The EMBO Journal Vol.18 No.4 pp.1003–1013, 1999

In vivo expression of the nucleolar group I intron-encoded I-*Dir*I homing endonuclease involves the removal of a spliceosomal intron

Anna Vader[1], Henrik Nielsen[2] and Steinar Johansen

Department of Molecular Cell Biology, Institute of Medical Biology, University of Tromsø, N-9037 Tromsø, Norway and [2]Department of Medical Biochemistry and Genetics, Panum Institute, Blegdamsvej 3C, DK-2200 Copenhagen N, Denmark

[1]Corresponding author
e-mail: annav@fagmed.uit.no

The *Didymium iridis* DiSSU1 intron is located in the nuclear SSU rDNA and has an unusual twin-ribozyme organization. One of the ribozymes (DiGIR2) catalyses intron excision and exon ligation. The other ribozyme (DiGIR1), which along with the endonuclease-encoding I-*Dir*I open reading frame (ORF) is inserted in DiGIR2, carries out hydrolysis at internal processing sites (IPS1 and IPS2) located at its 3′ end. Examination of the *in vivo* expression of DiSSU1 shows that after excision, DiSSU1 is matured further into the I-*Dir*I mRNA by internal DiGIR1-catalysed cleavage upstream of the ORF 5′ end, as well as truncation and polyadenylation downstream of the ORF 3′ end. A spliceosomal intron, the first to be reported within a group I intron and the rDNA, is removed before the I-*Dir*I mRNA associates with the polysomes. Taken together, our results imply that DiSSU1 uses a unique combination of intron-supplied ribozyme activity and adaptation to the general RNA polymerase II pathway of mRNA expression to allow a protein to be produced from the RNA polymerase I-transcribed rDNA.

Keywords: *Didymium iridis*/group I intron/homing endonuclease/mRNA processing/ribozyme

Introduction

Many group I introns contain open reading frames (ORFs) in their peripheral loop regions. Most commonly, the ORFs encode sequence-specific endonucleases that are involved in intron mobility, but they can also encode maturases that aid in intron splicing or, in the case of a few mitochondrial introns, structural genes (see Johansen *et al*., 1997). Whereas the great majority of the mitochondrial and phage introns contain ORFs, very few examples exist among the nuclear group I introns. PpLSU3 and DiSSU1 from the myxomycetes *Physarum polycephalum* and *Didymium iridis* encode the homing endonucleases I-*Ppo*I and I-*Dir*I, respectively (Muscarella and Vogt, 1989; Johansen and Vogt, 1994; Johansen *et al*., 1997). Site-specific endonucleases are also encoded by the NaSSU1 introns from several species of the amoeba flagellate *Naegleria* (Elde *et al*., 1999). Finally, an ORF with an unknown function has been described recently within a group I intron in the nuclear SSU of the archiascomycetous fungus *Protomyces pachydermus* (Nishida *et al*., 1998).

Compared with the organellar and phage group I intron ORFs, the nuclear ORFs face two additional challenges. First, the eukaryotic nucleus contains three highly specialized RNA polymerases. As a rule, protein-coding genes are only transcribed by RNA polymerase (pol) II. Enzymes associated with the pol II complex provide the mRNA with a 7-methyl guanosine (m^7G) 5′ cap, which is critical for its stability and important at the initiation of translation (reviewed in Lewis *et al*., 1995). All nuclear introns are found in the pol I-transcribed rRNA genes. Secondly, in order to be translated, the eukaryotic mRNA has to be transported from its place of synthesis in the nucleus to the cytoplasm. The 5′ cap has been implicated in this process as well. It is thus a paradox how an endonuclease can be expressed from an ORF encoded by a group I intron residing in the pol I-transcribed rDNA.

The only nuclear group I intron for which *in vivo* expression has been studied extensively is the prototype group I intron TtLSU1 from *Tetrahymena*. The intron is inserted in the LSU rDNA and does not contain an ORF. In isolated nuclei of *Tetrahymena thermophila*, accumulation of TtLSU1 was shown to be α-amanitin resistant, suggesting that the intron RNA is indeed a product of pol I transcription (Zaug and Cech, 1980). Furthermore, it was demonstrated that most pre-RNA molecules are spliced before transcription is terminated (Din *et al*., 1979), and that the excised intron was degraded rapidly in the nucleus (Brehm and Cech, 1983). The PpLSU3 intron from *P.polycephalum* has a very high degree of sequence similarity to TtLSU1 and is inserted at the same position in the LSU rDNA. However, PpLSU3 contains an ORF encoding the homing endonuclease I-*Ppo*I which renders the intron mobile at the DNA level (Muscarella and Vogt, 1989). After excision of the PpLSU3 RNA *in vitro*, the ORF is split from the catalytic part of the intron by autocatalytic cleavage at an internal processing site (Ruoff *et al*., 1992). The free intron as well as the two intron halves are also observed when whole-cell RNA is studied by Northern blot analyses (Ellison, 1994; Rocheleau and Woodson, 1995).

We have reported previously that the *D.iridis* DiSSU1 intron encodes the I-*Dir*I endonuclease and is mobile in genetic crosses between intron-containing and intron-lacking strains (Johansen *et al*., 1997). DiSSU1 is located in the SSU rDNA and has an unusual twin-ribozyme organization (Johansen and Vogt, 1994; Decatur *et al*., 1995). One of the ribozymes (DiGIR2) catalyses intron excision and circularization as well as exon ligation. The other ribozyme (DiGIR1), which along with the ORF is inserted in DiGIR2, carries out hydrolysis at internal processing sites (IPS1 and IPS2) located at its 3′ end (Decatur *et al*., 1995; Einvik *et al*., 1998). To address the

question of how a protein can be expressed from a group I intron located in the nuclear rDNA, we have examined the *in vivo* expression of DiSSU1.

Results

In vivo intron processing

We have reported previously that the 1436 nucleotide DiSSU1 intron self-splices *in vitro*, yielding an excised intron and ligated rRNA exons (Johansen and Vogt, 1994). Other reactions catalysed by the intron include site-specific hydrolysis at its 3′ splice site, the formation of a full-length circle and site-specific cleavage by DiGIR1 at the internal processing sites, IPS1 and IPS2 (Decatur *et al*., 1995; Einvik *et al*., 1998). A processing pathway has been proposed where the excised intron circularizes by G-exchange, re-opens by 3′ hydrolysis and finally is cleaved at IPS1/2 by DiGIR1. The result is a shorter ORF-containing RNA species that has been hypothesized to be the I-*Dir*I mRNA (Decatur *et al*., 1995).

To identify the *in vivo* pattern of pre-rRNA splicing and DiSSU1 processing, whole-cell RNA was isolated from the intron-containing *D.iridis* strain Lat3-5. RNA from the intron-lacking strain S7 was analysed in parallel. Northern analysis with a probe complementary to the ORF revealed six bands corresponding to ~10, 7.4, 3.9, 1.4, 1.2 and 0.9 kb (Figure 1B). It is important to note that the relative signal intensities of the bands varied greatly according to the physiological environment and life stage of the organism at the time of RNA isolation (A.Vader, unpublished results). We believe the 10 kb band to result from the rRNA precursor since it also hybridizes to an SSU rDNA probe (data not shown). The strong band corresponding to 7.4 kb appears to contain only intron sequences as it did not hybridize to the SSU probe (data not shown). DiSSU1 produces full-length circles *in vitro* (Johansen and Vogt, 1994). We therefore believe the 7.4 kb band to represent a circular form of the intron which has been retarded in the gel. Direct sequencing of an RT–PCR product obtained using primers flanking the circularization site showed that full-length circles indeed are present *in vivo* (A.Vader and S.Johansen, unpublished results). The 3.9 kb band was observed frequently in whole-cell and especially nuclear RNA. We do not know the identity of this RNA species (see below).

DiSSU1 produces a complex RNA processing pattern *in vitro* (Johansen and Vogt, 1994; Decatur *et al*., 1995). In order to identify the smaller RNA species observed *in vivo*, Lat3-5 RNA was separated on a 4% denaturing polyacrylamide gel, blotted and hybridized to DNA probes complementary to DiGIR1, ORF and DiGIR2 sequences, respectively (Figure 1C). Furthermore, the 5′ ends of the RNA species were determined by primer extension. The 1.4 kb band hybridized to all probes, identifying it as the full-length excised intron (RNA1). In accordance with the splicing mechanism of group I introns (see Cech and Herschlag, 1996), a non-coded nucleotide is present at the 5′ end of the intron (Figure 2A). We have reported previously that *in vitro* DiGIR1 cleaves at two processing sites in an obligate sequential order (Einvik *et al*., 1998). The sites, IPS1 and IPS2, are located just 3′ of DiGIR1′s catalytic core and only three nucleotides apart. The band at 1.2 kb hybridized to the ORF and DiGIR2 probes but not to the DiGIR1 probe, and was thus identified as the 3′ product of DiGIR1 hydrolysis (RNA2). The expected complementary product, the 0.24 kb DiGIR1 fragment, was not observed *in vivo*, suggesting that it is degraded shortly after excision. *In vitro*, DiGIR1 cleavage is dependent on high salt concentrations, and both IPS1 and IPS2 can be detected by primer extension analyses (Einvik *et al*., 1998). *In vivo*, on the other hand, only IPS2 was seen (Figure 2B). One explanation for this is that cellular factors facilitate DiGIR1 activity in such a way that IPS2 cleavage immediately follows IPS1 hydrolysis. Alternatively, perhaps only IPS2 is used *in vivo*. This latter explanation seems less likely, as it was shown *in vitro* that no IPS2 cleavage would take place if IPS1 was inactivated by the mutation of surrounding nucleotides (see Einvik *et al*., 1998).

The 0.9 kb band has not been observed *in vitro*. This RNA species, termed RNA3, hybridized only to the ORF probe, and thus seems to arise by RNA2 being cleaved 3′ of the ORF. The fact that the DiGIR2 probe detected a 0.3 kb band supports the conclusion that an additional internal processing site, IPS3, exists. Primer extension placed IPS3 55 nucleotides downstream of the ORF 3′ end (position 1129, Figure 2C). The fact that no RNA species containing only DiGIR1 and an ORF was found by Northern blotting suggests that IPS1/2 hydrolysis takes place prior to IPS3 cleavage.

Taken together, the results imply that *in vivo* processing resembles the *in vitro* pattern in that after splicing, the excised intron may be cleaved at IPS1/2 and that full-length intron circles also are formed in the cell. However, we did not detect any full-length intron without the external G (see Figure 2A), as would be expected if the circles re-open by site-specific hydrolysis. Thus, either these molecules are removed immediately by DiGIR1 cleavage at IPS1/2, or the circle is not part of the I-*Dir*I mRNA formation pathway. The *in vivo* processing pathway also differs from the *in vitro* situation in that RNA2 is cleaved further at IPS3, an additional *in vivo* processing site located downstream of the ORF. The result is an RNA species containing little more than the I-*Dir*I ORF sequence.

The I-DirI ORF contains a 51 nucleotide spliceosomal intron

RT–PCR of Lat3-5 cellular RNA with ORF-specific primers consistently generated a small PCR product in addition to the one expected from the DNA sequence (see Figure 5B, whole-cell RNA). Sequencing of the smaller product showed that a 51 nucleotide sequence is absent within the 5′ half of the ORF (position 502–552, Figure 3A). The elimination of this sequence is expected to produce a 17 amino acid shorter version of the I-*Dir*I protein, but not to alter the reading frame (Figure 3B). The removed sequence starts with GT and ends with AG, consistent with the consensus sequence of the GT–AG class of spliceosomal introns. Closer inspection revealed that the sequence similarities extend further into the intron as well as to a putative intron branch site. No mRNAs have been characterized previously from *D.iridis*, but sequence information from the related species *P.polycephalum* suggests that small (45–100 nucleotides) spliceosomal introns are common in the myxomycetes.

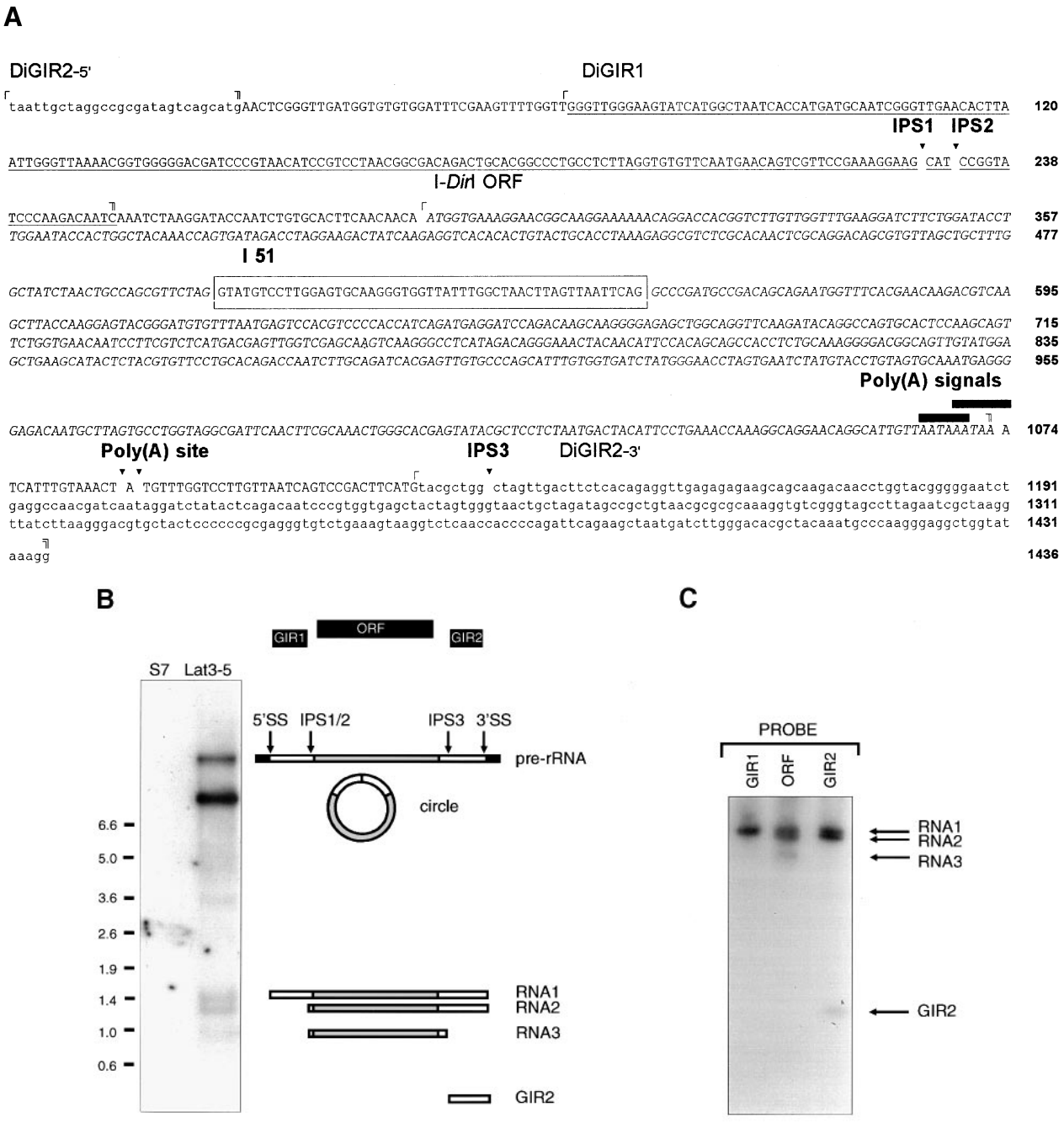

Fig. 1. Characterization of DiSSU1 RNA from *D.iridis* cells. (**A**) Overview of the DiSSU1 intron. The DiGIR2 ribozyme sequence and its insertion are shown in lower and upper case, respectively. The DiGIR1 ribozyme sequence is underlined, the I-*Dir*I ORF sequence is in italics, and the I51 spliceosomal intron is boxed. The internal processing sites IPS1, IPS2 and IPS3, as well as the polyadenylation site are denoted by arrowheads, and the poly(A) signals are marked with thick black lines. (**B**) Northern blot of Lat3-5 and S7 whole-cell RNA. A 4 μg aliquot of RNA from the intron-containing (Lat3-5) and intron-lacking (S7) strains was run on a 1% denaturing agarose gel and analysed by Northern analysis using the ORF probe. The positions of the ORF as well as the GIR1 and GIR2 probes are indicated by thick black lines. An overview of the observed intron RNA species is shown on the right. Exon, ORF and intron sequences are indicated in black, grey and white, respectively. The position of the 5′ and 3′ splice sites (SS) as well as the internal processing sites (IPS) are indicated. The size indications are derived from the G319 RNA marker (Promega). (**C**) Northern blot of Lat3-5 whole-cell RNA. A 4 μg aliquot of RNA from Lat3-5 was run on a 4% denaturing polyacrylamide gel and analysed by Northern blotting. The positions of the GIR1, ORF and GIR2 probes are shown in (B). S7 whole-cell RNA, which was run in parallel, did not hybridize to any of the probes (data not shown).

Thus, we infer that the excised 51 nucleotide RNA, I51, is a spliceosomal intron. It is the first reported example of a spliceosomal intron embedded in a group I intron as well as the first example of a spliceosomal intron in the pol I-transcribed rDNA.

To investigate the order of the processing events, RT–PCR was carried out with a primer located immediately downstream of the I51 intron, combined with different upstream primers. While upstream primers complementary to DiGIR1 produced solely a +I51 type of PCR product, primers complementary to the ORF led to the amplification of both intron-containing and intron-lacking DNAs (data not shown). The simplest explanation for these results is that IPS1/2 hydrolysis precedes I51 removal. However, a

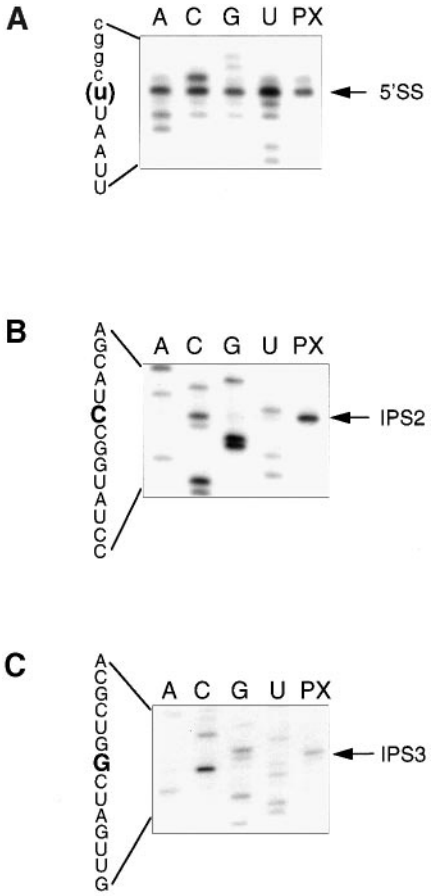

Fig. 2. Primer extension analyses of Lat3-5 whole-cell RNA. The 5′ splice site (5′SS) (**A**), and the internal processing sites IPS2 (**B**) and IPS3 (**C**) were identified using primers OP3, OP12 and OP178, respectively. The position of IPS3 was confirmed by a second primer extension reaction using OP 236 (data not shown). *In vitro* transcribed DiSSU1 RNA was used as a template to obtain RNA sequencing ladders in all primer extension reactions. The reactions were analysed on an 8% denaturing polyacrylamide gel. A, C, G and U denote the RNA sequence, i.e. reactions terminated with ddT, ddG, ddC and ddA, respectively. PX denotes primer extension of cellular RNA.

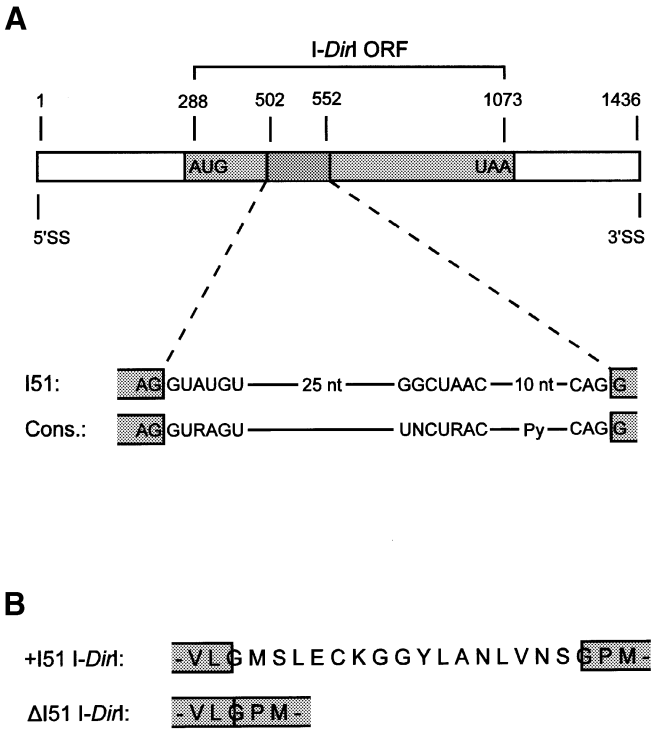

Fig. 3. The I51 spliceosomal intron. (**A**) Schematic presentation of the I51 spliceosomal intron. I-*Dir*I ORF sequences are shaded. Numbers are relative to the start of the DiSSU1 intron. The AUG and UAA translation initiation and termination codons of the I-*Dir*I ORF are indicated. 5′SS and 3′SS denote the 5′ and 3′ splice sites, respectively. Parts of the I51 intron sequence are aligned with the consensus sequence of the mammalian GT–AG spliceosomal introns (see Kreivi and Lamond, 1996). The numbers between the aligned sequences indicate the number of nucleotides. (**B**) Reading frame of the I-*Dir*I protein with (+I51 I-*Dir*I) and without (ΔI51 I-*Dir*I) the I51 intron.

scenario where cleavage follows immediately after splicing would produce the same RNA species.

In addition to the I51-containing and I51-lacking RT–PCR products, an RNA species ~20 nucleotides larger than the +I51 product was amplified (Figure 5B, asterisk). This product was observed only when the I51 band was present, suggesting that it arises during intron splicing. Presently, we do not know the identity of this band.

The processed ORF-containing RNA is polyadenylated

Nearly every fully processed eukaryotic mRNA has a polyadenylated 3′ end. The poly(A) tail has been implicated in the stabilization and nuclear export of mRNA as well as the cytoplasmic processes of translation and mRNA degradation (reviewed by Baker, 1997). The I-*Dir*I ORF has two overlapping consensus AAUAAA polyadenylation signals at its 3′ end (position 1065–1074, Figure 4A). An RT–PCR product was generated using a downstream poly(dT)-containing primer and an upstream primer complementary to ORF sequences (Figure 4B). This demonstrates that at least some of the I-*Dir*I ORF RNA is polyadenylated. The fact that no PCR product was amplified when the upstream primer was designed to hybridize to DiGIR1 sequences indicates that only IPS1/2-cleaved ORF RNA is polyadenylated *in vivo*. Direct sequencing of the PCR product mapped the polyadenylation site to position 1087 or 1088, before or after the adenosine residue 15 nucleotides downstream of the ORF 3′ end (Figure 4C).

Re-amplification of the polyadenylated I-*Dir*I ORF product with I51-flanking primers yielded predominantly a PCR product lacking the intron, but traces of intron-containing poly(A) RNA also were observed (data not shown). This suggests that although polyadenylation probably takes place first, it is followed rapidly by I51 splicing. The poly(A) site is upstream of IPS3, implying that both RNA2 and the IPS3-cleaved RNA3 may be substrates for polyadenylation. However, while IPS3 is readily detected by primer extension (Figure 2C), the same primers fail to produce evidence for a poly(A) site-cleaved RNA (data not shown). Thus, cleavage at IPS3 seems to be a prerequisite for polyadenylation.

The different ORF RNAs have distinct intracellular distribution patterns

In order to study in which cellular compartment the different processing steps take place, Lat3-5 cells were fractionated and RNA isolated from the nucleus, cytosol

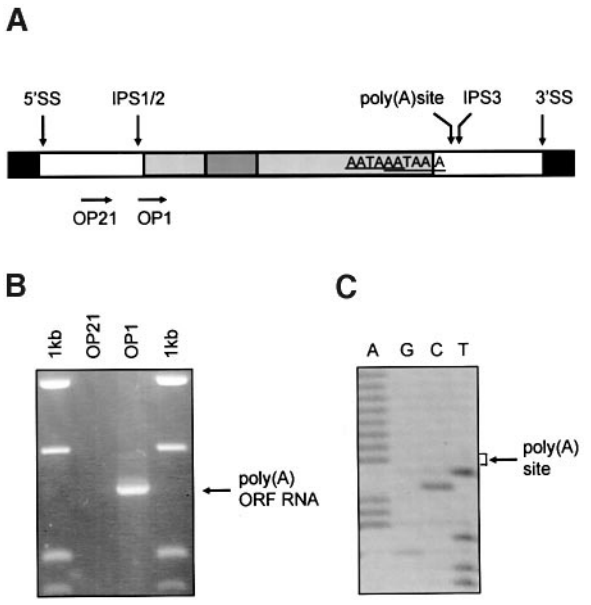

Fig. 4. Polyadenylation. (**A**) Schematic presentation of the DiSSU1 intron showing the AAUAAA polyadenylation signals, the polyadenylation site and the internal processing site downstream of the ORF (IPS3). The positions of the upstream primers applied for RT–PCR in (B) are indicated. (**B**) Detection of polyadenylated RNA by RT–PCR. A 2 µg aliquot of Lat3-5 whole-cell RNA was subjected to reverse transcription with a poly(T) primer (OP41). Primer OP313, which is identical to the specific sequence at the 5′ end of primer OP41 (see Table I), was used as the downstream primer to PCR amplify the cDNA. Primers OP1 or OP21 were used as upstream primers. The PCR products were analysed on a 2.5% agarose gel; 1 kb denotes the 1 kb DNA size marker (Gibco-BRL). (**C**) DNA sequence of the polyadenylated RT–PCR product. OP 283 was used to prime the reaction.

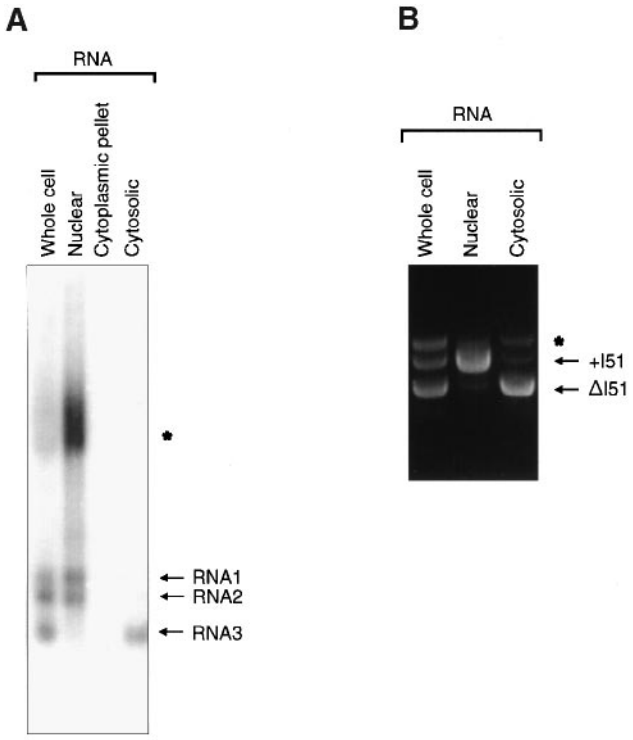

Fig. 5. Intracellular distribution of the ORF RNAs. (**A**) Whole-cell, nuclear, cytoplasmic pellet or cytosolic RNA from 10^6 Lat3-5 cells was run on a 1% denaturing agarose gel and analysed by Northern hybridization using the ORF probe described in Figure 1. An equal amount of whole-cell S7 RNA, which was run in parallel, did not produce any signal upon hybridization with the same probe (data not shown). The identity of the signal marked * is unknown. (**B**) RT–PCR of whole-cell, nuclear and cytosolic Lat3-5 RNA. Oligo OP42 was used for reverse transcription, followed by PCR with oligos OP1 and OP42. The PCR products resulting from amplification of intron-containing and intron-lacking I-*Dir*I ORF RNA are marked +I51 and I51, respectively. The origin of the band marked * has not been determined.

and a cytoplasmic pellet fraction containing mitochondria and other organelles. Northern analysis showed the different ORF-containing RNAs to have distinct intracellular distribution patterns (Figure 5A). While the full-length intron (RNA1) and RNA2 dominate in the nucleus, RNA3 is found almost exclusively in the cytosol. As expected, no ORF RNA was found in the cytoplasmic pellet. The blurred hybridization signal corresponding in size to the 3.9 kb LSU rRNA was observed repeatedly in whole-cell and especially nuclear RNA, and may represent a splicing intermediate distorted by co-migration with the rRNA. A similar signal was observed in Northern analysis of the PpLSU2 group I intron (Rocheleau and Woodson, 1995). In this case, the hybridization at 3.9 kb was explained as a non-covalent interaction between the free intron and the LSU rRNA.

As a rule, spliceosomal introns must be removed by the nucleoplasmic splicing machinery before mRNAs can be exported to the cytoplasm for translation. To examine whether this is also the case for the I-*Dir*I mRNA, Lat3-5 nuclear and cytosolic RNA was subjected to RT–PCR using primers flanking the I51 intron. The results showed that while the nuclear fraction mainly contained +I51 RNA, I51 RNA dominated in the cytosolic fraction (Figure 5B).

RNA3 is the I-DirI ORF mRNA

The polysomal distribution of a transcript can be used to determine whether it is translated *in vivo*. Actively translated mRNAs are present on the polysomes, while stored or masked mRNAs exist as messenger ribonucleoprotein complexes (mRNPs). Generally the two populations have sufficient size difference to allow separation on an Mg^{2+}-containing sucrose density gradient. If the Mg^{2+} is replaced with EDTA, the polysomes will dissociate. Thus, a translated mRNA can be detected by its shift from the polysomes in the presence of Mg^{2+} to smaller particles in the presence of EDTA. Untranslated mRNAs usually exhibit an unchanged sedimentation profile (Meyuhas *et al*., 1996).

We have shown previously that DiSSU1 is mobile when haploid intron-containing and intron-less amoebae are mated (Johansen *et al*., 1997). The intron-encoded protein I-*Dir*I is an endonuclease that cleaves the intron-less rDNA allele at the intron insertion site, presumably initiating the intron homing process. In order to determine which RNA species is the I-*Dir*I mRNA, a cytosolic extract from Lat3-5 amoebae was centrifuged through parallel Mg^{2+}- and EDTA-containing 20–47% linear sucrose gradients (Figure 6A). Slot-blot analysis of RNA from the resulting fractions showed a shift of I-*Dir*I ORF RNA from the polysomes in the Mg^{2+} gradient to smaller particles in the EDTA gradient (Figure 6B). This strongly suggests that the I-*Dir*I ORF RNA is translated. However,

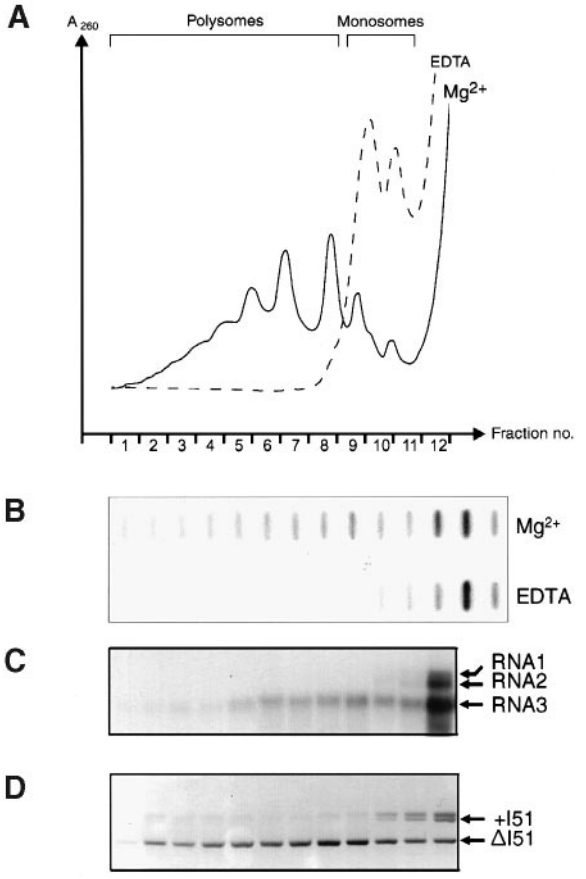

Fig. 6. Polysomal association of the I-*Dir*I mRNA. (**A**) A_{260} sedimentation profile of equal amounts of cytoplasmic extract in 20–47% (w/w) sucrose gradients in 5 mM $MgCl_2$ (Mg^{2+}) or 10 mM EDTA (EDTA). The 80S monosome and the 60S and 40S ribosomal subunits are denoted 'monosomes'. The rRNA species were identified from the ethidium bromide-stained gel used for Northern analysis in (C) (data not shown). (**B**) Slot-blot analysis of I-*Dir*I ORF RNA from one-fifth of the volume of the Mg^{2+} and EDTA fractions. The ORF probe was the same as in Figure 1. (**C**) Northern blot analysis of RNA from one-fifth of the volume of the Mg^{2+} fractions using the ORF probe. The RNA was run on a 1% denaturing agarose gel. (**D**) RT–PCR of RNA from the Mg^{2+} fractions. Oligo OP2 was used for reverse transcription, followed by PCR with oligos OP1 and OP2. The PCR products were analysed on a 2.5% agarose gel. The negative of the photographed gel is shown.

most of the ORF RNA was found in the fractions containing the ribosomal subunits (Figure 6B, 'monosomes') as well as in the fractions containing even smaller particles. This suggests that the loading of the I-*Dir*I mRNA onto the polysomes is inefficient. In order to examine which RNA species are translated, the Mg^{2+} fractions were subjected to further analysis. Only RNA3 was detected on the polysomes by Northern analysis (Figure 6C). RT–PCR using I51-flanking ORF-specific primers showed the polysome-associated RNA3 to lack the I51 intron (Figure 6D). This was also true for the polyadenylated RNA, as determined by RT–PCR with an oligo(dT) primer and re-amplification with primers flanking the I51 intron (data not shown). Taken together, the results imply that the intron-less polyadenylated form of RNA3 is the I-*Dir*I mRNA.

The subpolysome fractions also contained the I51-containing ORF species RNA1 and RNA2. Their presence may represent leakage from nuclei during cell fractionation. This possibility is supported by the fact that some of the polyadenylated RNA in these fractions was unspliced (data not shown). Alternatively, the cytoplasmic +I51 ORF RNA may represent genuine cytoplasmic species that exist as mRNPs and therefore are present in the subpolysome fractions in both the Mg^{2+} and EDTA gradients. Presently, we cannot exclude the possibility that these RNAs are translated at other life stages of the organism or under different physiological conditions. In any case, the absence of RNA1 and RNA2 from the polysomes suggests that they are not substrates for translation in the amoeba cells investigated in the present study.

Discussion

We have examined the *in vivo* expression of the I-*Dir*I ORF, which is encoded by the nucleolar DiSSU1 group I intron. Several novel features were found. First, the I-*Dir*I mRNA seems to be a rare example of a natural eukaryotic mRNA transcribed by pol I as part of the pre-rRNA transcript. Secondly, the DiGIR1 ribozyme functions in the *in vivo* maturation of the I-*Dir*I mRNA. It is thus the only known group I ribozyme with an *in vivo* function other than splicing. Thirdly, the I-*Dir*I messenger precursor resembles pol II transcripts in that it contains a 51 nucleotide spliceosomal intron (I51) and polyadenylation signals. The removal of I51 seems to be an obligatory step in the formation of the mRNA. The 3′ end of the mRNA is processed by cleavage at IPS3, an additional *in vivo* internal processing site located downstream of the ORF, and by polyadenylation. Based on its association with the cytoplasmic polysomes, we propose that the I51-lacking polyadenylated RNA3 is the I-*Dir*I mRNA. Taken together, our results suggest that DiSSU1 uses a unique combination of intron-supplied ribozyme activity and adaptation to the general pol II pathway of mRNA expression to allow a protein to be produced from the pol I rRNA transcript.

The I-DirI mRNA is synthesized as part of the pre-rRNA transcript

We propose that the I-*Dir*I mRNA is synthesized as part of the pol I-transcribed pre-rRNA and matured by processing of the excised DiSSU1 intron (Figure 7A). This conclusion is supported by several lines of evidence. First, the dominant I-*Dir*I-coding RNA in the cytoplasm is RNA3, which is the proposed end-product of the DiSSU1 maturation pathway. The processing intermediates, RNA1 and RNA2, are also detected readily in the *Didymium* cells. Previous reports on pol I-promoted protein production (Fleischer and Grummt, 1983; Grummt and Skinner, 1985) have been disputed by Lopata *et al.* (1986) who found that although pol I transcripts were present, only a subpopulation of aberrantly initiated pol II transcripts were associated with the polysomes. However, it has been demonstrated subsequently that exogenous pol I and pol III transcripts can be loaded onto the polysomes and translated in yeast and human cells, respectively (Gunnery and Mathews, 1995; Lo *et al.*, 1998). Although we cannot rigorously exclude the existence of cryptic pol II promoters, we have not detected any aberrantly initiated I-*Dir*I transcripts.

Experimentally, pol I transcription is often distinguished

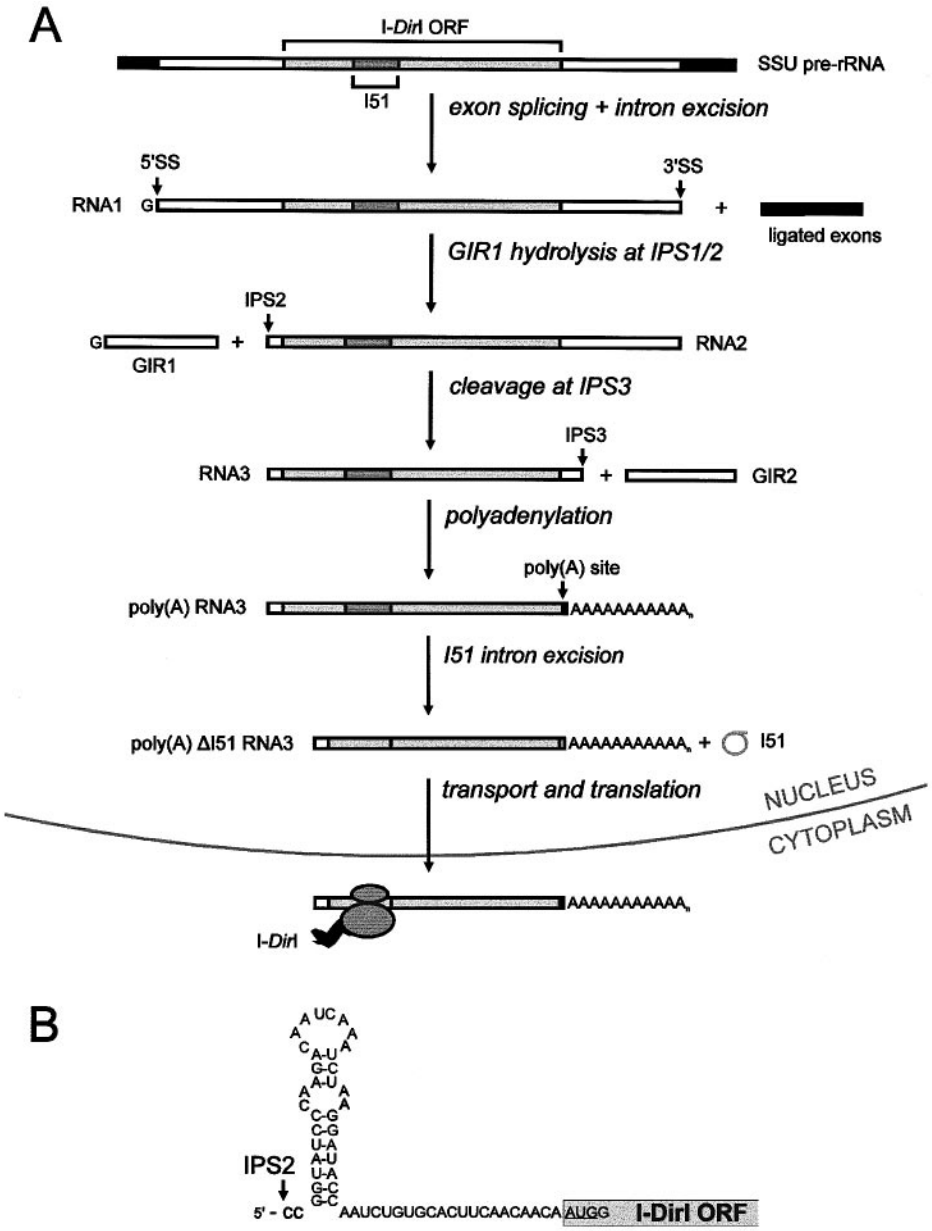

Fig. 7. The I-*Dir*I mRNA. (**A**) The proposed *in vivo* DiSSU1 splicing and processing pathway. The G at the 5′ end of the full-length intron is the exogenous guanosine cofactor added during intron splicing. Exon, I-*Dir*I ORF, I51 spliceosomal intron and ribozyme sequences are shown in black, light grey, dark grey and white, respectively. The 5′ and 3′ splice sites (SS), internal processing sites (IPS) and polyadenylation site [poly(A)] are indicated. (**B**) Putative secondary structure of the I-*Dir*I mRNA 5′ end. The AUG start codon is underlined.

from that of pol II by the resistance and sensitivity, respectively, of the two enzymes to the fungal toxin α-amanitin. However, the toxin is not taken up readily by myxomycetes (Braun and Seebeck, 1982), making it necessary to test polymerase specificities by alternative strategies. The I-*Ppo*I homing endonuclease from the myxomycete *P.polycephalum* can be generated from PpLSU3 integrated in yeast rDNA (Muscarella and Vogt, 1993). Recently, a yeast strain temperature-sensitive for pol I was used to demonstrate that the I-*Ppo*I mRNA is indeed the product of pol I-promoted rRNA transcription in yeast (Lin and Vogt, 1998).

Finally, protein production from pol I transcripts is not totally unprecedented. Exogenously introduced reporter genes have been expressed by pol I in kinetoplastids (Rudenko *et al*., 1991; Zomerdijk *et al*., 1991), mammalian cells (Palmer *et al*., 1993) and yeast (Lo *et al*., 1998). However, as far as we are aware, natural pol I-promoted mRNAs have only been identified in *Trypanosoma brucei*, where several highly expressed genes display α-amanitin-resistant transcription (reviewed in Lee and Van der Ploeg, 1997). The fact that the kinetoplastid mRNAs are matured by trans-splicing (the addition of a capped mini exon onto the 5′ end of the main coding exon) has been used to explain how mRNA can be synthesized by pol I in these organisms.

DiGIR1—a group I ribozyme with an in vivo function in mRNA 5′ end formation

The *Didymium* and *Naegleria* GIR1 ribozymes are unusual in their small size and compact structure as well as in catalysing a cleavage reaction (see Einvik *et al*., 1998).

Furthermore, they are unique in that they exist as inserts within other group I ribozymes. In this report, we show that *Didymium* GIR1 function *in vivo* parallels its *in vitro* reaction in that it cleaves the DiSSU1 intron at an IPS upstream of the I-*Dir*I ORF. The specific association of the DiGIR1-detached RNA3 with the polysomes implies that cleavage at IPS2 is obligatory for the production of the I-*Dir*I mRNA. In agreement with this, recent studies in yeast have shown that mutations in NaGIR1 that destroy its ribozyme activity prevent protein production from the downstream endonuclease ORF (W.Decatur, S.Johansen and V.M.Vogt, in preparation). DiGIR1 RNA is degraded rapidly after separation from the rest of the intron RNA, substantiating that its function is completed upon cleavage. GIR1 is the only known group I ribozyme with an *in vivo* function other than splicing.

After DiGIR1 cleavage, the I-*Dir*I mRNA contains a 55 nucleotide leader upstream of the ORF initiation codon. Interestingly, this sequence can be folded into a stabile stem–loop structure only two nucleotides from the 5′ end of the transcript (Figure 7B). *In vitro* results suggest that this structure can only be formed after IPS2 cleavage, since some of its sequences are part of the essential P10 helix in the catalytically active DiGIR1 (Einvik *et al*., 1998). Other types of intron-encoded RNAs, namely some of the box C/D and box H/ACA snoRNAs, also contain ends that are produced by processing events. The uncapped 5′ ends of these molecules form stem–loop structures which are believed to be important for their stability (Bachellerie *et al*., 1995; Balakin *et al*., 1996). All DiSSU1 RNAs are present at low amounts compared with the SSU rRNA from which they are excised, suggesting that the intron RNAs in general have short half-lives. However, RNA3 is detected readily in transcriptionally inactive *Didymium* cysts (A.Vader, unpublished results), implying that this particular ORF species is stabilized specifically. We propose that the putative stem–loop at the 5′ untranslated region (UTR) of the cleaved I-*Dir*I ORF RNA acts to stabilize the mRNA (Figure 7B), possibly by binding a protein.

The I-DirI mRNA resembles an RNA polymerase II transcript

A spliceosomal intron (I51) was found within the I-*Dir*I ORF. Only the intron-less version of the mRNA is associated with the polysomes in the analysed vegetative amoeba cells. This is the first report of a spliceosomal intron within a group I intron and an rRNA transcript. What can be its significance? One possibility is that I51 splicing is needed to ensure efficient transport of the mRNA from the nucleus to the cytoplasm. Several examples of intron-dependent gene expression have been reported, including the well-described mammalian β-globin gene (Buchman and Berg, 1988; Collis *et al*., 1990). Here, transcripts are not exported to the cytoplasm unless the precursor mRNA contains at least one intron. In *Didymium*, unspliced and spliced I-*Dir*I RNAs are sharply divided between the nucleus and cytoplasm, respectively (see Figure 5), supporting a possible intron-dependent expression pathway for the I-*Dir*I gene. Alternatively, the presence of I51 may allow the formation of different I-*Dir*I protein isoforms through facultative mRNA splicing. A provocative observation is that although I51 splicing removes 17 amino acid codons, it does not alter the I-*Dir*I reading frame (see Figure 3B). Thus, two versions of the I-*Dir*I protein, with perhaps slightly different functions, may be synthesized. While the results indicate that the spliced I-*Dir*I dominates at the haploid amoeba stage (see Figure 6D), an unspliced version of the protein may be produced during other developmental stages of the myxomycete life cycle (see Johansen *et al*., 1997). In any case, the fact that I51 is the only spliceosomal intron reported within a group I intron strongly suggests that it has been inserted into the pre-existing DiSSU1 intron or I-*Dir*I ORF. As such, its presence can be taken as an argument in favour of the ‘intron-late’ hypothesis, which states that spliceosomal introns have invaded pre-assembled protein-coding genes late in evolution (see Logsdon *et al*., 1998).

During *in vivo* DiSSU1 processing, the intron is cleaved at an internal processing site (IPS3) located downstream of the I-*Dir*I ORF. Our results suggest that IPS3 processing takes place in the nucleus, prior to polyadenylation. Cleavage at IPS3 has not been observed *in vitro*, implying that cellular factors facilitate the reaction. Its importance is suggested by the fact that the PpLSU3 and NaSSU1 introns are also cleaved downstream of their respective ORFs *in vivo* (Ellison, 1994; W.Decatur, S.Johansen and V.M.Vogt, in preparation). While the splitting of PpLSU3 is a G-addition reaction which is catalysed by the ribozyme itself *in vitro* (Ruoff *et al*., 1992), NaSSU1 processing is similar to that of DiSSU1 in that it only takes place *in vivo*. The biological function of the internal cleavage reaction is unclear. In the case of DiSSU1, IPS3 cleavage removes the compactly folded DiGIR2 RNA, possibly making the sequences immediately downstream of the I-*Dir*I ORF available to the polyadenylation machinery.

The 3′ end of the mature I-*Dir*I mRNA resembles most eukaryotic mRNAs in that it contains a poly(A) tail. Polyadenylation is usually restricted to pol II transcripts, and the C-terminal domain (CTD) of pol II has even been reported to be essential for polyadenylation (McCracken *et al*., 1997; Hirose and Manley, 1998). Nevertheless, poly(A) tails have been found on both natural and exogenous pol I-transcribed trans-spliced mRNAs in the kinetoplastids (Rudenko *et al*., 1991; Zomerdijk *et al*., 1991; Lodes *et al*., 1995). Similarly, exogenous pol I transcripts containing poly(A) signals were polyadenylated in mammalian and yeast cells (Palmer *et al*., 1993; Lo *et al*., 1998). Thus, polyadenylation does not seem to be coupled exclusively to pol II transcription, but rather to depend upon the presence of a poly(A) signal in the transcript. The I-*Dir*I ORF contains two overlapping conserved AAUAAA polyadenylation signals at its 3′ end.

The 5′ cap and the 3′ poly(A) tail are believed to facilitate translational initiation (reviewed by Sachs *et al*., 1997). However, the fact that functional protein could be produced from an exogenous pol III-transcribed mRNA that lacked a cap and poly(A) tail (Gunnery and Mathews, 1995) shows that there is no absolute requirement for either. Recently, it has been shown that a poly(A) tail on an mRNA will stimulate its translation in yeast extracts and yeast cells, even in the absence of a 5′ cap (Iizuka *et al*., 1994; Tarun and Sachs, 1995; Preiss and Hentze, 1998). This has led to the proposal of a translational initiation model where both the poly(A) tail and cap may

act as translational promoters in the recruitment of the small ribosomal subunit, while the cap docks the initiator complex at the translational start site (Sachs *et al*., 1997; Preiss and Hentze, 1998). Thus, the poly(A) tail of the I-*Dir*I mRNA may aid in recruiting ribosomes to the messenger. However, the fact that only a small amount of the I-*Dir*I mRNA was associated with the polysomes suggests that the poly(A) tail is not sufficient to ensure efficient I-*Dir*I production. This is in agreement with studies in yeast, where it was found that an exogenous capless, polyadenylated pol I transcript was translated at <10% of the level of its pol II-transcribed capped counterpart (Lo *et al*., 1998). Intron-encoded homing endonucleases are generally present at very low concentrations in the cells of their natural host organism. Thus, copious protein production of the I-*Dir*I endonuclease is probably not needed and may even be deleterious to the *Didymium* cells.

Host adaptation

In conclusion, expression of I-*Dir*I combines intron-specific strategies (DiGIR1-catalysed mRNA 5′ end formation) with adaptation to the nuclear mRNA expression pathway (spliceosomal intron and polyadenylation). Inspection of the other group I introns for which expression data are available suggests that a combined expression strategy is a general feature. Although only DiSSU1 and NaSSU1 harbour GIR1 ribozymes, several introns use alternative splicing to put an internal ORF in-frame with the upstream exon. The ND1-I4 and cox1-I7 introns from *Podospora anserina* mitochondria both harbour two ORFs, of which only one is continuous with the upstream exon. While intron splicing will remove both ORFs, it has been reported that an alternative internal 3′ splice site can be used *in vivo* to place the downstream ORF in-frame with the upstream exon (Sellem and Belcour, 1994).

Examples of host features which have been adopted by introns include a conserved yeast mitochondrial 3′ UTR sequence in the ScLSU1 (previously termed ω) and aI5β introns (Zhu *et al*., 1987), and a poly(A) tail in the sea anemone mitochondrial MsND5i1 intron (Beagley *et al*., 1996). The ORFs of the T4 phage td and SunY introns resemble host genes in containing their own intron-internal T4 promoters, and in their mode of regulation. Similar to several T4 late genes, the lack of translation from the early pre-mRNAs can be explained by the presence of stable secondary structures that occlude the ribosome-binding site. These structures are absent when transcription is initiated from the late promoters (Gott *et al*., 1988). Host adaptation has been taken to the extreme in several mitochondrial introns which express structural genes that are important for host function. Examples include the ribosomal protein S5 which is encoded by LSU rDNA introns in *Neurospora crassa* and other fungi (Burke and RajBhandary, 1982; Cummings *et al*., 1989; Yamamoto *et al*., 1995), and two NADH dehydrogenase subunits which are expressed solely from introns in the *ND5* gene of three sea anemone species (Beagley *et al*., 1996). We propose that group I intron RNAs, by incorporating characteristic features of cellular transcripts, gain access to the normal mRNA maturation and transport pathways in the different compartments of their host cells. To address questions concerning the biological implications of the IPSs, the I51 spliceosomal intron and the polyadenylation signal of the DiSSU1 intron, functional analysis will have to be performed in a transformable organism.

Table I. Oligonucleotides

Oligo	Sequence	Position[a]	Direction[b]
OP1	5′-CACTTCTAGAACCATGGTGAA-AGGAACG-3′	301	F
OP2	5′-TGTCTGGATCCTCATCTG-3′	644	R
OP3	5′-CGAAATCCACACACCATC-3′	38	R
OP9	5′-CAACCATGGATCATTTGTAAA-CTATGTTT-3′	1092	F
OP12	5′-TCACCATGGTTGTTGAAGTG-CACAGATTG-3′	265	R
OP18	5′-TGGTGGATCCTGCCAGTAGT-3′	E24	F
OP21	5′-GGGTTAAAACGGTGGGGGA-3′	141	F
OP25	5′-CTCGAATTCGCTCTTGGAGCT-GGAATTA-3′	E638	R
OP41	5′-CGACGCATGCACGCA$(T)_{15}$-3′	poly(A)	
OP42	5′-ACACGTAGAGTATGCTTCAGCT-CCA-3′	831	R
OP66	5′-TACCTTTATACCAGC-3′	1423	R
OP178	5′-TTCCCCCGTACCAGGT-3′	1172	R
OP236	5′-CAGTTACCCACTAGTA-3′	1239	R
OP283	5′-TACGCTCCTCTAATGACTAC-3′	1029	F
T7	5′-AATTTAATACGACTCACTATA-3′	plasmid	

[a]The number is the position of the 3′ nucleotide of the oligo in the DiSSU1 sequence (Johansen and Vogt, 1994). OP18 and OP25 hybridize to exon sequences upstream of the DiSSU1 intron and are numbered according to the SSU rRNA gene exon sequence (E) of *P.polycephalum* (Johansen *et al*., 1988).
[b]F and R denote forward and reverse oligo primers, respectively.

Materials and methods

RNA isolation and Northern blotting

The *D.iridis* strains Lat3-5 (intron-containing) and S7 (intron-lacking) have been described previously by Johansen *et al*. (1997). A total of 10^7 amoeba were harvested by 5 min centrifugation at 400 *g*. The pellet was dissolved in 1 ml of Trizol reagent (Gibco-BRL) and RNA extracted according to the manufacturer's instructions. RNA was separated on a 2% formaldehyde/agarose gel in formaldehyde/1× MOPS (40 mM MOPS, 10 mM NaAc, 2 mM EDTA, 0.04% HAc) buffer and transferred to a nylon membrane by capillary action. Alternatively, RNA was electrophoresed on an 8 M urea–4% polyacrylamide gel in 1× TBE buffer (Mighty Small, Hoeffer) and transferred to a nylon membrane by electroblotting (Hoeffer; 45 min, 150 mA, 4°C). Prior to gel electrophoresis, the RNA was denatured for 15 min at 65°C in formaldehyde and formamide. For the agarose gels, 10 ng/μl ethidium bromide was added to the RNA samples prior to electrophoresis. Hybridization was carried out either in Rapid Hyb solution (Amersham) according to the manufacturer's protocol, or in 5× SSPE, 5× Denhardt's solution, 0.1% SDS and 50 μg/ml salmon sperm DNA (Stratagene) at 65°C overnight. In the latter case, post-hybridization washing was for 20 min at room temperature in 1× SSPE/0.1% SDS followed by 20 min at room temperature and finally 20 min at 65°C in 0.1× SSPE/0.1% SDS.

The ORF and SSU probes were amplified from Lat3-5 genomic DNA (Johansen *et al*., 1997) by PCR using oligos OP1 and OP2, and OP18 and OP25 (see Table I), respectively. The PCR products subsequently were purified on S-300 spin columns (Pharmacia). Oligos T7 and OP12 were used to amplify the first part of the DiSSU1 intron from plasmid pDiSSU-Δ353 (Johansen and Vogt, 1994). The PCR product was digested with *Ava*I and *Msp*I, yielding the 202 bp GIR1 probe. The 193 bp GIR2 probe was produced by amplification of pDiSSU-Δ353 with OP9 and OP66 and subsequent digestion with *Hph*I. Both GIR1 and GIR2 probes were gel-purified using Spin-X centrifuge filter units (Costar). The Mega Prime kit (Amersham) was used to label the probes with [α-^{32}P]dCTP (10 μCi/μl; Amersham).

RT–PCR, RNA sequencing and primer extension

For RT–PCR, the first strand cDNA synthesis kit (Pharmacia) was used to reverse transcribe 2 μg or more of RNA. One-tenth of the cDNA was amplified using Dynazyme II DNA polymerase (Finnzymes Oy).

For RNA sequencing, gel-purified oligos were labelled with [γ-^{32}P]ATP (10 μCi/μl; Amersham) using T4 polynucleotide kinase (Gibco-BRL). A 1 μg aliquot of *in vitro* transcribed RNA (see Johansen and Vogt, 1994) was added to 10 pmol of labelled oligo in 1× RT buffer [50 mM Tris–HCl pH 8, 60 mM KCl, 10 mM $MgCl_2$, 1 mM dithiothreitol (DTT)] in a total volume of 20 μl, denatured at 80°C for 2 min and incubated at 45°C for 10 min. Subsequently, 4 μl of RNA-oligo mixture was added to each of four tubes containing 1 U of avian myeloblastosis virus (AMV) reverse transcriptase (Pharmacia), 1 U of RNasin (Promega), 0.2 mM dATP, dCTP and dTTP and 0.4 mM dGTP (Pharmacia), and either one of 0.2 mM ddATP, ddCTP or ddTTP or 0.4 mM ddGTP (Pharmacia) in 1× RT buffer. The reaction was incubated 1 h at 40°C before being stopped by the addition of 5 μl of formamide loading buffer.

The protocol for primer extension analysis was similar to that described above except that 2 pmol of labelled oligo was incubated with RNA in 1× RT buffer in a total volume of 5 μl, and that no ddNTP was added during the reverse transcription reaction. The products of RNA sequencing and primer extension were denatured by heating at 100°C for 1 min before separation on an 8 M urea–8% polyacrylamide gel.

Cloning and DNA sequencing

PCR products were cloned into the *Sma*I site of pUC18 using the Pharmacia SureClone kit. For direct sequencing, 5 μl of RT–PCR product was incubated with 10 U of exonuclease I and 2 U of shrimp alkaline phosphatase (USB) for 15 min at 37°C. The enzymes were heat inactivated by a 15 min incubation at 80°C before 0.5 μl of the treated PCR product was subjected to sequencing using the Cycle Sequencing Kit (USB Corporation) and [α-^{33}P]dCTP (10 μCi/μl; Amersham).

Cell fractionation and polysome analysis

A total of 2×10^7 Lat3-5 cells were added to DS/2 growth medium (Johansen *et al*., 1997) to a total volume of 80 ml and centrifuged in two tubes at 300 *g* for 5 min. After resuspending the cell pellets in 2× 14 ml of DS/2, the cells were re-centrifuged. The pellet from one tube was dissolved in 1 ml of Trizol (see above), producing whole-cell RNA. The cells in the other tube were resuspended in 4 ml of ice-cold nuclear isolation buffer (0.25 M sucrose, 5 mM EDTA, 10 mM Tris–HCl pH 8.0, 1 mM spermine, 0.1% Triton X-100, 2 mM $MgCl_2$), lysed in a blender for 10 s and centrifuged at 1000 *g*, 4°C for 30 min. The pelleted nuclei were dissolved in Trizol (nuclear RNA), while the supernatant was re-centrifuged at 10 000 *g*, 4°C for 10 min. The resulting pellet was dissolved in 1 ml of Trizol (cytoplasmic pellet RNA), while the RNA in the supernatant was precipitated by ethanol prior to adding 1 ml of Trizol (cytosolic RNA).

For analysis of the polysomes, 5×10^7 Lat3-5 cells were added to ice-cold DS/2 with 10 μg/ml cycloheximide to a total volume of 50 ml, and centrifuged at 1000 r.p.m., 4°C for 5 min. The cell pellet was resuspended in 500 μl of lysis buffer (20 mM Tris–HCl pH 8.0, 1.5 mM $MgCl_2$, 140 mM KCl, 1.5 mM DTT, 1 mM $CaCl_2$, 0.1 mM EDTA, 0.16 mM cycloheximide, 0.5% NP-40, 500 U/ml RNasin) and centrifuged at 10 000 *g*, 4°C for 10 min. The supernatant was applied to a linear 20–47% sucrose gradient in 20 mM Tris–HCl pH 8.0, 140 mM KCl with 5 mM $MgCl_2$ or 10 mM EDTA, respectively, and centrifuged for 2 h and 15 min, 4°C at 40 000 r.p.m. in a Beckman SW41 rotor. Fractions of ~1 ml were collected with concominant measurement of the absorbance at 260 nm, and immediately frozen at –70°C. RNA was isolated by phenol/chloroform extraction. Slot-blotting was carried out according to Sambrook *et al*. (1989) using nylon filters (Zeta-probe, Bio-Rad).

Acknowledgements

We thank Mogens Engelhardt, Jan Christiansen and Finn Haugli for advice on cell separation and polysome analyses, Kari Haugli and Franz Frenzel for technical assistance, and Jue Lin, Wayne Decatur and Volker M.Vogt for communicating unpublished results. Volker M.Vogt is also gratefully acknowledged for critical reading of the manuscript. This work was supported by grants from NORFA (A.V.), The Danish Research Council for Natural Sciences (H.N.), The NOVO Foundation (H.N.), The Norwegian Research Council (S.J.), The Norwegian Cancer Society (S.J.) and The Aakre Foundation for Cancer Research (S.J.).

References

Bachellerie,J.P., Michot,B., Nicoloso,M., Balakin,A., Ni,J. and Fournier,M.J. (1995) Antisense snoRNAs: a family of nucleolar RNAs with long complementarities to rRNA. *Trends Biochem. Sci.*, **20**, 261–264.

Baker,E.J. (1997) mRNA polyadenylation: functional implications. In Harford,J.B. and Morris,D.R. (eds), *mRNA Metabolism and Post-transcriptional Gene Regulation*. Wiley-Liss, New York, pp. 85–105.

Balakin,A.G., Smith,L. and Fournier,M.J. (1996) The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions. *Cell*, **86**, 823–834.

Beagley,C.T., Okada,N.A. and Wolstenholme,D.R. (1996) Two mitochondrial group I introns in a metazoan, the sea anemone *Metridium senile*: one intron contains genes for subunits 1 and 3 of NADH dehydrogenase. *Proc. Natl Acad. Sci. USA*, **93**, 5619–5623.

Braun,R. and Seebeck,T. (1982) RNA metabolism. In Aldrich,H.C. and Daniel,J.W. (eds), *Cell Biology of Physarum and Didymium*. Academic Press, New York, pp. 393–435.

Brehm,S.L. and Cech,T.R. (1983) Fate of an intervening sequence ribonucleic acid: excision and cyclization of the *Tetrahymena* ribosomal ribonucleic acid intervening sequence *in vivo*. *Biochemistry*, **22**, 2390–2397.

Buchman,A.R. and Berg,P. (1988) Comparison of intron-dependent and intron-independent gene expression. *Mol. Cell. Biol.*, **8**, 4395–4405.

Burke,J.M. and RajBhandary,U.L. (1982) Intron within the large rRNA gene of *N.crassa* mitochondria: a long open reading frame and a consensus sequence possibly important in splicing. *Cell*, **31**, 509–520.

Cech,T. and Herschlag,D. (1996) Group I ribozymes substrate recognition, catalytic strategies, and comparative mechanistic analysis. In Eckstein,F. and Lilley,J.M.D. (eds), *Catalytic RNA*. Springer-Verlag, Berlin, Germany, pp. 1–17.

Collis,P., Antoniou,M. and Grosveld,F. (1990) Definition of the minimal requirements within the human β-globin gene and the dominant control region for high level expression. *EMBO J.*, **9**, 233–240.

Cummings,D.J., Domenico,J.M., Nelson,J. and Sogin,M.L. (1989) DNA sequence, structure, and phylogenetic relationship of the small subunit rRNA coding region of mitochondrial DNA from *Podospora anserina*. *J. Mol. Evol.*, **28**, 232–241.

Decatur,W.A., Einvik,C., Johansen,S. and Vogt,V.M. (1995) Two group I ribozymes with different functions in a nuclear rDNA intron. *EMBO J.*, **14**, 4558–4568.

Din,N., Engberg,J., Kaffenberger,W. and Eckert,W.A. (1979) The intervening sequence in the 26S rRNA coding region of *T.thermophila* is transcribed within the largest stable precursor for rRNA. *Cell*, **18**, 525–532.

Einvik,C., Nielsen,H., Westhof,E., Michel,F. and Johansen,S. (1998) Group I-like ribozymes with a novel core organization perform obligate sequential hydrolytic cleavages at two processing sites. *RNA*, **4**, 530–541.

Elde,M., Haugen P., Willassen,N.P. and Johansen,S. (1999) I-*Nja*I, a nuclear intron-encoded homing endonuclease from *Naegleria*, generates a pentanucleotide 3′ cleavage-overhang within a 19 base pair partially symmetric DNA recognition site. *Eur. J. Biochem.*, **259**, 281–288.

Ellison,E. (1994) The nature of I-*Ppo*I and PpLSU3 *in vivo*. PhD Thesis, Cornell University, pp. 79–105.

Fleischer,S. and Grummt,I. (1983) Expression of an mRNA coding gene under the control of an RNA polymerase I promoter. *EMBO J.*, **2**, 2319–2322.

Gott,J.M., Zeeh,A., Bell Pedersen,D., Ehrenman,K., Belfort,M. and Shub,D.A. (1988) Genes within genes: independent expression of phage T4 intron open reading frames and the genes in which they reside. *Genes Dev.*, **2**, 1791–1799.

Grummt,I. and Skinner,J.A. (1985) Efficient transcription of a protein-coding gene from the RNA polymerase I promoter in transfected cells. *Proc. Natl Acad. Sci. USA*, **82**, 722–726.

Gunnery,S. and Mathews,M.B. (1995) Functional mRNA can be generated by RNA polymerase III. *Mol. Cell. Biol.*, **15**, 3597–3607.

Hirose,Y. and Manley,J.L. (1998) RNA polymerase II is an essential mRNA polyadenylation factor. *Nature*, **395**, 93–96.

Iizuka,N., Najita,L., Franzusoff,A. and Sarnow,P. (1994) Cap-dependent and cap-independent translation by internal initiation of mRNAs in cell extracts prepared from *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **14**, 7322–7330.

Johansen,S. and Vogt,V.M. (1994) An intron in the nuclear ribosomal DNA of *Didymium iridis* codes for a group I ribozyme and a novel ribozyme that cooperate in self-splicing. *Cell*, **76**, 725–734.

Johansen,S., Johansen,T. and Haugli,F. (1988) Nucleotide sequence of the *Physarum polycephalum* small subunit ribosomal RNA as inferred

from the gene sequence: secondary structure and evolutionary implications. *Curr. Genet.*, **14**, 265–273.
Johansen,S., Elde,M., Vader,A., Haugen,P., Haugli,K. and Haugli,F. (1997) *In vivo* mobility of a group I twintron in nuclear ribosomal DNA of the myxomycete *Didymium iridis*. *Mol. Microbiol.*, **24**, 737–745.
Kreivi,J.P. and Lamond,A.I. (1996) RNA splicing: unexpected spliceosome diversity. *Curr. Biol.*, **6**, 802–805.
Lee,M.G.-S. and Van der Ploeg,L.H.T. (1997) Transcription of protein-coding genes in trypanosomes by RNA polymerase I. *Annu. Rev. Microbiol.*, **51**, 463–489.
Lewis,J.D., Gunderson,S.I. and Mattaj,I.W. (1995) The influence of 5′ and 3′ end structures on pre-mRNA metabolism. *J. Cell Sci.*, **19**, 13–19.
Lin,J. and Vogt,V.M. (1998) I-*Ppo*I, the endonuclease encoded by the group I intron PpLSU3, is expressed from an RNA polymerase I transcript. *Mol. Cell. Biol.*, **18**, 5809–5817.
Lo,H.-J., Huang,H.-K. and Donahue,T.F. (1998) RNA polymerase I-promoted HIS4 expression yields uncapped, polyadenylated mRNA that is unstable and inefficiently translated in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **18**, 665–675.
Lodes,M.J., Merlin,G., deVos,T., Ghosh,A., Madhubala,R., Myler,P.J. and Stuart,K. (1995) Increased expression of LD1 genes transcribed by RNA polymerase I in *Leishmania donovani* as a result of duplication into the rRNA gene locus. *Mol. Cell. Biol.*, **15**, 6845–6853.
Logsdon,J.M.,Jr, Stoltzfus,A. and Doolittle,W.F. (1998) Recent cases of spliceosomal intron gain? *Curr. Biol.*, **8**, 560–563.
Lopata,M.A., Cleveland,D.W. and Sollner Webb,B. (1986) RNA polymerase specificity of mRNA production and enhancer action. *Proc. Natl Acad. Sci. USA*, **83**, 6677–6681.
McCracken,S., Fong,N., Yankulov,K., Ballantyne,S., Guohua,P., Greenblatt,J., Patterson,S.D., Wickens,M. and Bentley,D.L. (1997) The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature*, **385**, 357–361.
Meyuhas,O., Bibrerman,Y., Pierandrei-Amaldi,P. and Amaldi,F. (1996) Isolation and analysis of polysomal RNA. In Krieg,P.A. (ed.), *A Laboratory Guide to RNA: Isolation, Analysis and Synthesis*. Wiley-Liss, New York, pp. 65–81.
Muscarella,D.E. and Vogt,V.M. (1989) A mobile group I intron in the nuclear rDNA of *Physarum polycephalum*. *Cell*, **56**, 443–454.
Muscarella,D.E. and Vogt,V.M. (1993) A mobile group I intron from *Physarum polycephalum* can insert itself and induce point mutations in the nuclear ribosomal DNA of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **13**, 1023–1033.
Nishida,H., Tajiri,Y. and Sugiyama,J. (1998) Multiple origins of fungal group I introns located in the same position of nuclear SSU rRNA gene. *J. Mol. Evol.*, **46**, 442–448.
Palmer,T.D., Miller,A.D., Reeder,R.H. and McStay,B. (1993) Efficient expression of a protein coding gene under the control of an RNA polymerase I promoter. *Nucleic Acids Res.*, **21**, 3451–3457.
Preiss,T. and Hentze,M.W. (1998) Dual function of the messenger RNA cap structure in poly(A)-tail-promoted translation in yeast. *Nature*, **392**, 516–519.
Rocheleau,G.A. and Woodson,S.A. (1995) Enhanced self-splicing of *Physarum polycephalum* intron 3 by a second group I intron. *RNA*, **1**, 183–193.
Rudenko,G., Chung,H.M., Pham,V.P. and Van der Ploeg,L.H. (1991) RNA polymerase I can mediate expression of CAT and neo protein-coding genes in *Trypanosoma brucei*. *EMBO J.*, **10**, 3387–3397.
Ruoff,B., Johansen,S. and Vogt,V.M. (1992) Characterization of the self-splicing products of a mobile intron from the nuclear rDNA of *Physarum polycephalum*. *Nucleic Acids Res.*, **20**, 5899–5906.
Sachs,A.B., Sarnow,P. and Hentze,M.W. (1997) Starting at the beginning, middle, and end: translation initiation in eukaryotes. *Cell*, **89**, 831–838.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sellem,C.H. and Belcour,L. (1994) The *in vivo* use of alternate 3′-splice sites in group I introns. *Nucleic Acids Res.*, **22**, 1135–1137.
Tarun,S.Z.,Jr and Sachs,A.B. (1995) A common function for mRNA 5′ and 3′ ends in translation initiation in yeast. *Genes Dev.*, **9**, 2997–3007.
Yamamoto,H., Naruse,A., Ohsaki,T. and Sekiguchi,J. (1995) Nucleotide sequence and characterization of the large mitochondrial rRNA gene of *Penicillium urticae*, and its comparison with those of other filamentous fungi. *J. Biochem.*, **117**, 888–896.
Zaug,A.J. and Cech,T.R. (1980) *In vitro* splicing of the ribosomal RNA precursor in nuclei of *Tetrahymena*. *Cell*, **19**, 331–338.
Zhu,H., Macreadie,I.G. and Butow,R.A. (1987) RNA processing and expression of an intron-encoded protein in yeast mitochondria: role of a conserved dodecamer sequence. *Mol. Cell. Biol.*, **7**, 2530–2537.
Zomerdijk,J.C., Kieft,R. and Borst,P. (1991) Efficient production of functional mRNA mediated by RNA polymerase I in *Trypanosoma brucei*. *Nature*, **353**, 772–775.

Received October 13, 1998; revised and accepted December 14, 1998