Nucleic Acids Research, 1994, Vol. 22, No. 7 1135–1137

The *in vivo* use of alternate 3′-splice sites in group I introns

Carole H.Sellem* and Léon Belcour
Centre de Génétique Moléculaire, CNRS, Gif-sur-Yvette, F91198 Cedex, France

Received January 31, 1994; Accepted March 3, 1994

ABSTRACT

Alternative splicing of group I introns has been postulated as a possible mechanism that would ensure the translation of proteins encoded into intronic open reading frames, discontinous with the upstream exon and lacking an initiation signal. Alternate splice sites were previously depicted according to secondary structures of several group I introns. We present here strong evidence that, in the case of *Podospora anserina nad1-i4* and *cox1-i7* mitochondrial introns, alternative splicing events do occur *in vivo*. Indeed, by PCR experiments we have detected molecules whose sequence is precisely that expected if the predicted alternate 3′-splice sites were used.

INTRODUCTION

Many of the group I introns contain an open reading frame (ORF) encoding a protein proposed to be involved either in intron splicing as an RNA maturase, or in intron mobility as a highly specific DNA endonuclease (1). The translation of the intronic ORFs is generally achieved in continuity with the upstream exon or initiated at an intronic initiation codon (2). Alternative splicing was first invoked by Hensgens *et al.* (3) as a mean for the expression of a mitochondrial intronic ORF which neither is contiguous with the upstream exon nor contains any translation initiation signal. Alternative splicing, as normal splicing of group I introns, would be guided in the splice site selection by secondary structures of the intron and achieved by its catalytic core (4). The selection of the 3′-splice site was shown, on the basis of sequence examinations (5, 6) and in vitro experiments (7), to depend upon either one or both of two long distance pairings involving nucleotides immediately preceeding -P9-0- (8, 9, 10) and following -P10- (6, 11) the invariable G terminal nucleotide residue of the group I introns. The alternative selection of an internal 3′-splice site, concerning only a minority of the primary transcripts, would bring the discontinous ORF in frame with the upstream exon and thus ensure a low level translation of the intron encoded protein. Even though the presence of alternate P9-0 and P10 signals at the begining of several group I introns free standing ORFs has allowed precise prediction of alternative splicing events (12), it remained to be shown that they do operate *in vivo*.

For this purpose, we have employed Polymerase Chain Reaction (PCR) analysis to search for molecules resulting from the expected alternative splicing events of two mitochondrial group IB introns, *nad1-i4* and *cox1-i7* of the filamentous fungus *Podospora anserina*.

MATERIALS AND METHODS

Senescent cultures and DNA extraction

The *s* wild type strain of *Podospora anserina* was used. Senescent cultures were obtained after about 10 cm of vegetative growth of the mycelia. Nucleic acids were extracted according to Lecellier and Silar (13) from degenerative mycelia, about 2 cm before their growth arrest. Aliquots of the extracts corresponding to about 50 ng of mitochondrial DNA were submitted to RNAse treatment prior to PCR amplification.

PCR and sequencing

PCR experiments were performed in a Hybaid programmable thermal reactor as follows: 3 min. initial denaturation at 94°C, 30 cycles: 1 min. at 94°C (denaturation), 5 sec. at 65°C, 1 min. at 5°C below the fusion temperature of the primers (annealing), 5 sec. at 65°C, 1 min. at 72°C (polymerization) and finally one cycle of 10 min. at 72°C (final elongation). Each reaction contains 100 pmole of each synthetic oligonucleotide used as primer, 50 ng of mitochondrial DNA, 10 nmole of the four dNTPs, 5 units of Taq DNA polymerase (Appligene) in a final volume of 50 µl, 10 mM Tris−HCl pH 9.0, 50 mM KCl, 1.5 mM $MgCl_2$, 0.1% Triton X100, 0.02% gelatin. 50 µl of mineral oil are overlayed prior to initiate the program. Amplified DNA fragments were loaded on 1.2% agarose gels containing ethidium bromide. The fragments to be cloned were purified from equivalent low melting agarose gels and ligated into a dephosphorylated linearized Puc 18 vector. After transformation, plasmids are isolated as previously described (14). Sequence determination was carried out using the BRL Sequenase kit; primers were those previously used for PCR amplifications (Fig. 1a).

RESULTS AND DISCUSSION

In order to demonstrate the *in vivo* use of alternate 3′-splice sites, we chose to study two mitochondrial group IB introns, *nad1-i4* and *cox1-i7* of the fungus *Podospora anserina* (Fig. 1a). For both introns, having the particularity of possessing two distinct long ORFs, the secondary structure has predicted a potential alternative

*To whom correspondence should be addressed

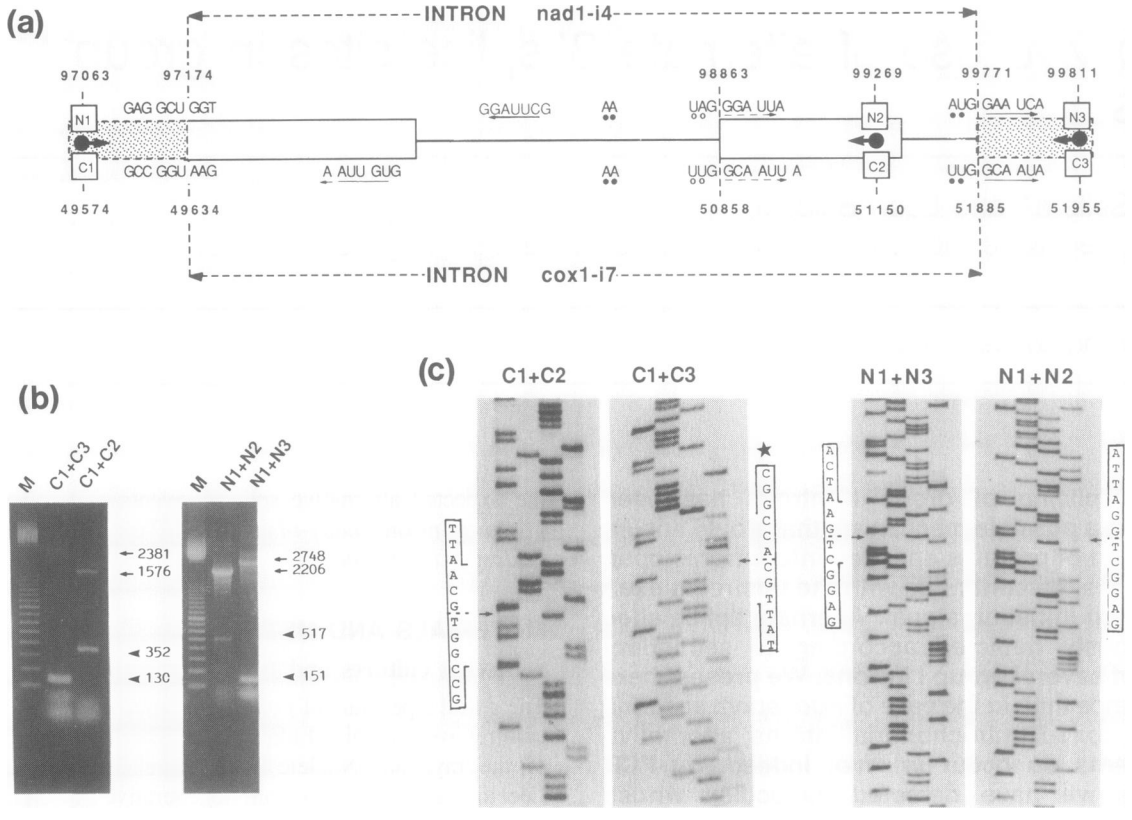

Figure 1. (a) Structure of the *Podospora anserina nad1-i4* and *cox1-i7* introns. *nad1-i4* and *cox1-i7* are two mitochondrial group IB introns already described (15, 16) possessing two distinct long open reading frames (open boxes). The first ones, in frame with their 5′ exon (hatched boxes) loop out of the secondary stucture and exhibit, in both introns, the two LAGLI-DADG dodecapeptides motifs. The downstream ORFs are discontinous with their 5′ exon; that of *nad1-i4* loops out of the P9-1 helix and exhibits a GIY-YIG motif while that of *cox1-i7* lies downstream the P9-1 helix and exhibits a two dodecapeptides motif (15, 16). All the presented coding sequences are divided into triplets. Numbering is according to Cummings *et al.* (22). Alternate 3′-splice sites selection involves the pairing of P10 (←) with proximal P10′ (----→) in place of distal one (→) and of P9-0 (●) with proximal P9-0′ (○) in place of distal one (●). These tertiary interactions (dots under the bases) may involve A.A pairings, known and proposed to occur in the structure of several group I introns (23, 24). (b) PCR experiments. DNA molecules amplified by PCR using N1+N2, C1+C2, N1+N3 and C1+C3 as couples of primers were loaded on 1.2% agarose gels. The 5′ nucleotide of each primer (●→) is numbered and their 5′-3′ orientation is given (see part a). The lower bands (←) were purified, cloned and sequenced. (c) Sequence determination of the cloned junctions corresponding to the use of different 3′-splice sites. N1+N3 and C1+C3 show full splicing of *nad1-i4* and *cox1-i7* introns respectively; N1+N2 and C1+C2 reveal alternative splicing bringing the downstream ORFs of both introns in frame with their upstream exon. The GATC order is respected for each presented sequence. Sequences were initiated from the sense primers (C1 or N1) except in one case (★), primed from the antisense oligonucleotide C3.

splicing that would ensure the expression of the discontinous downstream ORF (15, 16). An increase in the mitochondrial reverse transcriptase activity during the senescent stage of the fungus has allowed the detection, by classical PCR experiments, of cDNA complementary to mature mRNA (17). With the same strategy, we have looked for the low abundant RNA molecules corresponding to both of the expected alternate spliced introns.

PCR experiments, carried out using sense and antisense primers located in the upstream and downstream exons respectively (Fig.1a, primers N1+ N3 and C1+ C3), allowed the amplification, in nucleic acids extracted from young cultures, of mitochondrial DNA molecules containing the whole introns. In senescent cultures, an additional much smaller and more abundant DNA fragment (Fig. 1b), analogous to DNA molecules previously identified as devoid of their intron (17), was also detected. Similarly, the use of pairs of primers located, sense in the upstream exon and antisense in the downstream ORF, allowed the amplification of mitochondrial DNA molecules of the expected size corresponding to the presence of the upstream ORF (Fig. 1a, primers N1+ N2 and C1+ C2). In senescent cultures, a supplementary smaller molecule was also amplified (Fig. 1b), the length of which agreed with the deletion of both the upstream ORF and the catalytic core of the intron.

Cloning and sequencing of the amplified additional molecules were carried out for both introns and both kinds of deleted molecules. The complete deletion of the two introns leads to clear exon-exon junctions (N1+N3, C1+C3; Fig. 1c). The deletion of their upstream half corresponds exactly to the elimination, at the RNA level, of an alternative intron with the same 5′- but an alternate 3′-splice site resulting from the use of the predicted proximal P9-0 and P10 pairings (Fig. 1a). These deletions brought the *nad1-i4* and *cox1-i7* downstream ORFs into juxtaposition with their 5′ exon (N1+N2, C1+C2; Fig. 1c) restoring the frame needed for translation of the ORF-encoded proteins.

Such deleted molecules, easily amplifiable only in senescent cultures of the fungus, are interpreted as the result of the reverse transcription of alternate spliced RNAs, such as the molecules devoid of their entire intron have been previously interpreted as the reverse trancripts of the mature RNAs (17). However, an alternative hypothesis would have been that they result from a deletion, at the DNA level, of the upstream ORF-core region

of the intron. This possibility seems to be very unlikely since no repeated sequence is detectable at the ends of the deleted fragments. Furthermore, it would imply an increase in the frequency of the deletion events in senescent cultures.

Our results, strongly supporting the *in vivo* use of intronic alternate 3'-splice sites, strenghten the idea that, within group I introns containing two distinct ORFs, the downstream one is indeed expressed. The question of the respective function of each ORF within the same intron could be raised. An RNA maturase activity that would guide the intron to become correctly folded (18) could be proposed. In some of the group I introns whose core is separated from the 3'-splice junction by a reading frame, the selection of distal rather than proximal splicing sites would require such an activity encoded into the ORF. The distal choice could be achieved in two different ways leading to the self-regulation of the encoded protein. This protein would act either positively, ensuring the RNA–RNA peripherical interactions that brings the distal 3'-splice site into close proximity with the catalytic core (19) or negatively, masking proximal sites as soon as they are transcribed.

On the basis of the mobility of both introns (manuscript in preparation), an endonuclease activity could also be suggested. Since it still remains unclear whether both RNA maturase and DNA endonuclease activities could be encoded into the same reading frame (20, 21), it could simply be proposed that, in the case of introns containing two distinct ORFs, each activity is encoded in a different reading frame. The maturase activity, proposed for the downstream ORF product of both *nad1-i4* and *cox1-i7* introns agrees with the fact that they are optional (manuscript in preparation) so that, in strains devoid of such an ORF, the intron would not require a protein to become correctly folded.

ACKNOWLEDGEMENTS

We thank A. Sainsard-Chanet and R. Driscoll for critical reading of the manuscript. This work was supported by grants from the Ligue Nationale Franìaise Contre le Cancer, the Association Franìaise contre les Myopathies and the Institut National de la Santé et de la Recherche Médicale (grant no.: 920101).

REFERENCES

1. Lambowitz, A. M. and Belfort, M. (1993) *Annu. Rev. Biochem.* **62**, 587–622.
2. Perlman, P. S. and Butow, R. A. (1989) *Science* **246**, 1106–1109.
3. Hensgens, L.A.M., Bonen, L., de Haan, M.,van der Horst, G. and Grivell, L.A. (1983) *Cell* **32**, 379–389.
4. Cech, T.R. (1990) *Annu. Rev. Biochem.* **59**, 543–568.
5. Michel, F., Jacquier, A. and Dujon, B. (1982) *Biochimie*, **64**, 867–881.
6. Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A. and Scazzocchio, C. (1982) *Nature* **300**, 719–724.
7. Michel, F. , Netter, P., Xu, M-Q. and Shub, D.A. (1990) *Gene & Dev.* **4**, 777–788.
8. Tanner, N. K. and Cech, T. R. (1987) *Biochemistry* **26**, 3330–3340.
9. Burke, J. M. (1989) *FEBS Letters* **250**, 129–133.
10. Michel, F., Hanna, M., Green, R., Bartel, D. P. and Szostak, J. W. (1989) *Nature* **342**, 391–395.
11. Suh, E. R. and Waring, R. B. (1990) *Mol. Cell. Biol.* **10**, 2960–2965.
12. Michel, F. and Westhof, E. (1990) *J. Mol. Biol.* **216**, 585–610.
13. Lecellier, G. and Silar (1994) *Curr. Genet.* **25**,122–123..
14. Stephen, D., Jones, C. and Schofield, J. P. (1990) *Nucleic Acids Res.* **18**, 7463–7464.
15. Cummings, D. J., Domenico, J. M. and Michel, F. (1988) *Curr. Genet.* **14**, 253–264.
16. Cummings, D. J., Michel, F. and Mc Nally, K. L. (1989) *Curr. Genet.* **16**, 407–418.
17. Sainsard-Chanet, A., Begel, O. and Belcour, L. (1993) *J. Mol. Biol.* **234**, 1–7.
18. Burke, J. M. (1988) *Gene* **73**, 273–294.
19. Michel, F. Jaeger, L., Westhof, E., Kuras, R., Tihy, F., Xu, M-Q. and Shub, D. A. (1992) *Genes & Develop.* **6**, 1373–1385.
20. Goguel, V., Delahodde, A. and Jacq, C. (1992) *Mol. Cell. Biol. 12*, 696–705.
21. Lazowka, J., Szczepanek, T., Macadre, C. and Dokova, M. (1992) *C.R. Acad. Sci.* **315**, 37–41.
22. Cummings, D. J., Mc Nally, K. L., Domenico, J. M. and Matsuura, E.T. (1990) *Curr. Genet.* **17**, 375–402.
23. Saenger, W. (1983) Principles of Nucleic Acid Structure, Springer-Verlag, New York.
24. Kim, S-H. and Cech, T. R. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8788–8792.