl- or NH_4Cl-dependent spliced-exon reopening reaction that is evident for the self-splicing group IIB introns of yeast mitochondria. Like the group IIB intron aI5γ, the domain 4 of aI1 can be largely deleted in *cis*, without blocking splicing; also, *trans*-splicing of half molecules interrupted in domain 4 occurs. This is the first report of a maturase-encoding intron of either group I or group II that self-splices *in vitro*.

INTRODUCTION

Introns in fungal mitochondria, plant mitochondria and chloroplasts have been classified as group I or group II on the basis of group-specific short conserved sequences and secondary structure elements (reviewed in 1, 2). Examples of each intron type self-splice by a two-step transesterification mechanism. Group I introns, such as the *Tetrahymena* nuclear large rRNA gene intron, require a guanosine cofactor to initiate the first step in splicing and the excised linear intron is subsequently reconfigured as a circle (reviewed in 3). In contrast, group II intron splicing occurs independent of a nucleotide cofactor and splicing is initiated by the 2′-hydroxyl group of an internal adenosine residue leading to the formation of a branched intron 'lariat' similar to that formed during splicing of nuclear pre-mRNA introns (reviewed in 4).

While some representatives of both intron groups readily self-splice *in vitro*, others do not (see ref. 4) The reason for this difference is not entirely clear but may be due to the absence of protein cofactors. There is growing evidence that nuclear gene products are needed for group I and group II intron splicing in a number of organisms (reviewed in 4–6). In addition, strong genetic and biochemical evidence shows that some group I and group II introns encode polypeptides (maturases) that are required for their own splicing (c.f., ref. 4). It has been hypothesized that maturases participate in splicing by stabilizing some secondary or tertiary structure needed for splicing. Their exact function, however, remains uncharacterized.

There are four group II introns in yeast mitochondria, three in the cytochrome c oxidase subunit I (*coxI*) gene and one in the apocytochrome b (*cob*) gene. Both introns assigned to the IIB subgroup (*coxI* intron 5γ [aI5γ] and *cob* intron 1 [bI1]) self-splice *in vitro* (7–9). The two remaining introns, *coxI* intron 1 (aI1) and intron 2 (aI2), are of the IIA subtype (see ref. 2). They are about three times larger (2.4 kb vs 0.9 kb) and contain long open reading frames (ORFs) that code for maturases (10–12). In this paper, we report that the maturase-encoding aI1 self splices efficiently *in vitro* under a subset of conditions suitable for aI5γ and bI1 self-splicing and characterize the main products of those reactions including some novel post-splicing reactions.

MATERIALS AND METHODS

Plasmids and Yeast Strains

Mitochondrial DNA (mtDNA) from strain ID41-6/161 ϱ+ was isolated as described in ref. 13. A 3124 nt HpaII–Eco RI fragment contains intron 1 of the *coxI* gene plus some flanking sequences (see Figure 1). That fragment was purified and ligated into pBSM13+ (Strategene, Inc.) that had been cleaved with Acc I and Eco RI to yield plasmid pSH2. The DNA sequence of most of that fragment from strain D273-10B is published (14). We have sequenced much of that region from our strain and have detected a modest number of presumably strain-specific sequence changes that will be compiled separately. There is a 95 bp gap in the published sequence immediately following the Hpa II site (summarized by de Zamaroczy and Bernardi (15)); we have

determined that sequence from strain ID41-6/161 and report it in the legend to Figure 1. Plasmid pSHDAC was constructed by cleaving pSH2 with Acc I and Cla I and religating the complementary ends; the predicted sequence at the junction between the Acc I and Cla I sites was confirmed by sequencing. The ClaI−Eco RI insert fragment from pSH2 (see Figure 1) was subcloned into Acc I and Eco RI cleaved pBSM13+ to yield plasmid pSH4.

Enzyme Reagents

Restriction enzymes were purchased from Bethesda Research Laboratories and New England Biolabs. T3 RNA polymerase was purchased from Stratagene, Inc. Ribonuclease T1 and phage T4 RNA ligase were purchased from Pharmacia. Reverse transcriptase was purchased from Seikagaku America, Inc. Human debranching enzyme was provided by Dr. Mary Edmonds (University of Pittsburgh).

Transcription and Purification of RNA

RNA was transcribed with T3 RNA polymerase and purified as described in ref. 16. Unless otherwise indicated, the plasmids used to synthesize RNA were linearized in the downstream polylinker with Eco RI.

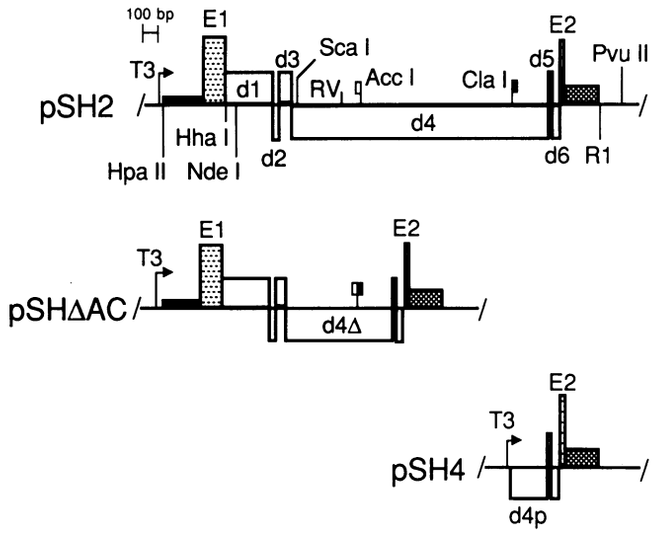

Figure 1: Diagram of portions of the *coxI* gene cloned in plasmids used in this study. Plasmid pSH2 contains an insert of 3124 bp of yeast mtDNA containing *coxI* intron 1 (aI1) plus the flanking sequences shown. Regions encoding portions of coxI protein are lightly shaded (E1 and E2); the six domains of intron 1 are indicated along with those restriction sites thatwere used in the studies reported here. The thick line between the T3 promoter and E1 represents part of the 5′ leader of the *coxI* gene included in this construction. The heavily shaded box following E2 is part of the *cox I* intron 2 reading frame. Plasmid pSHΔAC is deleted for most of intron domain 4 but is otherwise identical to pSH2. Plasmid pSH4 contains the extreme 3′ end of the insert of pSH2 including part of domain 4, domains 5 and 6, exon 2 and part of intron 2. The DNA sequence of this portion of the *coxI* gene from strain D273-10B is published (14) except for the first 95 bp of the insert of pSH2, immediately following the Hpa II site indicated; the sequence of that region was determined in this study and is as follows:

CCGGCCCGCC CCCGCGGGAC CCCTTTTTAA GAAGGAGTAT TCATTAATAA
TAAATATTAT AAATTCAACT ATTGTTATAT TTATAAATAG AATAATATAT

The sequence between the underlined residues is denoted as gap 10B in the sequence compilation of de Zamaroczy and Bernardi (15).

Splicing Reactions

For most studies reported here RNA was diluted in water and mixed with an equal volume of buffer stock solution to achieve a final concentration of 40 mM Tris-HCl (pH 7.6), 100 mM $MgCl_2$, and 1M NH_4Cl. Reactions were incubated at 40°C for the times indicated. Samples were then ethanol precipitated, the pellet vacuum dried and dissolved in 10 μl of water. An equal volume of loading buffer was added and samples loaded directly onto 3.5% polyacrylamide/8M urea gels, containing acrylamide and bis-acrylamide in a ratio of 39:1. In some experiments alternate salts, salt concentrations, reaction times, temperatures and gel conditions were used, as noted.

Analysis of the ends of intron RNAs

Primer extension was used to map the 5′ ends of IVS-LAR and IVS-BL from NH_4Cl reactions and of 5′ IVS F1 and F2 and IVS-BL* from KCl reactions as described in ref. 7. The oligonucleotide primer used for the full-length excised intron RNAs and the small KCl-specific intron fragments is complementary to nucleotides 340 to 361 of the Bonitz et al. sequence of this gene (14). The primers used to map the 5′ ends of the larger and smaller IVS-BL* species are complementary to nucleotides 560−583 and 798−821, respectively (14). The 3′ ends of 5′ IVS F1 and F2 were mapped by S1 nuclease protection (17). The Nde I end of the Nde 1 to Eco RV or ScaI fragment of pSH2 (see Figure 1) was filled in with α-^{32}P-dATP in the presence of unlabeled dTTP. It was hybridized with gel purified 5′ IVS F1 or F2 RNA and incubated with the amounts of S1 nuclease noted in Figure 5.

Other Methods

Debranching experiments were done as described earlier (7). RNA sequence analysis was done by the dideoxy chain termination method as described in ref. 18. Phage T4 RNA ligase was used to label 3′ ends of RNA molecules as described in ref. 19. T1 endonuclease analysis was conducted as described in ref. 16.

RESULTS

CoxI Intron 1 Self-splices *in vitro*

Plasmid pSH2 was constructed by cloning a fragment of yeast mtDNA containing aI1 and flanking exon sequences behind the T3 promoter of pBSM13+ (Figure 1). *In vitro* transcription of pSH2 linearized with Eco RI generates a 3151 nt precursor-mRNA analog containing a 413 nt 5′ exon (comprised of 27 nt of vector sequence, 216 nt of *coxI* gene untranslated leader sequence and all of *coxI* exon 1), the entire intron 1 (2448 nt), and a 290 nt 3′ exon (comprised of the entire 36 nt exon 2 coding sequence fused to 254 nt of intron 2). Our initial survey of reaction conditions known to be suitable for self-splicing of aI5γ and bI1 (7, 9, 16, 20) revealed that the full-length transcript of pSH2 self-splices actively under some, but not all of those conditions. It is inactive in low levels of Mg^{2+}, and is dramatically activated by addition of certain salts of monovalent cations.

Figure 2 shows the effects of increasing concentrations of salts on reaction extent and product array in the presence of 100 mM Mg^{2+} (pH = 7.6 and 40°C, see below). In these initial figures, product RNAs are labeled according to their identity defined in a later section. Panel A shows that 1M NH_4Cl stimulates the reaction maximally (lane 6) and that lower concentrations of that

salt yield substantially less reaction (lanes 3–5). Interestingly, some reaction is detected even at the lowest level of added salt, 50 mM, but in that case the reaction is a partial one (lane 2); the most obvious products are about 470 and 2700 nt long, shown below to be 5′ exon (5′ E) and linear intron plus 3′ exon (IVS-3′ E) RNAs, respectively. While some of those products are present at all levels of added NH_4Cl, it is clear that low levels of salt favor this partial reaction while higher levels yield mostly spliced exons (SE) plus excised intron (both lariat and linear [or broken], IVS-LAR and IVS-BL, respectively).

Added KCl also stimulates reactions of aI1 RNA (Figure 2B). Over half of the input RNA is converted to product in a 60 min incubation with 1 M KCl (lane 6). Some reaction is detected even at very low levels of KCl; as was seen using NH_4Cl, there is

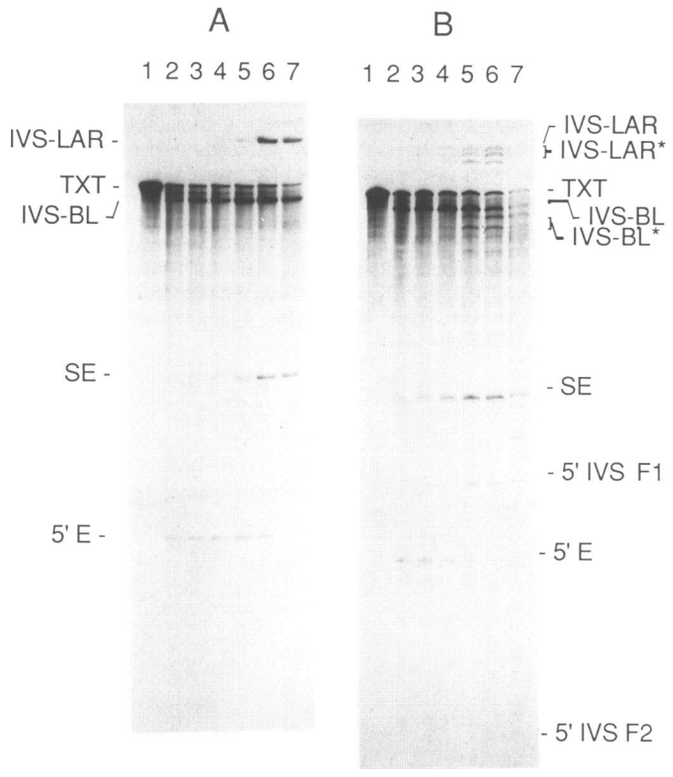

Figure 2. Self-splicing reactions of full-length transcripts of pSH2. A: Stimulation by added NH_4Cl. Radioactive transcripts were prepared from pSH2 linearized with Eco RI. Each sample contained about 10^5 cpm of transcript and was incubated at 40°C for 60 min. prior to electrophoresis. The reaction buffer contained 100 mM $MgCl_2$, 50 mM Tris-Cl (pH 7.6), and NH_4Cl at 50, 100, 250, 500mM, 1M or 2M in lanes 2–7, respectively. Lane 1 is unreacted precursor RNA. Bands in this and subsequent figures are labeled according to their sequence content: intron lariat (IVS-LAR); primary transcript (TXT); broken lariat/ linear intron (IVS-BL); spliced exons (SE); and 5′ exon (5′E). The linear intermediate containing intron plus 3′ exon (IVS-3′E) is not labeled but is located between TXT and IVS-BL and is especially evident at low levels of NH_4Cl. B: Stimulation by added KCl. The experiment was conducted as in panel A except that KCl was used in place of NH_4Cl. KCl concentrations are 50, 100, 250, 500mM, 1M or 2M in lanes 2–7, respectively. Lane 1 is unreacted precursor RNA. All of the products evident in panel A are present here along with some new products. The major condition-specific products are labeled here based on their identity as characterized in later sections of this paper (see below). Those novel products are IVS-LAR*, shortened lariat RNAs; IVS-BL*, linear or broken forms of IVS-LAR*; 5′ IVS F1 and F2, fragments of excised intron RNA containing the 5′ end of the intron and different 3′ ends. As in panel A, some IVS-3′E is evident between TXT and IVS-BL

a preference for the partial reaction yielding 5′E and linear IVS-3′E as the main products (lanes 2 and 3). At higher concentrations of KCl (lanes 4–6), spliced exons are clearly present along with linear (or broken lariat) intron RNA. And finally, lane 7 shows that very high levels of added KCl support splicing but also promote other cleavages of precursor and (probably) product RNAs. Interestingly, where KCl reaction extents are high (e.g., Fig. 2B, lanes 5 and 6) the array of products is more complex than that seen with comparable extents of reaction with NH_4Cl. A number of products clearly evident in lanes 5–7 of panel B are absent from analogous samples in panel A; those include slowly-migrating RNAs (lariats, see below) and long and short, apparently, linear RNAs.

Added 0.5 M $(NH_4)_2SO_4$ greatly enhances the splicing reactions of aI5γ and bI1 yielding spliced exons and excised intron lariat as the major products. At 50 and 100 mM $(NH_4)_2SO_4$, some 5′ exon RNA is released from the aI1 precursor RNA, apparently without associated lariat formation (not shown). At 250–1,000 mM salt, some splicing and lariat formation are detected and the maximum extent of reaction is obtained at 2 M salt. The rate and extent of reaction of aI1 RNA in 2 M $(NH_4)_2SO_4$ are much less than in 1 M NH_4Cl; but, both salts yield the same array of major products.

The temperature dependence of aI1 self-splicing at optimized salt concentrations was determined (not shown). Reactions are readily detected between 35° and 42°C in 1 M KCl and 1 M NH_4Cl while a clear optimum of 42°C was noted in 2 M $(NH_4)_2SO_4$ All three reaction conditions are essentially devoid of activity at 30°C. The pH is not a significant variable between values of 6.5 and 8.5 (not shown) and all of the reactions shown here were carried out at pH 7.6.

Characterization of the Major Products of NH_4Cl Reactions

Reactions in 100 mM Mg^{2+} supplemented with 1 M NH_4Cl yield the highest rates and extents of reactions with the simplest array of products. Since most of those products are also obtained in reactions containing the other two salts, we have characterized the major products of NH_4Cl reactions in some detail. Products of NH_4Cl reactions that contain exon sequences were first identified by probing blots of unlabeled reaction products with exon-specific oligonucleotides. Products identified as 5′ exon (5′ E) and spliced exons (SE) in Figure 2A hybridize to the 5′ exon-specific probe. The band identified as spliced exons also hybridizes to the 3′ exon probe. Neither 5′E nor SE hybridizes to the intron-specific oligonucleotide used. The analysis of the spliced exon product was completed by dideoxy sequencing across the splice junction (Figure 3A). An end-labeled 3′ exon-specific primer was annealed to purified SE RNA and extended using reverse transcriptase in the presence of unlabeled dNTPs and ddNTPs (see Materials and Methods). The cDNA sequence, 5′ TACTAAAACATTAAAT, is complementary to the splice site predicted by Michel et al. (2).

The product between TXT and IVS-BL in Figure 2 is identified as the reaction intermediate containing the intron and 3′ exon. It is about 450 nt smaller than the precursor and contains the 3′ end of the precursor RNA. It is the only intron-containing product that remains labeled when 3′ end-labeled precursor is reacted (not shown). In that experiment no product that contains only 3′ exon sequences (expected to be 290 nt long) was observed. The absence of free 3′ exon as a product was confirmed by experiments using precursor RNA with a 150 nt 3′ extension (using as template plasmid DNA linearized at the vector Pvu II

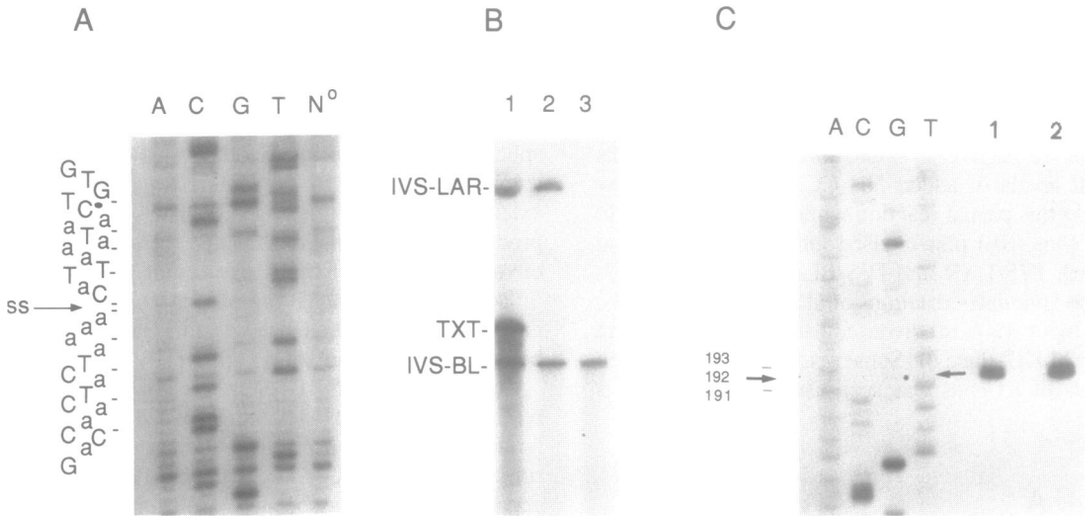

Figure 3: Characterization of products of NH_4Cl reactions. A: cDNA sequence of the *in vitro* splice site: The product RNA identified as spliced exons in NH_4Cl reactions was purified from a reaction mixture. A 5′ end-labeled oligodeoxyribonucleotide complementary to a sequence in the 3′ exon was annealed to the spliced exon RNA and extended with reverse transcriptase using unlabeled dNTPs and ddNTPs. The cDNA sequence across the splice junction is shown; the splice site is indicated by an arrow labeled SS. Ambiguous nucleotides are labeled with lower case letters while upper case letters are used for unambiguous residues. N^o is an extension reaction in the absence of ddNTPs. The sequence shown is the complement of the published exon sequence (14) with one exception: the C residue denoted by a dot is a T in strain D273; we have confirmed this silent third position change (CAA to CAG on the mRNA-like strand) by sequencing this region of plasmid pSH2. B. Debranching of the aI1 lariat RNA: IVS-LAR RNA was purified from an NH_4Cl reaction and mock treated (lane 2) or treated (lane 3) with human debranching enzyme. In the treated sample, IVS-LAR RNA was completely converted to a faster migrating RNA species which comigrates with IVS-BL. Lane 1 is an NH_4Cl reaction of aI1 precursor RNA. C: Definition of the 5′ end of excised intron RNA: A 5′-end-labeled intron-specific oligodeoxyribonucleotide complementary to nucleotides +340 to +361 of the *coxI* gene (14) was annealed to purified IVS-LAR (lane 1) and IVS-BL (lane 2) and extended with reverse transcriptase. Each reaction yields a prominent product 192 nt long based on the DNA sequencing ladder provided.

site, see Figure 1). The smallest RNA that shifted to lower mobility in that experiment was spliced exons (not shown).

Excised intron lariat (IVS-LAR) was identified by the results of several experiments. First, labeling the 3′ ends of unlabeled reaction products showed that that RNA has a free 3′-hydroxyl end and is, therefore, not a circle (not shown). Second, aI1 lariat RNA is sensitive to human debranching enzyme known to hydrolyze 2′-5′ phosphodiester bonds of lariat RNAs (21) including group II intron lariats (7, 16). As shown in Figure 3B, purified IVS-LAR comigrates with the slowest migrating product of the reaction (compare lanes 1 and 2); some apparently is broken during the purification to yield an RNA that comigrates with another major product of the reaction, identified as IVS-BL, for broken lariat plus linear intron. Treatment of purified IVS-LAR with human debranching enzyme converted it to a faster migrating RNA that comigrates with IVS-BL (compare lanes 2 and 3). In most debranching experiments of this sort, all of lariat RNA is converted to linear intron; however, that does not always result in an increased signal at the position of IVS-BL because broken lariat RNA is also debranched yielding linear products of various sizes (see ref. 7 for a full analysis of this issue). Given the specificity of the extract used, the junction appears to be a 2′-5′ phosphodiester bond similar to that found in excised nuclear introns and other group II introns.

The 5′ ends of IVS-LAR and IVS-BL obtained under this reaction condition were determined by extension of an intron-specific oligonucleotide with reverse transcriptase. Both IVS-LAR and IVS-BL yield the same 192 nt extension product (Figure 3C). The strong stop to reverse transcription is at the position of the first nucleotide of the intron showing that the 5′ ends of both forms of excised intron RNA result from cleavage at the 5′ splice

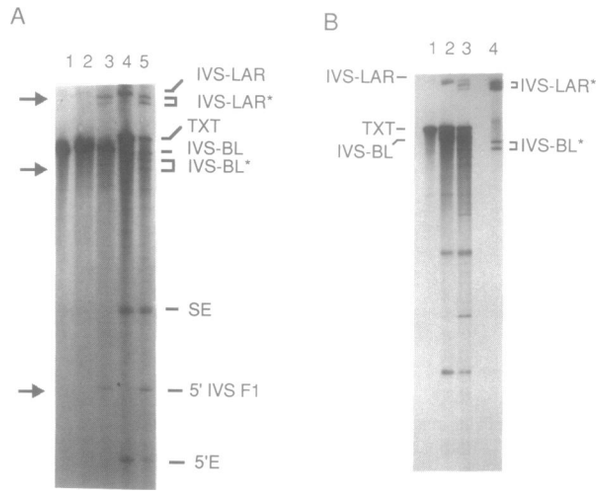

Figure 4. Identification of large novel products of KCl reactions. A. IVS-BL from NH_4Cl reactions is reactive in KCl buffer: Linear excised intron RNA was gel purified from products of an NH_4Cl reaction (lane 1). An aliquot of that material was incubated for 30 min in NH_4Cl containing reaction buffer (lane 2) and in KCl containing buffer (lane 3). Lanes 4 and 5 are precursor RNA incubated for 30 min. under NH_4Cl and KCl conditions, respectively. All of the major products including those specific to KCl reactions noted in Figure 2B are labeled. Note that IVS-BL yields the large novel lariat (IVS-LAR*) and linear (IVS-BL*) products and the short linear intron fragments (5′IVS F1) in KCl reactions but not in NH_4Cl ones; in this experiment 5′ IVS F2 ran off of the gel. B. Gel analysis of purified novel lariats from KCl reactions: Control samples are unreacted aI1 precursor RNA (lane 1), reactions products in NH_4Cl buffer (lane 2) and products in KCl buffer (lane 3). The doublet of lariat species typical of KCl reactions (see lane 3) was excised from a preparative gel, extracted and electrophoresed again (lane 4).

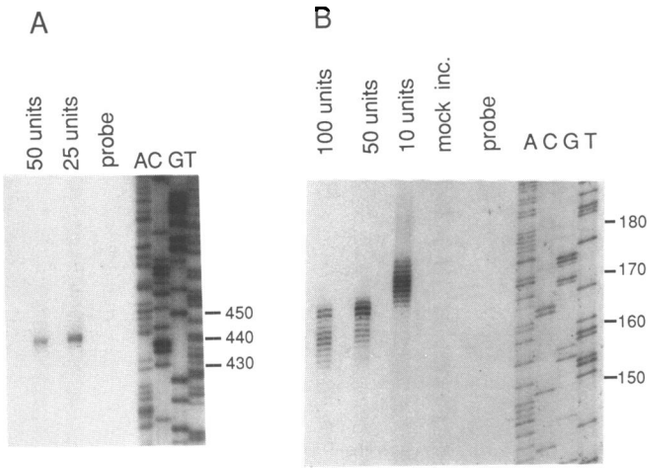

Figure 5: Characterization of KCl-specific branching sites: A. Mapping the 3′ end of 5′ IVS F1: 5′ IVS F1 RNA was gel purified from a KCl reaction, annealed with the end-labeled probe noted in Methods, treated with 0, 25 or 50 units of S1 nuclease, and the products fractionated alongside of a DNA sequencing ladder on a 5% polyacrylamide/8M urea gel. The sequencing ladder shows aI1 sequence starting from an oligonucleotide complementary to positions 798−821 (14). Major protected fragments of the probe 441−444 nt long are evident in both treated samples. The intron sequence where this cleavage occurs is indicated in Figure 6. B. Mapping the 3′ end of 5′ IVS F2: 5′ IVS F2 RNA was gel purified from a KCl reaction, annealed with the end-labeled probe noted in Methods, treated with 0, 10, 50 and 100 units of S1 nuclease and analyzed on an 8% polyacrylamide/8M urea gel. The same sequencing reactions as in panel A were used to calibrate this experiment. Protected fragments 171−180 nt long are present in the sample incubated with 10 units of S1 nuclease; higher amounts of nuclease yield somewhat smaller products probably due to overdigestion of this very A+T/U-rich hybrid region. The intron sequence where this cleavage occurs is indicated in Figure 6.

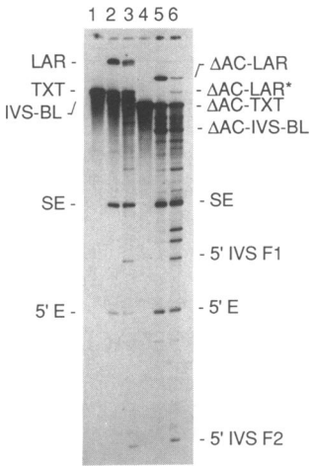

Figure 6: Location of novel intron cleavage sites: Panel A shows a portion of the RNA sequence of domain 1 of aI1 from position 413 to 441 based on the primary sequence of Bonitz et al. (14). That region comprises the D3(ii) helix and loop of domain 1, a region that contains the EBS1 sequence involved in 5′ splice site selection (see ref. 2). The S1 nuclease protection experiment of Figure 5B locates the 3′ end of 5′ IVS F2 to the region just preceeding EBS1. Primer extension data (not shown) indicate that the 5′ end of the larger IVS-BL* RNA (the other product of the post-splicing reaction in this region) is near the 5′ edge of EBS1. Panel B shows a portion of the RNA sequence of domain 4 of aI1 from position 680−700 (see ref. 14). This region of the intron plays no known role in the self-splicing of this intron and no model for the structure of this region has been proposed. The location of the 3′ end of 5′ IVS F1 was determined by the S1 nuclease protection experiment of Figure 5A. Primer extension data (not shown) indicate that the 5′ end of the smaller IVS-BL* RNA (the other product of the post-splicing reaction in this region) is near the end of 5′ IVS F1.

junction. Because of the very low mobility of the IVS-LAR RNA, it is inefficiently recovered from gel slices so that we were unable to obtain it in sufficient quantity to map its branch point. Construction of a shorter form of the intron that retains self-splicing activity (see below) largely solved this problem.

Novel Products of KCl Reactions

In this section, we summarize experiments defining the major products specific to the KCl reaction condition. In time course experiments in 1 M KCl, full-length excised intron and spliced exons are the predominant products at short reaction times. The novel products become prominent only at late times, where much of the precursor RNA has been converted to products. This indicates that the KCl-specific products probably result from secondary reactions. This inference is confirmed by the experiment of Figure 4A where excised linear intron from a short-term KCl reaction was purified (lane 1) and incubated for an additional 60 min in KCl and NH_4Cl containing buffers. That RNA is stable and unreactive in the NH_4Cl buffer (lane 2) but yields a number of products upon further incubation in KCl (lane 3). Lanes 4 and 5 are control reactions of precursor RNA in NH_4Cl and KCl conditions, respectively. Comparing lanes 3 and 5, it is clear that excised linear RNA yields all of the intron products specific to that reaction condition. This result leads us to conclude that those RNAs are products of post-splicing, secondary reactions of the intron RNA.

There are at least two slowly migrating products (denoted together as IVS-LAR* in Fig. 2B) both smaller than the lariat RNA formed in NH_4Cl reactions (Fig. 2B, lanes 5 and 6). They were found to be lariats by showing that they can be 3′ end-labeled by RNA ligase; they are not, however, substrates for debranching enzyme (not shown) suggesting some atypical feature at the branch. The other large novel products of this reaction condition, denoted IVS-BL* in Fig. 2B, were found to be linear or broken IVS-LAR* RNAs. As shown in Figure 4B, a portion of purified IVS-LAR* RNA comigrates with IVS-LAR* of primary reactions and the remainder migrates faster, comigrating with the IVS-BL* species. It is likely that some of the IVS-BL* in primary reactions is authentic linear, shortened intron RNA.

The small novel products specific to KCl reactions (5′IVS F1 and F2, see Fig. 2B and Fig. 7) hybridize with probes specific to the 5′ end of the intron (not shown). Their identity as 5′ fragments of excised intron RNA was confirmed by purifying each small RNA from a preparative gel and mapping each 5′ end by primer extension. Each yields the same 192 nt primer extension product obtained earlier using the large intron RNAs from NH_4Cl reactions (see Figure 3C); this indicates that the 5′ end of each is the first nucleotide of the intron (not shown). And finally, we mapped the 3′ ends of 5′IVS F1 and F2 by S1 nuclease protection (Figure 5A and B, respectively). The 3′ end of 5′IVS F2 is located in the D3(ii) element of domain 1 (see ref. 2), in a sequence adjacent to EBS1 (see Figure 6). The S1 nuclease yields a fairly broad signal, probably because the sequence of the hybrid there is very A+T/U rich. As is also noted in Figure 6, the 3′ end of 5′ IVS F1 is in a region of domain 4 of unknown secondary structure (see ref. 2). A preliminary primer extension experiment using unreacted RNA and RNA

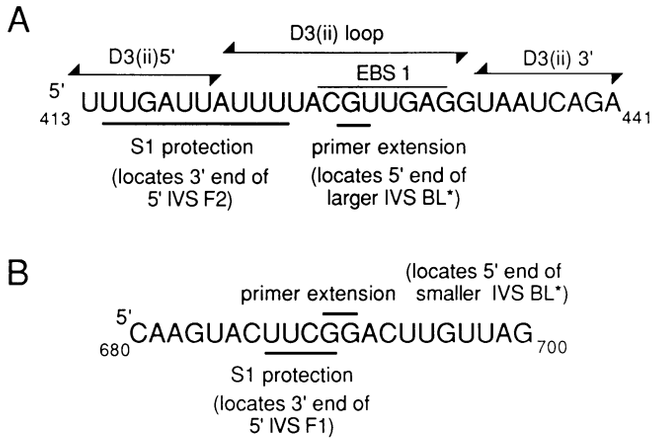

Figure 7: *In vitro* reactions of aI1 RNA deleted for most of domain 4. Radioactive transcripts from Eco RI cleaved pSH2 (lanes 1–3) and pSHDAC (lanes 4–6) were reacted at 40°C for 45 min. in reaction buffer containing 1 M NH_4Cl (lanes 2 and 5) or 1 M KCl (lanes 3 and 6). Unreacted precursor RNA is displayed in lanes 1 and 4. Products of pSHΔAC RNA reactions that do not comigrate with products of full-length intron reactions are labeled on the right-hand side of the figure with special notations (ΔAC -LAR, ΔAC-LAR*, etc.).

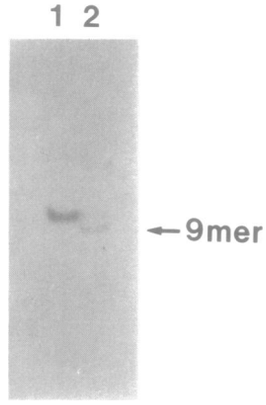

Figure 8: The branch site is near the 3′ end of the intron lariat. Unlabeled IVS-LAR and IVS-BL from an NH_4Cl reaction were purified from gel slices and 3′-end-labeled with pCp (with ^{32}P at the 5′ position). Each sample was digested to completion with RNase T1 and fractionated on a 25% polyacrylamide/8M urea gel (lanes 1 and 2, respectively).

reacted in NH_4Cl or KCl as templates located two KCl-specific stops to reverse transcription in the vicinity of the two sites defined by the S1 nuclease protection experiments (noted in Figure 6). Those strong stops to primer extension probably delineate the 5′ ends of the two major IVS BL* species. Further primer extension experiments of IVS-LAR* and/or IVS-BL* species were not pursued because those RNA species are difficult to isolate in quantity uncontaminated by other large RNA molecules.

Intact Intron Domain 4 is not Required for Self-Splicing of aI1

Jarrell et al. (22) showed that an intact domain 4 is not essential for the *in vitro* reactions of aI5γ; even a form of aI5γ completely deleted for domain 4 remains highly reactive in unimolecular self-splicing reactions. Domain 4 of aI1 is 1,872 nt long and contains most of the open reading frame that encodes the maturase protein required *in vivo*. *CoxI* intron 5g has a much shorter domain 4 (128 nt) and does not encode a maturase. We tested whether the reading frame within domain 4 of aI1 is dispensable *in vitro* even though it plays an essential role *in vivo*.

We deleted 1107 nt (60%) of domain 4 of aI1 by cleaving plasmid pSH2 at its unique Acc I and Cla I sites and religating the complementary ends to form plasmid pSHDAC (see Figure 1). As shown in Figure 7, lanes 4–6, full-length transcript of that plasmid was found to be highly reactive relative to the control substrate (compare lane 2 to 5 and lane 3 to 6) yielding spliced exons and intron RNAs analogous to those obtained with the full-length intron. The presence of spliced exons was confirmed by sequencing across the splice junction as in Figure 3A (not shown). The product of the NH_4Cl reaction specific to this shorter construct that migrates more slowly than the deleted precursor (see lane 5) was confirmed to be intron lariat by a debranching experiment (not shown).

Jarrell et al. also demonstrated that 'half molecules' of aI5γ, interrupted within or deleted for part of domain 4 *trans*-splice *in vitro* (22). We have confirmed that analogous half molecules of aI1 precursor readily *trans*-splice in KCl and NH_4Cl reactions (not shown, see ref. 23). For those experiments one new plasmid, pSH4 (see Figure 1) was made as template for 3′ half molecules; 5′ half molecules were obtained from pSH2 linearized at Sca I, Acc I or Cla I sites. None of the partial precursor RNAs was reactive under these conditions when incubated alone, and all three 5′ molecules (deleted for 1584, 1107 and 0 nt of domain 4, respectively) were comparably reactive in bimolecular reactions with the transcript of pSH4. Thus, deletion of nearly 85% of domain 4 (1584 of 1872 nt) fails to block such *trans* reactions of this intron.

Location of the Branch Site in aI1 Lariat RNA

In both group II introns characterized previously, the branch point adenosine is located very near the 3′ end of the intron in domain 6 (8, 9). The presumed branch site of intron 1 is an adenosine residue 7 nt from the 3′ end of the intron (5′ GCUA*UUUCAU 3′) (see ref. 2). That adenosine is contained in a 9 nt long RNAse T1 product that also contains the 3′ end of the excised intron.

The shortened ΔAC-LAR RNA was found to be retrieved from gel slices more efficiently than the full-length lariat, apparently due to its smaller size. This finding permitted isolation of a sufficient amount of it to locate the branch site. Unlabeled shortened precursor RNA (transcribed from pSHΔAC, see Fig. 1) was reacted in 1 M NH_4Cl, the products 3′ end-labeled, fractionated on a polyacrylamide gel and IVS-LAR and IVS-BL RNAs extracted from gel slices. Those RNAs were digested to completion with RNAse T1 and the products analyzed on a 25% polyacrylamide gel with 5′ end-labeled oligodeoxyribonucleotides as rough size standards (Figure 8). IVS-BL yields mostly a fragment migrating with the expected mobility (9 nt) and a small amount of a second fragment with slower mobility (lane 2). IVS-LAR yields a single fragment comigrating with the minor product from IVS-BL (lane 1). We interpret these data to mean that most of the IVS-BL of aI1 is unbranched and that the branch point in IVS-LAR is present in the last 9 nt of the intron (near or at the predicted site).

DISCUSSION

In this paper we report that the first intron of the *coxI* gene of yeast mtDNA self-splices *in vitro* to yield excised intron lariat and accurately spliced exons. It is the fifth example of self-splicing

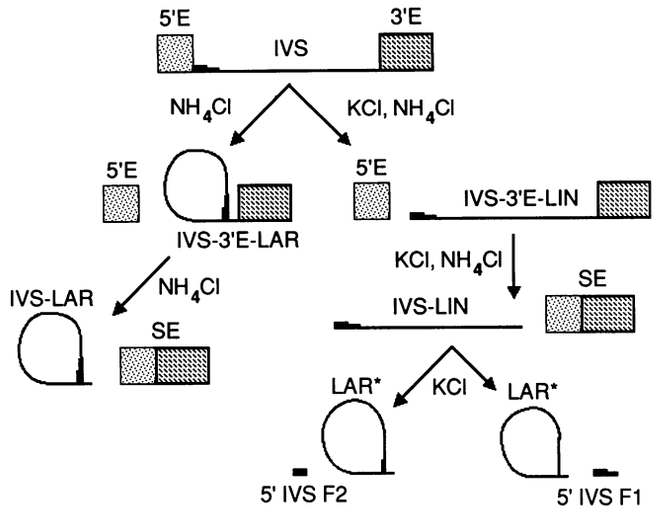

Figure 9: Diagram of pathways leading to the main products of reactions of aI1 precursor RNA. The top line is a diagram of the precursor RNA used in these studies. The 5′ and 3′ exons are shaded differently, as labeled. Two segments of the 5′ end of the intron are represented by lines of different thickness to indicate sites within the intron at which specific reactions occur in KCl stimulated reactions. Each arrow indicates a distinct reaction step and each arrow is labeled to indicate the reaction condition(s) where that step occurs efficiently enough for its products to be prominent on gels. There are some other products of KCl reactions, especially evident in reactions using the ΔAC construction (see Fig. 7), that were not characterized in this study. The intermediates of the first reaction steps (IVS-3′E-LAR and IVS-3′E-LIN) are not labeled on the gels shown, but are present as minor products in some of them. Linear excised intron RNA is labeled IVS-LIN here but is denoted IVS-BL in the figures since that band contains both linear intron RNA and broken lariat in actual experiments. The post-splicing reactions specific for KCl reactions shown at the bottom right hand side of the figure as transesterification (branching) reactions probably also occur as hydrolysis reactions (yielding LIN* RNAs smaller than full-length excised intron RNA).

group II intron and is only the second group IIA intron shown to self-splice (7−9, 24, 25). It is the first maturase-encoding intron of either group I or II shown to self-splice *in vitro*. While a number of self-splicing group I introns contain reading frames (e.g., 26−30) none of them has been shown to encode a maturase *in vivo*.

The *in vitro* splicing conditions for aI1 are somewhat different from those optimal for the self-splicing group IIB introns, aI5γ and bI1. Specifically, the temperature optimum for aI1 is around 40°C while aI5γ and bI1 have an optimum of 45°C. In the presence of 2 mM spermidine, aI5γ and bI1 are reactive throughout the range of 10−100 mM Mg^{2+}; aI1 is inactive in 10 mM Mg^{2+} (plus spermidine) and has only slight activity in 100 mM. Certain salts added to 100 mM Mg^{2+} stimulate reactions of aI5γ and bI1 and added salts are quite essential for ready detection of self-splicing reactions of aI1. The optimum salt requirement for aI1 reactions is 2−4-fold higher than for aI5γ and bI1 (depending on the salt). Also, at lower salt concentrations aI1 carries out the first reaction step but appears blocked for the second step. While aI5γ and bI1 are very reactive in the presence of 0.5 M $(NH_4)_2SO_4$, aI1 RNA is not as reactive in the presence of that salt, even at the highest concentration tested (2 M).

Under KCl reaction conditions, aI5γ and bI1 splice with little branch formation and the splicing products engage in efficient post-splicing reactions. One such reaction, spliced exon reopening (SER), accounts for the nearly quantitative conversion of the spliced exon product into separate 5′ and 3′ exon RNAs (16, 20). Our studies of aI1 RNAs failed to identify a 290 nt long 3′ exon product even in KCl reactions, suggesting that aI1 RNA does not carry out the SER reaction. Direct attempts to detect that activity using reconstruction experiments (such as used for aI5γ in ref. 16) detected no cleavage products of spliced exon RNA. Potassium chloride reactions of aI1 RNA yield a number of products in addition to spliced exons and excised intron; the most abundant ones result from post-splicing reactions in which the linear excised intron RNA yields lariats at cryptic sites with associated release of linear fragments of the 5′ end of the intron.

The reactions of the aI1 precursor RNA characterized in this paper are summarized diagramatically in Figure 9. In NH_4Cl stimulated reactions, most molecules follow the pathway on the left half of the figure, leading to spliced exons and excised intron lariat. For a fraction of molecules the first reaction step occurs without branching (see Fig. 8) so that the spliced exon product is obtained together with linear excised intron RNA. In KCl stimulated reactions, most molecules follow the reaction path on the right side of the figure. A key difference between the two reaction conditions involves the fate of excised linear intron RNA: in NH_4Cl, that product is stable and accumulates while in KCl it is reactive yielding post-splicing intron RNA products. The linear intron RNA engages in intramolecular branching reactions at two major sites which were localized here; those reactions yield shorter lariat RNAs (IVS-LAR*) and short linear RNAs containing the 5′ end of the full-length excised intron RNA (5′IVS F1 and F2). These cryptic branching sites of this intron are highly reactive in KCl reaction buffers. The two most prominent sites of reaction were mapped, with main emphasis on defining the ends of the 5′ intron fragments; those sites do not resemble each other in any obvious way nor do they resemble other functionally important sequences of this intron (e.g., IBS1, EBS1, 5′ or 3′ ends of the intron). No analogous side reactions of this sort have been noted for the wild-type forms of any of the other self-splicing group II introns.

All group II introns have a highly conserved secondary structure consisting of a central wheel with six helical spokes ('domains', numbered 1 through 6, 5′- 3′) (2). Intron domains 5 and 6 are highly conserved while intron domain 4 is the most variable (as small as 25 nt in bI1 or as large as 1872 nt in aI1). In every case where an intron of this group contains a reading frame, most of it is 'inserted' in domain 4 (2). It has been demonstrated that all of domain 4 of aI5γ (which has no ORF) can be deleted without blocking *in vitro* splicing (22). The present study extends that finding to a group IIA intron: large portions of the ORF (up to nearly 85% of domain 4) of aI1 were deleted with little effect on *trans*-splicing; also, a *cis* deletion of 60% of domain 4 was not inhibitory.

We have shown that the *in vivo* dependence of the aI1 intron on its own maturase does not preclude its self-splicing *in vitro*. The question then arises: what is the role of the maturase *in vivo*? Garriga and Lambowitz have shown that the *Neurospora* mitochondrial large rRNA gene group I intron undergoes splicing *in vivo* by essentially the same trans-esterification reaction as *in vitro* (31). That intron, however, does not self-splice under *in vitro* conditions suitable for other *Neurospora* group I introns (26) and requires at least one nuclear encoded protein to facilitate formation of the productive secondary structure (32). Perhaps more analogous to our situation, *Neurospora cob*I1 (26) and yeast *cob*I5 (33, 34) both self-splice but depend on a nuclear gene for their splicing *in vivo* (35, 36); it is unknown whether those nuclear

encoded proteins provide functions analogous to maturase proteins.

It is likely for both group I and group II introns that self-splicing is the fundamental splicing mechanism *in vivo* and that *trans*-acting factors required *in vivo* facilitate those reactions by inducing or stabilizing a highly reactive structure, thereby increasing the reaction rate. The need for high concentrations of salt (100 mM magnesium ion plus potassium or ammonium ion at 50 mM or higher) for efficient *in vitro* self-splicing of yeast group II introns is probably due in part to a requirement for electrostatic screening. In the absence of salts, the high negative charge of the RNA may destabilize some secondary or tertiary structure crucial for the self-splicing reaction. These salts may mimic functions provided by maturases and other proteins that are involved *in vivo*. It remains unexplained, however, why some salts (e.g., LiCl; see ref. 16) do not stimulate group II intron splicing and others (e.g., KCl) yield salt-specific side reactions.

ACKNOWLEDGEMENTS

We thank Craig Peebles, Kevin Jarrell, Rosemary Dietrich, and Sulayman Dib-Hajj for valuable comments and advice during these studies. We thank Mary Edmonds (University of Pittsburgh) for the generous donation of purified human debranching enzyme. This research was supported by a grant (GM31480) from the National Institutes of Health. S.M.B. is a Fellow of the Robert A.Welch Foundation (Grant I-1211).

REFERENCES

1. Cech, T. R. (1988) Gene **73**, 259–271.
2. Michel, F., Umesono, K., and Ozeki,H. (1989) Gene **82**, 5–30.
3. Cech, T. R. (1990) Ann. Rev. Biochem. **59**, 543–568.
4. Perlman, P. S., Peebles, C. L., and Daniels, C. (1990) IN: *Intervening Sequences in Evolution and Development* (E. M. Stone and R. J. Schwartz, eds.). Oxford University Press, New York, pp. 112–161.
5. Perlman, P. S. (1990) Methods Enzymol. **181**, 539–557.
6. Lambowitz, A. M. and Perlman, P. S. (1990) TIBS **15** 440–444.
7. Peebles, C. L., Perlman, P. S., Mecklenburg, K. L. , Petrillo, M. L., Tabor, J. H., Jarrell, K. A., and Cheng, H. L. (1986) Cell **44**, 213–223.
8. van der Veen, R., Arnberg, A. C. , van der Horst, G., Bonen, L., Tabak, H. R., and Grivell, L. A. (1986) Cell **44**, 225–234.
9. Schmelzer, C. and Schweyen, R. J. (1986) Cell **46**, 537–565.
10. Carignani, G., Groudinsky, O., Frezza, D., Schiavon, E., Bergantino, E., and Slonimski, P. P. (1983) Cell **35**, 733–742.
11. Carignani, G., Netter, P., Bergantino, E., and Robineau, S. (1986) Curr. Genet. **11**, 55–63.
12. Mecklenburg, K. L. (1987) Ph.D. Thesis. The Ohio State University, Columbus, Ohio.
13. Hudspeth, M., Shumard, D., Tati, K., and Grossman, L. I. (1980) Biochem. Biophys. Acta **610**, 221–228.
14. Bonitz, S. G., Coruzzi, G, Thalenfeld, B. E., and Tzagoloff, A. (1980) J. Biol. Chem. **255**, 11927–11941.
15. de Zamaroczy, M. and Bernardi, G. (1986) Gene **47**, 155–177.
16. Jarrell, K. A., Peebles, C. L., Dietrich, R. C., Romiti, S. L., and Perlman, P. S. (1988) J. Biol. Chem. **263**, 3432–3439.
17. Sharp, P. A., Berk, A. J., and Berget, S. (1980) Methods Enzymology. **65**, 750–768.
18. Lane, D. J., Pace, B., Olsen, G. J., Stahl, D. A., Sogin, M. L., and Pace, N. R. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 6955–6959.
19. England, T. E., Bruce, A. G., and Uhlenbeck, O. C. (1980) Methods Enzymol. **65**, 65–74.
20. Muller, M. W., Schweyen,R. J., and Schmelzer, C. (1988) Nucleic Acids Res. **16**, 7383–7395.
21. Ruskin, B. and Green, M. R. (1985) Science **229**, 135–140.
22. Jarrell, K. A., Dietrich, R. C., and Perlman, P. S. (1988) Mol. Cell. Biol. **8**, 2361–2366.
23. Hebbar, S. K. (1989) Ph.D. Thesis, The Ohio State University, Columbus, Ohio.
24. Kuck, U., Godehardt, I., and Schmidt, U. (1990) Nucleid Acids Res. **18**, 2691–2697.
25. Schmidt, U., Riederer, B., Morl, M., Schmelzer, C., and Stahl, U. (1990) EMBO J. **9**, 2289–2298.
26. Garriga, G. and Lambowitz, A. M. (1984) Cell **39**, 631–641.
27. van der Horst, G. and Tabak, H. R. (1985) Cell **40**, 759–766.
28. Gott, J. M., Shub, D. A., and Belfort, M. (1986) Cell **47**, 81–87.
29. Tabak, H. R., van der Horst, G., Kamps, A. M. J. E., and Arnberg, A. C. (1987) Cell **48**, 101–110.
30. Muscarella, D. and Vogt, V. (1989) Cell **56**, 443–454.
31. Garriga, G. and Lambowitz, A. M. (1983) J. Biol. Chem. **258**, 14745–14748.
32. Garriga, G. and Lambowitz, A. M. (1986) Cell **46**, 669–680.
33. Gampel, A. and Tzagoloff, A. (1987) Mol. Cell. Biol. **7**, 2545–2551.
34. Partono, S. and Lewin, A. (1988) Mol. Cell. Biol. **18**, 2562–2571.
35. McGraw, P. and Tzagoloff, A. (1983) J. Biol. Chem. **258**, 9459–9468.
36. Gampel, A., Nishikimi, M., and Tzagoloff, A. (1989) Mol. Cell. Biol. **9**, 5424–5433.