

U2 and U6 snRNA genes in the microsporidian *Nosema locustae*: evidence for a functional spliceosome

Naomi M. Fast*, Andrew J. Roger¹, Claire A. Richardson and W. Ford Doolittle

Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia B3H 4H7, Canada and

¹Marine Biological Laboratory, 7 MBL Street, Woods Hole, MA 02543, USA

Received March 14, 1998; Revised and Accepted May 7, 1998

DDBJ/EMBL/GenBank accession nos AF053588, AF053589

ABSTRACT

The removal of introns from pre-messenger RNA is mediated by the spliceosome, a large complex composed of many proteins and five small nuclear RNAs (snRNAs). Of the snRNAs, the U6 and U2 snRNAs are the most conserved in sequence, as they interact extensively with each other and also with the intron, in several base pairings that are necessary for splicing. We have isolated and sequenced the genes encoding both U6 and U2 snRNAs from the intracellularly parasitic microsporidian *Nosema locustae*. Both genes are expressed. Both RNAs can be folded into secondary structures typical of other known U6 and U2 snRNAs. In addition, the *N.locustae* U6 and U2 snRNAs have the potential to base pair in the functional intermolecular interactions that have been characterized by extensive analyses in yeast and mammalian systems. These results indicate that the *N.locustae* U6 and U2 snRNAs may be functional components of an active spliceosome, even though introns have not yet been found in microsporidian genes.

INTRODUCTION

The removal of intervening sequences, or introns, from RNA involves two transesterification reactions mediated by the spliceosome. An extremely large ribonucleoprotein complex (comparable in size to the bacterial ribosome), the spliceosome is composed of many proteins and five small nuclear RNAs (snRNAs): U1, U2, U4, U5 and U6 (1). Each of these RNAs associates with its own group of specific proteins, along with the common core Sm proteins, to form a small nuclear ribonucleoprotein (snRNP) complex. The U1 snRNP complex recognizes the 5' splice site, while the U2 complex interacts with the intron branch point, aiding in presenting the nucleophilic adenosine. U4, U5 and U6 enter as a tri-snRNP to complete assembly of the spliceosome, and hence, allow the two transesterification reactions to proceed. The enzymatic activity responsible for the reactions of splicing is generally attributed to the RNA, due to the similarity of the chemical reactions involved in splicing both spliceosomal and autocatalytic Group II introns. However, a substantial catalytic role for the proteinaceous component of the spliceosome cannot be ruled out. These possibilities, along with a comprehensive review of the interactions within the spliceosome, are presented in Nilsen (2) and Staley and Guthrie (3).

Of the snRNA components of the spliceosome, U6 and U2 show the highest degree of sequence conservation (4), where U6 is by far the most highly conserved in sequence, as it base pairs extensively with U4, and also with U2 and the intron itself (5,6). All known U6 snRNA sequences also contain the conserved U6 intramolecular helix, which further constrains the sequence (7). By comparison, the U2 snRNA sequence is less conserved. In addition to the regions of pairing with U6, U2 snRNAs (with the exception of those from the *trans*-splicing trypanosomes) contain the conserved GUAGUA which pairs with the intron branch site, along with several nucleotides in regions of conserved secondary structure, such as those recognized by proteins (8).

In 1996, DiMaria and co-workers described a highly unusual U2 snRNA homolog from the microsporidian parasite *Vairimorpha necatrix* (9). This RNA reportedly had no Sm binding site and no conventional tri-methyl guanosine cap structure at its 5'-end. In addition, the secondary structure they proposed showed many unusual features, including the absence and alteration of otherwise highly conserved structures and the introduction of unique stem-loops.

The relatively odd nature of this snRNA left its role in splicing in some doubt and added to an already lengthy list of uniquely odd features that characterizes the microsporidia. As eukaryotic intracellular parasites, microsporidia survive outside a host cell as hardy spores surrounded by a thick layer of chitin and protein. Spores infect a host cell by an unusual mechanism: an organelle known as the polar tube everts from its tightly wound position within the spore to pierce the host's cell membrane and allow the parasite to inject itself into the host. Microsporidia appear to lack such typical eukaryotic features as mitochondria, stacked Golgi, peroxisomes and 80S ribosomes (microsporidian ribosomes are 70S) (10). Another unique feature is the tiny size of microsporidian genomes. At the extreme, the 2.9 Mb genome of *Encephalitozoon cuniculi* is the smallest known nuclear genome, smaller than the genomes of many bacteria (11). Other microsporidia possess genomes on the order of 5–6 Mb, still extremely small by eukaryotic standards, and thus posing interesting questions with respect to genome organization and the possible reduction or loss of non-coding regions (11,12).

No introns have been found in microsporidian genomes, but fewer than 25 microsporidian gene sequences have so far been reported. For microsporidia and other protists with few sequenced genes, the presence of potentially functional components of the spliceosome may provide the best, albeit indirect, evidence for the occurrence of splicing. Given the unusual structure proposed for the *V.necatrix* U2 snRNA, we chose to search for genes for that snRNA and its partner, U6 snRNA, in a second parasitic

*To whom correspondence should be addressed. Tel: +1 902 494 2968; Fax: +1 902 494 1355; Email: nfast@is2.dal.ca

microsporidian, *Nosema locustae*. Both snRNA genes were found, both are expressed and both RNAs can form the intramolecular and intermolecular secondary structures typical for these RNAs in organisms known to have splicing.

MATERIALS AND METHODS

Sequences reported here have been deposited in GenBank with accession nos AF053588 and AF053589.

Nucleic acid extraction and preparation

Washed *N.locustae* spores were obtained from L.Mearril of the M&R Durango Biocontrol Company (Colorado), and nucleic acid was extracted as follows.

DNA extraction and preparation. Approximately 10^{10} spores were pelleted and resuspended in 500 μ l extraction buffer (50 mM Tris-HCl, pH 7.5, 50 mM EDTA, 3% SDS and 1% 2-mercaptoethanol). Spores were subsequently ground with a mortar and pestle in liquid nitrogen until ~75% of the spores were broken, as determined by light microscopy. After resuspending the ground spores in 1 ml extraction buffer, proteinase K was added to a final concentration of 200 μ g/ml and incubated for 1 h at 50°C. DNA was then ethanol precipitated following standard sequential phenol and chloroform extractions. To remove polysaccharides, DNA was purified by CTAB extraction (13).

RNA extraction. Nucleic acid material was released from *N.locustae* spores as described above. Following sequential phenol and chloroform extractions, RNA was selectively precipitated with LiCl (2 M final concentration).

Genomic library construction and screening

A *N.locustae* genomic library was constructed with the Predigested Zap Express *Bam*HI/CIAP Vector Cloning Kit and the Gigapack II Plus Kit (both from Stratagene) using 6–10 kb DNA fragments generated from *Sau*3A partial digests of 6 μ g *N.locustae* genomic DNA. Size selection was achieved by agarose gel isolation (Prep-A-Gene; Bio-Rad).

The genomic library was screened for the U6 snRNA gene with a 36 nt oligomer probe that was designed to match a highly conserved region of U6, U6-36 (5'-GCAGGGGCCATGCTAATC-TTCTCTGTATAATTCCAA-3'). The probe was labelled by incubating 2 pmol oligomer with 20 U terminal deoxynucleotidyl-transferase (TdT; Promega) in the presence of [α - 32 P]dATP for 30 min at 37°C. Hybridization took place overnight at 60°C in 4 \times SSC, 1% SDS, 0.5% instant skimmed milk powder and 50 μ g/ml poly(A). Prior to autoradiography, membranes were washed in 2 \times SSC and 0.5% SDS twice for 5 min at room temperature and three times for 15 min at 60°C.

The genomic library was screened for the U2 snRNA gene using a 15 nt oligomer designed to hybridize to the highly conserved branch point recognition site of U2, U2-L15 (5'-CAG-ATACTACACTTG-3') (14). The probe was labelled by 3'-tailing as described above. Overnight hybridization took place at 37°C in 5 \times SSC, 5 mM EDTA, 1% SDS and 0.5% instant skimmed milk powder. Washes of 5 \times SSC and 0.5% SDS were 15 min long and repeated twice: first at room temperature, followed by 32°C.

DNA subcloning and sequencing

Genomic clones were digested with restriction endonucleases, blotted and probed to identify independent clones and appropriate restriction fragments for subcloning. Two independent U6 genomic clones and four independent U2 genomic clones were selected for further analysis. In all cases subclones were constructed in pBlueScript SK+. Each subclone was sequenced on both strands by automated LiCor sequencing employing the M13 -20 and M13 Reverse primers.

Northern blot analysis

Approximately 1 μ g *N.locustae* RNA was separated on a 6% polyacrylamide gel and blotted overnight by alkaline transfer with 5 mM NaOH. Exact match forward and reverse direction PCR primers were designed to amplify internal regions of the predicted *N.locustae* U6 and U2 coding regions and radioactive probes were created by amplifying that region from the appropriate genomic subclone in a standard PCR reaction in the presence of [α - 32 P]dCTP (35 cycles of 1 min at 92°C, 1 min at 50°C, 1 min at 72°C). Primer sequences were as follows: U6ampF, 5'-TTAGTTTGGAAACAACACTGAG-3'; U6ampR, 5'-CACCTC-TCAAAGAAAGATG-3'; U2ampF, 5'-AGCCCTCCACCTCTC-AAAGC-3'; U2ampR, 5'-CCAGCATATTCTAGCTCCAAG-3'. Labelling efficiency was verified by hybridizing each probe to a blot containing sample plasmid DNA and then used in the northern hybridization. Blots were hybridized overnight at 65°C in Church's buffer (0.36 M Na₂HPO₄, 0.14 M NaH₂PO₄, 1 mM EDTA and 7% SDS) and washed in 1 \times SSC and 0.5% SDS, once at room temperature and once at 65°C.

RESULTS

The *N.locustae* U6 snRNA gene sequence

Screening of the *N.locustae* genomic library with the probe designed to a highly conserved region of U6 produced several positive signals and subsequent independent genomic clones. Of these, two were chosen for further analysis. Restriction digestion and Southern hybridization with the U6 probe allowed ~500 bp subclones to be generated containing the DNA sequence of interest. Sequencing of the subclones revealed a single sequence corresponding to a clear homolog of the U6 snRNA gene, sharing extensive identity with others. The 5'-end of the molecule was assigned based on secondary structure predictions (discussed below) and confirmed by primer extension (results not shown). The 3'-end assignment is based upon length estimation from northern blot analysis (discussed below).

Figure 1 shows the *N.locustae* U6 gene sequence aligned with known homologs from selected taxa to emphasize overall conservation, particularly in the central region or 'catalytic core' of the molecule. One striking exception is nucleotide T34, a unique base otherwise conserved as an A at that position in almost all other taxa. In fact, this nucleotide is part of the so-called 'phylogenetically invariant' ACAGAGA heptad (nucleotide in question underlined) and its functional significance has been shown by mutational analyses in both mammalian and yeast systems (15–18).

U6 snRNA genes are transcribed by RNA polymerase III. All have been found to have upstream promoter elements that consist of distal, proximal and TA-rich sequence elements (19,20). An intragenic A block element and a downstream B block element

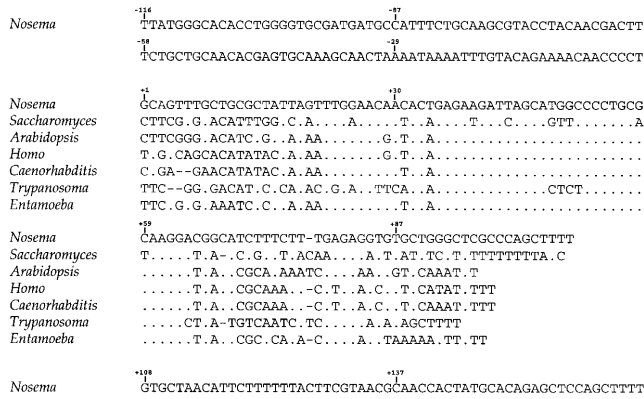


Figure 1. Alignment of a selection of U6 snRNA gene homologs with the *N. locustae* U6 snRNA coding sequence. Numbering of nucleotides is shown for the *N. locustae* sequence, with the predicted 5' start site labelled as +1 (see Materials and Methods). Nucleotides matching those of *N. locustae* are indicated with dots (.) and gaps introduced into the alignment are shown as hyphens (-). Upstream and downstream regions of the *N. locustae* sequence are also shown.

have been characterized in yeast, but appear to be absent from other well-studied (i.e. mostly vertebrate) U6 genes (19). Examining the upstream sequence of the *N. locustae* U6 homolog reveals a prominent TA-rich element, TAAAATAAAA, found at -22 to -31 (Fig. 1), which is consistent in both sequence and position with those previously identified. A possible A block element is the highly conserved AGATTAGCATGG found at position +39 to +50 (Fig. 1), identical in sequence to that recently predicted for *Entamoeba histolytica* (21). Although the proximal sequence element (PSE) should be included in the sequence examined (generally found at approximately -45), it could not be identified, possibly because the PSE does not always have a high degree of phylogenetic conservation (22,23). The presence of the distal sequence element (located at approximately -200 to -300) and the downstream B block element (located at approximately +250) were not addressed in this study, as they are not expected to be included in the examined clones.

The *N. locustae* U2 snRNA gene sequence

The sequence of the U2 snRNA gene does not show the same degree of overall conservation as U6. However, a short region of the molecule corresponding to that which recognizes the intron branch site is extremely well conserved and can be used as a probe (14). Screening of the *N. locustae* genomic library with such a probe (U2-L15) produced many signals and subsequent genomic clones that were analyzed by restriction digestion and Southern blotting. Four different, independent clones were then selected for further subcloning and sequencing. Sequencing both strands of all clones revealed a single sequence with a high degree of sequence conservation in the region surrounding and including the branch point recognition sequence, suggesting that this was a true U2 homolog. Additional evidence for this suggestion comes from the greater degree of overall similarity to the microsporidian *Vnecatrix* U2 homolog compared with others. The conserved, 5'-region of the *N. locustae* U2 sequence is aligned with that from *Vnecatrix* and other representative taxa in Figure 2. Prediction of the *N. locustae* U2 snRNA 5' boundary is based upon the consensus position of the start site relative to the conserved

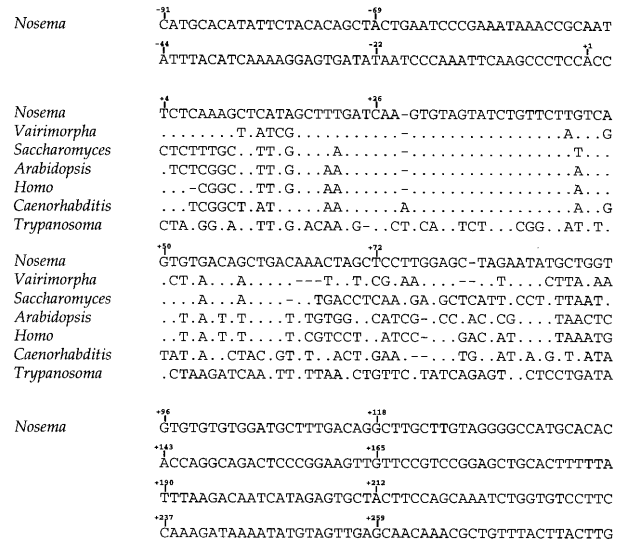


Figure 2. Alignment of the conserved 5'-regions of a selection of U2 snRNA gene homologs with that of *N. locustae*. Numbering and symbols shown are the same as those in Figure 1. Flanking 5'- and 3'-regions of the *N. locustae* U2 snRNA sequence are also presented.

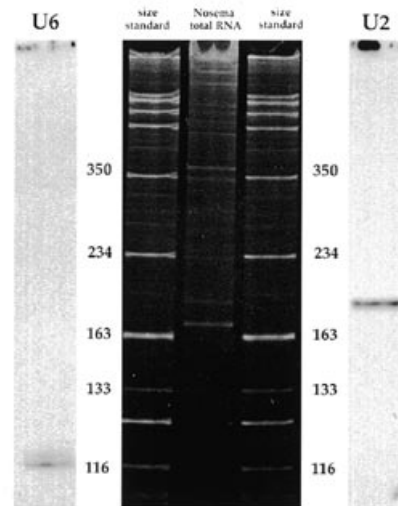


Figure 3. Expression of the *N. locustae* U6 and U2 snRNA genes indicated by northern hybridization analysis. LiCl-precipitated *N. locustae* RNA was run on a 6% polyacrylamide gel, blotted and probed with U6- and U2-specific [α -³²P]dCTP labelled probes (see Materials and Methods).

secondary structure. This predicted 5' boundary (Fig. 2) gains additional support from its close similarity (within 1 nt) to that of the sister microsporidian *Vnecatrix* (9). The 3'-end of the *N. locustae* U2 snRNA (+188) (Fig. 2) is estimated based on a rough size calculation from the northern analysis.

U2 snRNAs characteristically have an Sm binding site [general consensus RR(U)₂-₆RR]. Such a possible site in the *N. locustae* U2 sequence is located between nt 105 and 114 with the sequence GAUGCUUUGA (Fig. 2). This position of the site with respect to the predicted secondary structure of the *N. locustae* U2 snRNA is discussed later. U2 snRNA genes characterized to date also have an upstream TA-rich box, necessary for transcription of the

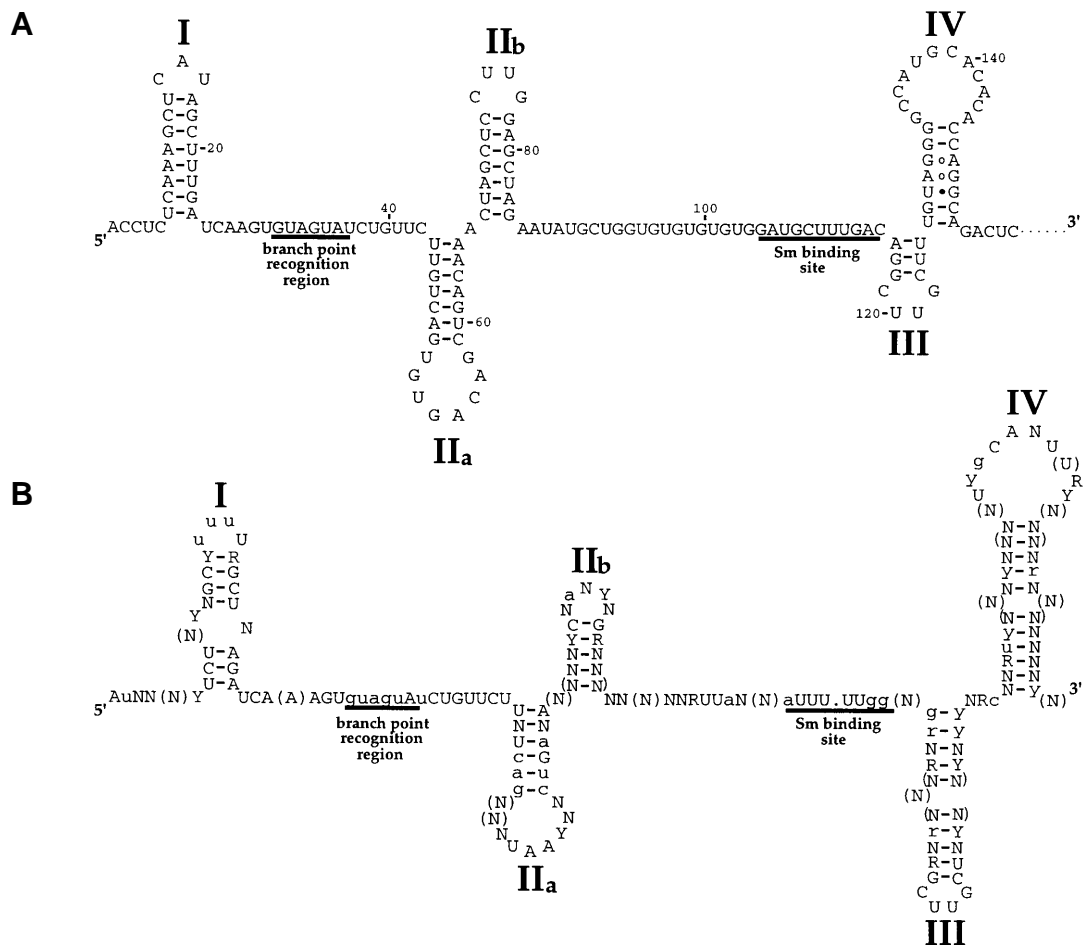


Figure 4. Secondary structure models for the *N. locustae* U2 snRNA and the consensus U2 snRNA. (A) The predicted *N. locustae* secondary structure was generated by the mfold software package and further folded by eye. Stem-loops are labelled based upon their corresponding structures in the consensus U2 folding. (B) U2 consensus secondary structure redrawn from Guthrie and Patterson (8) and based on an alignment of 12 U2 snRNA homologs. Invariant nucleotides are indicated in upper case, while those identical in all sequences except one are shown in lower case.

gene by RNA polymerase II. The TATAA element at -20 to -24 of the *N. locustae* sequence (Fig. 2) is a predicted TATA box.

Both the U6 and U2 snRNA gene sequences of *N. locustae* are clearly homologous to typical spliceosomal snRNAs, rather than to those involved in AT-AC intron splicing (24).

The *N. locustae* U6 and U2 snRNAs are expressed

To determine whether these snRNA genes are expressed by *N. locustae*, northern blot analysis was carried out using probes generated by PCR amplification of internal regions of the two genes. As shown in Figure 3, both RNA species are expressed. The *N. locustae* U6 snRNA length is estimated as 110–115 nt, consistent with other U6 snRNAs. When the degree of migration of the *Vnecatrix* U2 was calculated from the northern analysis shown by DiMaria *et al.* (9) and compared with the *N. locustae* U2 in this study, it was found that both migrate with RNA species of ~180–185 nt in length. This indicates that the *N. locustae* and *Vnecatrix* U2 snRNAs are very similar in length, assuming that such closely related microsporidia share the same cap structure.

Secondary structure modelling

Conservation of the secondary structure of both U6 and U2 snRNAs, as well as the conserved conformation of their intermolecular interaction with one another, is thought to reflect their catalytic significance in the spliceosome (4,5). Computer-assisted free energy minimalization using the mfold server (<http://www.ihc.wustl.edu/~zucker/rna/form1.cgi>), followed by additional folding by eye, generated the secondary structure for the *N. locustae* U2 homolog shown in Figure 4A. This structure is quite similar to that predicted for characterized functional U2s. When compared with the 'consensus' structure shown in Figure 4B, support for the positions of stem-loops I, IIa, IIb, III and IV with respect to the branch point recognition region and the Sm binding site is evident. In addition, *N. locustae* stem-loop I is supported by the first two conserved C-G and U-A pairings at the base along with the C-G and G-C pairings below the loop. Likewise, stem-loop IIa shares the conserved pairings along the stem and has a characteristic sized loop, while stem-loop IIb contains the conserved C-G pair nearest the loop. While an alternative structure for the 5'-region of U2 snRNA has been suggested (25), we could not find evidence for such a potential pairing in *N. locustae*. The 3'-portion of the structure is less similar in sequence, but does maintain the typical stem-loops. The predicted

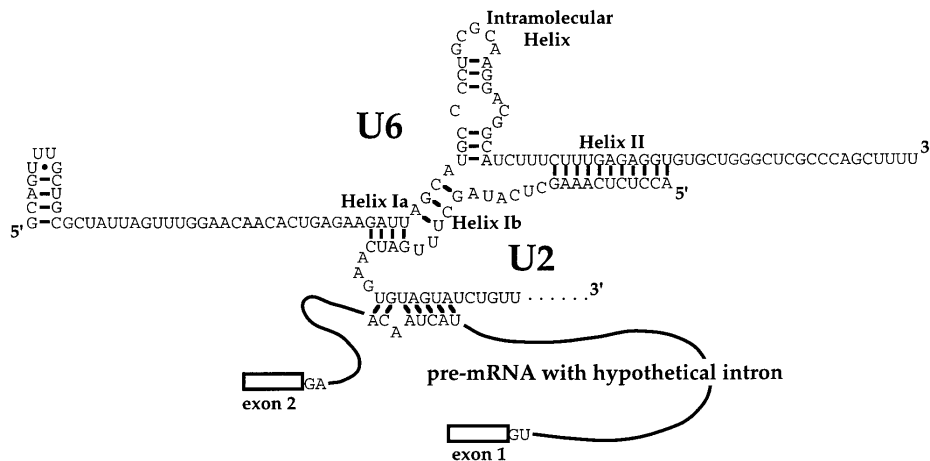


Figure 5. The potential intermolecular interactions between the *N.locustae* U6 and U2 snRNAs. Helices Ia, Ib and II are assigned based on homologous functional pairings characterized in other organisms. The conserved U6 intramolecular helix and 5' stem-loop, along with the potential interaction of U2 snRNA with a canonical intron branch site, are also shown.

Sm binding site gains support by its proximity to a stem-loop that, although shorter in stem length, shares the same loop sequence as stem-loop III, and is labelled as such in Figure 4A. Canonical stem-loop IV contains a YGCA sequence, a probable conserved protein binding motif (26). The UGCA in the large, 12 nt loop of the *N.locustae* U2 secondary structure is identical and, along with the size of the loop, lends credence to this being a true stem-loop IV.

Taken alone, the conservation of the *N.locustae* U2 snRNA structure would imply that it could function in pre-mRNA splicing. Additional evidence is provided by examining the potential intermolecular interactions between the *N.locustae* U2 and U6 snRNAs. Extensive mutational analyses have been carried out in yeast and mammalian systems, providing a framework of functional pairings between the two snRNAs against which the *N.locustae* interactions can be compared (4,15). The predicted folding of the *N.locustae* U6 and U2 snRNAs is depicted in Figure 5.

The conservation of both the *N.locustae* U6 and U2 sequences allows them to form the characteristic helix Ia and helix Ib. Of more functional significance is the formation of a strong helix II. The length and position of the *N.locustae* helix II are consistent with that found in mammals and yeast (27,28), however, the sequences in this region are not conserved among snRNAs. In fact, the sequences co-vary to maintain pairing, so identifying such an interaction indicates its functional value.

Aside from its interaction with U2, U6 also forms a typical intramolecular helix between helix I and helix II (7). This helix is evident in the folded *N.locustae* U6 and is conserved in both structure and many of the conserved nucleotides, including the 5'-most G and C and the 3'-most A within the loop. The 5' stem-loop of U6 snRNAs is also present.

DISCUSSION

We have characterized two integral spliceosomal snRNA components, U6 and U2, from the microsporidian *N.locustae*. Although the U2 snRNA is highly conserved in the 5'-region surrounding the branch point recognition sequence (GUAGUA), the 3'-region of the RNA is less well conserved and is not readily alignable with others. Nevertheless, the region can be folded into the characteristic stem-loops III and IV of known U2 snRNAs, although

the stem lengths and loop sizes are different from those of mammals and yeast. The *N.locustae* stem-loop IV is considerably shorter, but does contain the highly conserved YGCA sequence. This particular motif is found not only in yeast and human U2 stem-loop IV, but also in the U1 stem-loop II (8). Both of these loops are known to bind members of the U1A-U2B'' protein family (26,29), so such a motif in the *N.locustae* loop not only supports this structure, but also suggests that such a splicing protein factor could be present and functioning in *N.locustae*. The proposed stem-loop III has the characteristic loop sequence CUUG, but the stem length is much reduced compared with others. This alteration is probably not functionally significant, however, since stem-loop III can be removed from yeast without abolishing splicing activity (30). The position of the Sm binding site of the *N.locustae* U2 also supports the placement of this stem-loop.

In sequence, the *N.locustae* U2 snRNA is most like that of the related microsporidian *V.necatrix*. This similarity is most striking at the 5'-end of the RNAs, with identical nucleotide sequences in the regions proposed here to interact with U6. Downstream of this region, however, the similarity between the two sequences is much less, although the *V.necatrix* U2 can fold similarly to the structure presented here. [We favour this alternative folding, including a putative Sm binding site, which is described (not drawn) by DiMaria *et al.* (9).]

Assuming that the *N.locustae* U2 functions in the spliceosome, we predicted that we should also find a U6 snRNA which could pair with it. This prediction was accurate: *N.locustae* possesses a clear U6 snRNA homolog that is highly conserved in sequence.

One exception to this conservation is U34, an A in other U6 snRNAs and part of the highly conserved (often called 'phylogenetically invariant') ACAGAGA sequence, where position 34 is underlined. Due to the conserved nature and supposed functional significance of this sequence, it has been the target of considerable mutagenesis, both *in vitro* and *in vivo*, in yeast, as well as in human cell lines. In *in vivo* studies the A→U mutation is lethal in both yeast and human cell systems (16,18). The *in vitro* studies generally support this conclusion, although the mutant phenotypes do display variable degrees of splicing (50–100%) (15,17). In yeast it has been proposed that the ACA of the conserved heptad interacts with a UGU at the 5' splice site located 1 nt downstream of the canonical GU (31,32). As no introns have been characterized in the

microsporidia, it could be hypothesized that a currently undiscovered intron contains a trinucleotide other than UGU (for example AGU) that will satisfy the pairing. It is noteworthy that the intronic UGU sequence is not highly conserved outside yeast, raising the possibility that this region of the snRNA has a different or possibly redundant function, not clearly understood.

Both the U6 and the U2 snRNA genes of *N.locustae* are expressed. Further support for their functional significance is their potential pairing with each other. There are three regions of defined pairings between functional U6 and U2 snRNAs that can also be postulated for the *N.locustae* U6 and U2, as illustrated in Figure 5. These pairings have been identified as functional units in yeast and mammals with mutational analyses (4,15), although there is evidence for redundancy of the helices (33).

As depicted in Figure 5, the folding of the *N.locustae* U6 and U2 snRNAs and their potential to interact with an intron implies that they could be part of a functional spliceosome, removing introns from pre-mRNA. No such spliceosomal introns have yet been found in microsporidia and their existence in the genome of *N.locustae* would have to be at a very low density (introns per kb) based upon the absence of introns found in the ~10 kb of protein coding DNA sequenced in our laboratory. Considering the tiny nature of microsporidian genomes, a reduction in the number of non-coding elements such as introns may be a means to achieve such compaction. We think it unlikely, however, that the splicing machinery would be maintained if microsporidia have no introns to splice.

Another possibility is that the snRNAs have been maintained to act in spliced leader *trans*-splicing, a rare phenomenon that, to date, has only been observed in euglenozoans, nematodes and platyhelminths (34–36). Currently, there is no supportive evidence from the admittedly limited data about microsporidian gene organization. However, the possibility cannot be ruled out and further data on microsporidian mRNAs may be helpful in testing for the presence of *trans*-splicing in these organisms.

For the past decade, microsporidia have been considered 'ancient' (deeply diverging) eukaryotes (37). This notion was supported by their apparent lack of typical eukaryotic features, such as mitochondria, stacked Golgi, peroxisomes and spliceosomal introns, features also lacking in other putatively deep branching protists like the diplomonad *Giardia lamblia* (10), where spliceosomal snRNAs have been searched for unsuccessfully (38). More recently, the early divergence of the microsporidia has been called into question: genes of 'mitochondrial origin' have been found in microsporidian genomes (39,40) and protein phylogenies inferred from tubulin (41,42), HSP70 (39,40) and the largest subunit of RNA polymerase II (J.Logsdon, R.Hirt, T.M.Embley and W.F.Doolittle, personal communication) strongly suggest a recent divergence, with the microsporidia branching from within or as a sister group to the Fungi. The fact that *N.locustae* possesses two core spliceosomal components, U6 and U2 snRNAs capable of interacting with each other in a functional manner, is consistent with this phylogenetic reassignment. We suggest that microsporidia lost most (but probably not all) of their spliceosomal introns during radical genome size reduction, accompanying the adoption of an intracellular parasitic lifestyle.

ACKNOWLEDGEMENTS

We are grateful to D.Price, J.Norman and Y.Watanabe for materials and assistance and to G.McFadden for the use of his

laboratory facilities. Thanks are due to L.Mearril for providing *N.locustae* spores. We thank J.Logsdon and P.Keeling for providing comments on the manuscript and M.Schnare for his input throughout the course of this study. This work was supported by an operating grant from the Medical Research Council of Canada. N.M.F. is the recipient of post-graduate scholarships from the Natural Sciences and Engineering Research Council of Canada and the Killam Trust. W.F.D. is a fellow of the Canadian Institute for Advanced Research.

REFERENCES

- Will,C.L. and Luhrmann,R. (1997) *Curr. Opin. Cell Biol.*, **9**, 320–328.
- Nilsen,T.W. (1998) In Simons,R. and Grunberg-Manago,M. (eds), *RNA Structure and Function*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 279–307.
- Staley,J.P. and Guthrie,C. (1998) *Cell*, **92**, 315–326.
- Madhani,H.D. and Guthrie,C. (1992) *Cell*, **71**, 803–817.
- Madhani,H.D. and Guthrie,C. (1994) *Annu. Rev. Genet.*, **28**, 1–26.
- Brow,D.A. and Guthrie,C. (1988) *Nature*, **334**, 213–218.
- Wolff,T. and Bindereif,A. (1993) *Genes Dev.*, **7**, 1377–1389.
- Guthrie,C. and Patterson,B. (1988) *Annu. Rev. Genet.*, **22**, 387–419.
- DiMaria,P., Palic,B., Debrunner-Vossbrinck,B.A., Lapp,J. and Vossbrinck,C.R. (1996) *Nucleic Acids Res.*, **24**, 515–522.
- Cavalier-Smith,T. (1993) *Microbiol. Rev.*, **57**, 953–994.
- Biderre,C., Pages,M., Metenier,G., Canning,E.U. and Vivares,C.P. (1995) *Mol. Biochem. Parasitol.*, **74**, 229–231.
- Keeling,P.J. and McFadden,G.I. (1998) *Trends Microbiol.*, **6**, 19–23.
- Clark,C.G. (1992) In Lee,J.J. and Soldo,A. (eds), *Protocols in Protozoology*. Allen Press, Lawrence, KS, Vol. 1, pp. D3.1–D3.2.
- Ares,M., Jr (1986) *Cell*, **47**, 49–59.
- Wolff,T., Menssen,R., Hammel,J. and Bindereif,A. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 903–907.
- Datta,B. and Weiner,A.M. (1993) *Mol. Cell. Biol.*, **13**, 5377–5382.
- Fabrizio,P. and Abelson,J. (1990) *Science*, **250**, 404–409.
- Madhani,H.D., Bordonne,R. and Guthrie,C. (1990) *Genes Dev.*, **4**, 2264–2277.
- Eschenlauer,J.B., Kaiser,M.W., Gerlach,V.L. and Brow,D.A. (1993) *Mol. Cell. Biol.*, **13**, 3015–3026.
- Brow,D.A. and Guthrie,C. (1990) *Genes Dev.*, **4**, 1345–1356.
- Miranda,R., Salgado,L.M., Sanchez-Lopez,R., Alagon,A. and Lizardi,P.M. (1996) *Gene*, **180**, 37–42.
- Li,J.M., Haberman,R.P. and Marzluff,W.F. (1996) *Mol. Cell. Biol.*, **16**, 1275–1281.
- Thomas,J., Lea,K., Zucker-Aprison,E. and Blumenthal,T. (1990) *Nucleic Acids Res.*, **18**, 2633–2642.
- Tarn,W.Y. and Steitz,J.A. (1997) *Trends Biochem. Sci.*, **22**, 132–137.
- Zavanelli,M.I., Britton,J.S., Igel,A.H. and Ares,M., Jr (1994) *Mol. Cell. Biol.*, **14**, 1689–1697.
- Tang,J., Abovich,N. and Rosbash,M. (1996) *Mol. Cell. Biol.*, **16**, 2787–2795.
- Datta,B. and Weiner,A.M. (1991) *Nature*, **352**, 821–824.
- Wu,J.A. and Manley,J.L. (1991) *Nature*, **352**, 818–821.
- Scherly,D., Boelens,W., Dathan,N.A., van Venrooij,W.J. and Mattaj,I.W. (1990) *Nature*, **345**, 502–506.
- Shuster,E.O. and Guthrie,C. (1988) *Cell*, **55**, 41–48.
- Lesser,C.F. and Guthrie,C. (1993) *Science*, **262**, 1982–1988.
- Kandels-Lewis,S. and Seraphin,B. (1993) *Science*, **262**, 2035–2039.
- Field,D.J. and Friesen,J.D. (1996) *Genes Dev.*, **10**, 489–501.
- Bonen,L. (1993) *FASEB J.*, **7**, 40–46.
- Nilsen,T.W. (1993) *Annu. Rev. Microbiol.*, **47**, 413–440.
- Nilsen,T.W. (1995) *Mol. Biochem. Parasitol.*, **73**, 1–6.
- Vossbrinck,C.R., Maddox,J.V., Friedman,S., Debrunner-Vossbrinck,B.A. and Woese,C.R. (1987) *Nature*, **326**, 411–414.
- Niu,X.H., Hartshorne,T., He,X.Y. and Agabian,N. (1994) *Mol. Biochem. Parasitol.*, **66**, 49–57.
- Germot,A., Philippe,H. and Le Guyader,H. (1997) *Mol. Biochem. Parasitol.*, **87**, 159–168.
- Hirt,R.P., Healy,B., Vossbrinck,C.R., Canning,E.U. and Embley,T.M. (1997) *Curr. Biol.*, **7**, 995–998.
- Keeling,P.J. and Doolittle,W.F. (1996) *Mol. Biol. Evol.*, **13**, 1297–1305.
- Eldfind,T.D., Li,J., Visvesvara,G.S., Vodkin,M.H., McLaughlin,G.L. and Katiyar,S.K. (1996) *Mol. Phylogenet. Evol.*, **5**, 359–367.