Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 852–856, February 1999
Biochemistry

U1 small nuclear RNA and spliceosomal introns in *Euglena gracilis*

DAVID G. BRECKENRIDGE, YOH-ICHI WATANABE, SPENCER J. GREENWOOD, MICHAEL W. GRAY*, AND MURRAY N. SCHNARE

Program in Evolutionary Biology, Canadian Institute for Advanced Research, Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia B3H 4H7, Canada

Edited by Robert Haselkorn, The University of Chicago, Chicago, IL, and approved December 14, 1998 (received for review October 29, 1998)

ABSTRACT **In the flagellated protozoon *Euglena gracilis*, characterized nuclear genes harbor atypical introns that usually are flanked by short repeats, adopt complex secondary structures in pre-mRNA, and do not obey the GT-AG rule of conventional cis-spliced introns. In the nuclear fibrillarin gene of *E. gracilis*, we have identified three spliceosomal-type introns that have GT-AG consensus borders. Furthermore, we have isolated a small RNA from *E. gracilis* and propose, on the basis of primary and secondary structure comparisons, that it is a homolog of U1 small nuclear RNA, an essential component of the cis-spliceosome in higher eukaryotes. Conserved sequences at the 5′ splice sites of the fibrillarin introns can potentially base pair with *Euglena* U1 small nuclear RNA. Our observations demonstrate that spliceosomal GT-AG cis-splicing occurs in *Euglena*, in addition to the nonconventional cis-splicing and spliced leader trans-splicing previously recognized in this early diverging unicellular eukaryote.**

In most eukaryotes, maturation of mRNA precursors requires cis-splicing, the precise removal of intron sequences from pre-mRNA (1–3). Consensus sequences required for splicing are located at the 5′-end (the 5′-splice site) and the 3′-end (the 3′-splice site) of cis-spliced introns (1–3). Almost all introns have GT-AG splice junctions; i.e., the first two residues at the 5′-splice site are G and T, and the last two residues at the 3′-splice site are A and G (1–3). Cis-splicing also involves the participation of several small nuclear RNAs (snRNAs), designated U1, U2, U4, U5, and U6, which exist in the form of ribonucleoprotein (RNP) particles (2–7). Together with protein factors, these snRNPs assemble on pre-mRNA to form the spliceosome, within which the splicing reactions occur (2–7). Several base pairing interactions (snRNA:snRNA and snRNA:mRNA) have been identified as being important in spliceosome assembly and/or function (2, 8–10). For example, base pairing between a sequence near the 5′-end of U1 snRNA and the 5′-splice site in pre-mRNA plays an important role in 5′-splice site selection (11).

In the flagellated protozoon *Euglena gracilis*, three nuclear genes have been characterized to date, and all three contain atypical introns that lack the characteristic features of spliceosomal introns. These *Euglena* introns are very short and do not possess GT-AG consensus borders (12–16). In many cases, the introns are flanked by short direct repeats, making their precise boundaries difficult to determine (13–16). An interesting feature of the *Euglena* introns is their potential to form complex but stable secondary structures at the level of pre-mRNA (13–16). In all cases, the proposed secondary structures would bring the 5′- and 3′-ends of the intron together (13–16), a function that normally is carried out by the spliceosomal machinery in other eukaryotes. The observed abnormalities in these introns have prompted suggestions that, in *E. gracilis*, maturation of pre-mRNA involves a novel cis-splicing mechanism of unknown origin (13–16).

Another atypical feature of pre-mRNA processing in *E. gracilis* is that many of the mRNAs in this organism obtain their 5′-terminal sequences via trans-splicing (17, 18). The 5′-terminal spliced leader (SL) sequence initially is located at the 5′-end of a small SL RNA, which contains a typical 5′-splice site (17, 18). SL RNA trans-splicing of pre-mRNA also occurs in trypanosomatid protozoa, which are distantly but specifically related to *Euglena*, and in nematodes and trematodes (19–27). Spliceosomal snRNPs participate in SL RNA trans-splicing in a manner analogous to cis-splicing; however, U1 snRNP is thought not to be involved in trans-splicing (20, 21, 28, 29). No cis-spliced introns or U1 snRNA have yet been discovered in trypanosomes (19–23, 30). Nematodes do have a U1 snRNA (31, 32); however, their pre-mRNAs also undergo spliceosomal cis-splicing (21–24).

In this report, we document the isolation and characterization of a U1 snRNA from *E. gracilis*. In addition, we describe three introns in a *Euglena* nuclear gene that contain the GT-AG consensus boundary sequences expected for spliceosomal cis-spliced introns. The 5′ terminus of *Euglena* U1 snRNA has the potential to base pair with conserved sequences at the 5′ splice sites of these introns. Together, these data indicate that spliceosomal cis-splicing occurs in *Euglena*.

MATERIALS AND METHODS

Detailed procedures for RNA isolation, PCR amplification, RNA and DNA sequencing, and hybridization analysis are described elsewhere (33–35). Typical sequencing gels contained 7 M urea, although some chemical digests of 3′-end-labeled RNA and 5′-end-labeled DNA also were resolved in gels that contained 7 M urea plus 40% formamide (deionized), which helped to suppress band compressions (36). PCR amplification cycles were as follows: 94°C/5 min (1×); 94°C/30 sec, 55°C/30 sec, and 72°C/30 sec (30×); and 72°C/7 min (1×). The oligonucleotides used for PCR amplification corresponded to positions 7–27 and the complement of positions 150–170 of the U1 snRNA sequence. The latter oligonucleotide also was used for RT sequencing. Modified nucleotide analysis was performed as described (37, 38). All positions in the U1 snRNA sequence were determined by at least two independent approaches that included chemical/enzymatic sequencing of end-labeled RNA, reverse transcriptase sequencing of RNA, chemical/enzymatic sequencing of a PCR

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

This paper was submitted directly (Track II) to the *Proceedings* office.
Abbreviation: snRNA, small nuclear RNA; RNP, ribonucleoprotein; SL, spliced leader; TMG, N^2,N^2,7-trimethylguanosine.
Data deposition: The sequences reported in this paper have been deposited in the GenBank database [accession nos. U57366 (U1 snRNA) and AF102879 (fibrillarin genomic fragment/PCR product)].
*To whom reprint requests should be addressed. e-mail: M.W.Gray@Dal.Ca.

product, and terminal nucleotide analysis. The secondary structure diagram was generated with the computer program XRNA, developed by B. Weiser and H. Noller (University of California, Santa Cruz) and recently made available on the internet (ftp://fangio.ucsc.edu/pub/XRNA).

The 5′-region of the *E. gracilis* fibrillarin gene was amplified by using primers (GIBCO/BRL) based on the cDNA sequence (Y.W. and M.W.G.; GenBank accession no. AF110181). The 3′ primer, 5′-GGACTTTTGCACCGGGAGCAACC-3′, was used in conjunction with a 5′ primer, 5′-GAtcgGAtCcAT-GAAAGGCGACTTCGGA-3′ (lowercase nucleotides do not match the cDNA sequence). PCR mixtures using *Taq* polymerase and reaction buffer (GIBCO/BRL) included 0.1–1 μg of total *E. gracilis* DNA, 10 pmol of each primer, and 0.2 mM concentrations of each of the four dNTPs. $MgCl_2$ was added to a final concentration of 2.5 mM when the reactions reached 80°C. Cycle parameters were 94°C/2 min (1×); 94°C/45 sec, 60°C/30 sec, and 68°C/3 min (35×); and 72°C/7 min (1×). PCR products were resolved in a 1.5% agarose gel, were recovered by using the GlassMax DNA Isolation Spin Cartridge System (GIBCO/BRL), and were ligated into the pT7Blue T-Vector (Novagen). *E. coli* DH5α cells were transformed following the method of Inoue *et al.* (39), and plasmids were prepared by using the alkali lysis method (40). Both strands of three independent clones were sequenced by using the fmol DNA Cycle Sequencing System (Promega).

RESULTS AND DISCUSSION

E. gracilis U1 snRNA initially was identified as a contaminant of U3 small nucleolar RNA that was prepared by electrophoresis in 6% polyacrylamide/7 M urea gels (34). However, in 10% polyacrylamide/7 M urea gels, the U1 snRNA migrates as though it is considerably larger than the 180-nt U3 small nucleolar RNA (34). This artifactual mobility of U1 snRNA (which is 169 or 170 nt; see below) in 10% gels permitted its separation from U3 small nucleolar RNA and other contaminating nucleic acids (see Fig. 1). The unusually slow mobility of U1 snRNA in high percentage gels is indicative of stable secondary structure that persists in the presence of urea (41). Accordingly, we observed band compressions in sequencing gels, a common problem associated with sequencing of U1 snRNAs (42, 43), that were consistent with the proposed secondary structure (Fig. 2*A*). In fact, the structure of stem/loop III (positions 89–115; Fig. 2*A*) was so stable that, with the exception of a few residues at its base and in the loop, it was completely resistant to cleavage by RNases and alkali during enzymatic sequence analysis. Consequently, we were only able to determine the number of C residues (positions 97–100) by sequencing a PCR product.

In 10% gels, *E. gracilis* U1 snRNA often appears as two bands (Fig. 1), the smaller of which (169 nt) is missing one of the two 3′-terminal U residues present in the larger species (170 nt). 5′-end-labeling with polynucleotide kinase was enhanced greatly by prior treatment of the RNA with tobacco acid pyrophosphatase, indicating that *Euglena* U1 snRNA has a 5′-terminal cap, which tentatively was identified as N^2,N^2,7-trimethylguanosine (TMG) by immunoprecipitation with an anti-TMG antibody (Fig. 1). This cap identification is considered tentative because, in our hands, the anti-TMG antibody (44) also reacts to a significant extent with RNAs containing 7-methylguanosine. Enzymatic sequencing (see Fig. 3) suggested (45) the presence of a few post-transcriptionally modified residues at internal positions in the sequence. 5′-end-labeled partial alkali digestion products corresponding to each of these positions were isolated, and their labeled 5′-terminal nucleotides were identified by thin layer chromatography. This analysis verified the presence of $O^{2'}$-methyluridine at position 13 and $O^{2'}$-methyladenosine at positions 4 and 68. All U residues in the molecule were cleaved in the hydrazine reaction

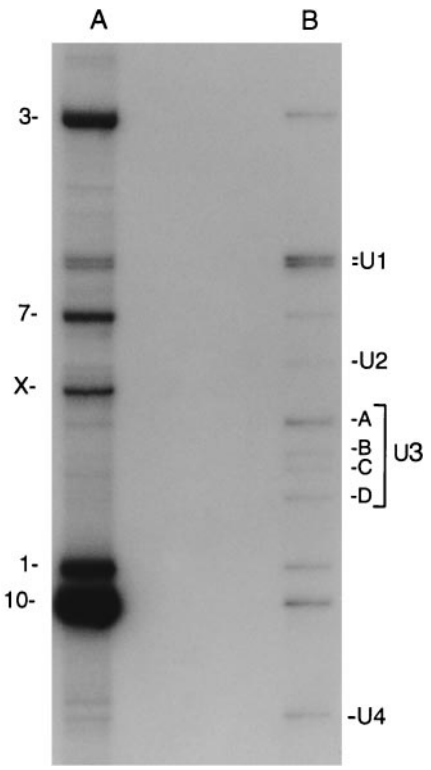
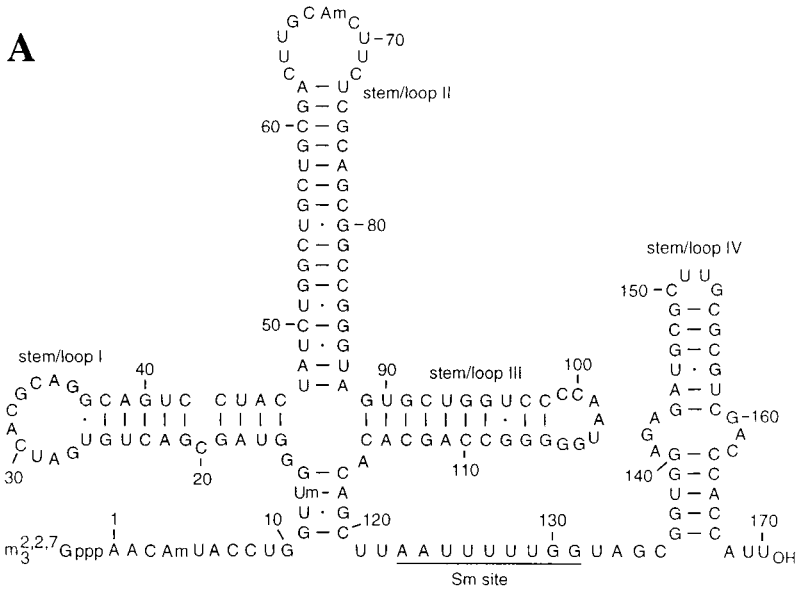


FIG. 1. Immunoprecipitation of TMG-capped *Euglena* small RNAs. Samples were 3′-end-labeled with [5′-^{32}P]pCp and RNA ligase and were resolved in a 10% polyacrylamide/7 M urea sequencing gel. (*A*) Total RNA. The most abundant labeled bands in this sample [1 (163 nt), 3 (350 nt), 7 (235 nt), and 10 (164 nt)] correspond to components of the fragmented large subunit rRNA (33). Band X (181 nt) was generated by ligation of large subunit rRNA species 12 to the 3′-end of large subunit rRNA species 14 (33). (*B*) Anti-TMG immunoprecipitate. Chemical sequencing revealed that the mAb enriched for U1 (see *Results and Discussion*), U2, and U4 (identified by sequence analysis and comparison with homologous sequences; M.N.S., unpublished results) snRNAs as well as four bands (A–D) containing variant forms of U3 small nucleolar RNA (34).

during chemical sequencing, indicating that none of these is modified to pseudouridine (Ψ); however, we cannot rule out the presence of other modified nucleoside constituents or partial conversion of U to Ψ.

The *Euglena* sequence presented here was identified as U1 snRNA by comparison with known sequences/structures as follows. The sequence can be modeled to fit a phylogenetically conserved and experimentally supported secondary structure (43, 46, 47). This model (Fig. 2*A*) contains a 4-bp long-range interaction enclosing a portion of the sequence that is represented by three hairpin structures (stem/loops I-III). A fourth hairpin (stem/loop IV) is located near the 3′-end of the sequence. This structure is possible even though the *Euglena* primary sequence has diverged substantially in helical regions compared with known U1 snRNA sequences from other organisms (48, 49). On the other hand, single-strand regions have been highly conserved (Fig. 2*B*) and can be correlated with known functional sites. The 5′-terminal single-strand region of U1 snRNA is known to pair with the 5′-splice site in pre-mRNAs (11). The loop regions of stem/loops I and II contain the major determinants for the binding of the mammalian U1 snRNP-specific proteins U1–70K and U1-A, respectively (4). The single-strand sequence at positions 123 to 131 represents the binding site for the Sm core protein complex, common to all spliceosomal snRNPs (4). Finally, the post-transcriptional modification $O^{2'}$-methyladenosine (residue 68) is also present at the same position in some other U1 snRNAs (42). With the U1 snRNA PCR product as a probe,

A



B

```
Positions      (2-16)              (26-37)          (63-72)

E. gracilis    ACAUACCUGGUUGGG / UGAUCACGCAGG / CUUGCACUUC
               ||  |||||||  ||| / |||||||| ||| /  |||||||  |
H. sapiens     ACUUACCUGGCAGGG / UGAUCACGAAGG / AUUGCACUCC

Positions      (96-108)         (116-134)              (147-156)

E. gracilis    UCCCCAAUGGGGG / ACAGCUUAAUUUUUGGUAG / GCGCUUGCGC
                |||  ||||  ||| / ||  ||  |||||| |||||| / |||  |  ||||
H. sapiens     CCCCAAAUGUGGG / ACUGCAUAAUUUGUGGUAG / GCGUUCGCGC
```

FIG. 2. (*A*) Primary sequence and secondary structure of *E. gracilis* U1 snRNA. Am, $O^{2'}$-methyladenosine; Um, $O^{2'}$-methyluridine. (*B*) Alignment of regions of *E. gracilis* and *Homo sapiens* (42) U1 snRNAs. Only regions that can be aligned unambiguously are shown (position numbers are based on the *E. gracilis* sequence). Identities between the two sequences are indicated (|).

Southern hybridization indicated that there is a single copy of the *Euglena* U1 snRNA gene. Only mature U1 snRNA was detected by Northern hybridization using a U1 snRNA-specific oligonucleotide as a probe (data not shown).

At first glance, the discovery of U1 snRNA in *Euglena* is surprising because all of the cis-spliced introns so far characterized in this organism appear to be nonconventional (13–16). However, there are several plausible explanations for the existence of U1 snRNA in *Euglena*. First, it is possible that U1 snRNA is involved in the excision of the atypical introns previously characterized in *Euglena*. This explanation seems unlikely because the 5′-terminus of the *Euglena* U1 snRNA cannot form extensive base pair interactions with the 5′-splice sites of any of these atypical introns. Alternatively, it is conceivable that the U1 snRNA is involved in the trans-splicing of the SL sequence to the 5′-termini of pre-mRNAs in *Euglena*; however, a number of studies have suggested that U1 snRNA may not be required for trans-splicing in other systems (20, 21, 28, 29). Finally, our discovery of a U1 snRNA in *Euglena* could be indicative of the presence of U1 snRNA-dependent spliceosomal cis-spliced introns in *Euglena*. Although no spliceosomal introns have yet been reported in *Euglena,* only three nuclear protein-coding genes (*gapC*, *rbcS*, and *lhcp2*) have been examined (12–16).

To initiate a search for spliceosomal introns in *E. gracilis*, we took advantage of a partial cDNA sequence encoding the *Euglena* fibrillarin protein (Y.W., unpublished results). When primers specific for the 5′ region of the cDNA (corresponding to amino acid positions 1–124 in the protein sequence) were used to amplify from total *Euglena* DNA, a product considerably larger than expected was obtained, suggesting that one or more introns were present in this region of the gene. We therefore cloned and sequenced three copies of the genomic PCR product, obtaining the consensus sequence shown in Fig. 4. Comparison of this genomic sequence with that of the cDNA revealed the presence of three short introns of lengths 46, 68, and 44 bp (Fig. 4). In contrast to the introns found in other *Euglena* nuclear genes, the three fibrillarin introns could not be modeled to form stable secondary structures that would juxtapose their splice junctions. Furthermore, all three introns contain GT-AG consensus borders, suggesting that they are spliceosomal (Fig. 4). The genomic PCR product was used as a probe in Southern blot analysis, which established that there is a single copy of the fibrillarin gene in *E. gracilis*, thereby ruling out the possibility that the PCR product represents a pseudogene (data not shown).

Comparison of the sequence around the 5′-splice sites of the three fibrillarin introns demonstrates that the first six nucleotides of introns A and B are identical but differ from that of intron C by A/G and A/T base changes at positions +3 and +6, respectively, relative to the 5′-splice junction (Fig. 5). As shown in Fig. 5, the 5′-splice sites of all three introns in the fibrillarin pre-mRNA could potentially base pair with the 5′-terminal sequence of *Euglena* U1 snRNA. Although, com-

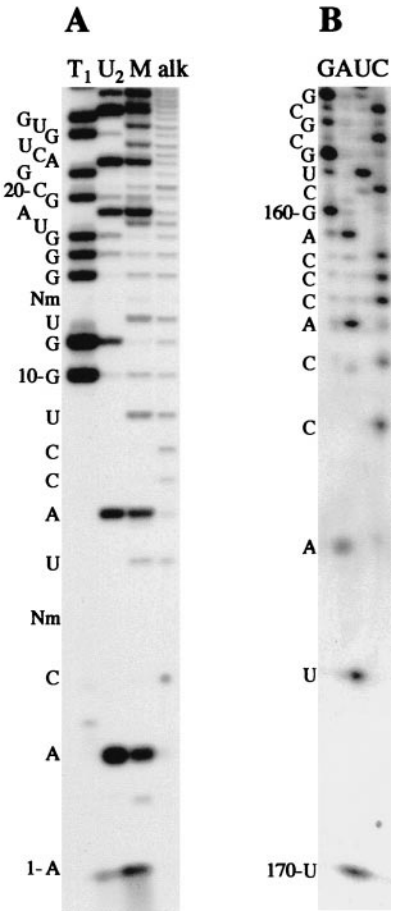


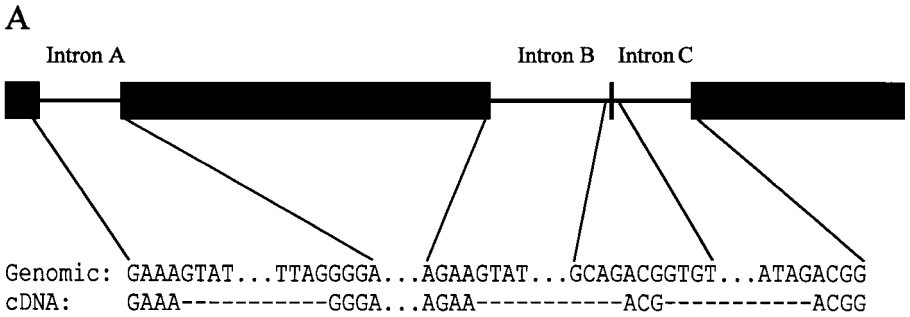
FIG. 3. Autoradiograms of representative sequencing gels (20% acrylamide/7 M urea). (*A*) Enzymatic sequencing of 5′-end-labeled *E. gracilis* U1 snRNA. T_1, RNase T_1 (cleaves after G); U_2, RNase U_2 (cleaves after A); M, RNase Phy M (cleaves after A>U); alk, alkali ladder; Nm, $O^{2'}$-methylnucleoside. (*B*) Chemical sequencing of the long version (170 nt) of *E. gracilis* U1 snRNA. Minor bands are the result of cross-contamination with the short version (169 nt) of this RNA species.

pared with other eukaryotes, the *Euglena* U1 snRNA has a single base change (U to A at position 4; Fig. 2*B*) in this 5′-terminal region, it still has the potential to form a strong interaction with each intron. The proposed U1 snRNA:5′-splice site interactions strongly suggest that the *Euglena* U1 snRNP can recognize each 5′-splice site and that the introns are excised from the pre-mRNA in a manner analogous to the spliceosomal reactions observed in other eukaryotes. A similar base pairing interaction is also possible between the 5′-terminus of U1 snRNA and the 5′-splice site of the SL RNA (Fig. 5).

```
U1 snRNA      3′...UC CAUACAA    5′
                  |  |||||
Intron A      5′...AA/GUAUGAA... 3′

U1 snRNA      3′...UC CAUACAA    5′
                  |  ||||| |
Intron B      5′...AA/GUAUGAU... 3′

U1 snRNA      3′...UC CAUACAA    5′
                  | ||·||||
Intron C      5′...CG/GUGUGUU... 3′

U1 snRNA      3′...UC CAUACAA    5′
                  | ||||  |
SL RNA        5′...CG/GUAUAUA... 3′
```

FIG. 5. Potential interactions between the 5′-terminal region of U1 snRNA and the 5′-splice sites of the three fibrillarin pre-mRNA introns and between the 5′-terminal region of U1 snRNA and the 5′-splice site of the SL RNA (17, 18). The cap at the 5′-end of U1 snRNA is not shown. The exon–intron boundaries are indicated by a slash (/), with the conserved GU at the 5′-end of each intron underlined.

Introns B and C are separated by a microexon of only three nucleotides (Fig. 4), which is among the shortest of recorded exons. Three-nucleotide exons (50, 51) and even a two-nucleotide exon (52) separated by spliceosomal introns have been reported. Dominski and Kole (53) have shown that exon skipping is prevented by strengthening 5′-splice sites down-



B

```
ATGAAAGGCG ACTTCGGAAA gtatgaaaac ttcagtttcg ttgcctttaa aattattcca  Intron A
ttttagGGGA AAAGGAAAGG GAGATGGAAA GGGCAAGGGG AAGGGGAAAG GAAAAGGCAA
AGGAAAAGGA AAGGGGAAAG GAAAGGGTGG AAAGGGCTCT GCGGTGGTCG TTGAACCGCA
CAAAGTCCAC GCTGGCATTT TTGTTTCCCG AGGAAAGAGC GAAGATTCAT TGGCCACCCT
CAATCTTGTC CCTGGAGTGT CCGTTTACGG TGAGAAGAGA GTGCAAACAG AAgtatgatc  Intron B
tgcacctgca aattctttcg acttctcttc tcttcataac gtcttcaacc gaacttgcag
ACGgtgtgtt gtcttctcct tggcttttag ctccttcaac ttcatagACG GACGCAAACG  Intron C
GAGCCCCAGA GAAGAAAGAA TATCGTGTCT GGAATCCGTA TCGATCAAAG TTGGCCGCTG
CAGTTGTTGG TGGCATTGAC AACTGCAAGG TTGCTCCCGG TGCAAAAGTC C
```

FIG. 4. (*A*) Schematic representation of a genomic PCR product corresponding to a segment of the *E. gracilis* fibrillarin gene and comparison with the cDNA sequence. (*B*) Primary sequence of the PCR product. Positions of the three introns (lower case) are indicated with the conserved GT and AG motifs at the 5′- and 3′-splice junctions underlined. Positions corresponding to the PCR primer binding sites are presented in bold.

stream of microexons. In this context, it is interesting that, downstream of the *Euglena* microexon, the 5′-splice site has the potential to form eight consecutive base pairs with the 5′-terminus of U1 snRNA (Fig. 5); as such, it is the strongest of the 5′-splice sites identified in this study.

CONCLUSIONS

The data presented here, together with previous studies on atypical *Euglena* introns (12–16), establish that two distinct cis-splicing pathways operate in *E. gracilis*. Whether the two cis-splicing mechanisms in *E. gracilis* share any splicing machinery remains unclear; however, the fact that none of the 5′-splice sites in the atypical introns can base pair with the 5′-terminus of U1 snRNA supports the idea that these introns may not be excised by a spliceosomal mechanism (13–16). On the other hand, the spliceosomal cis-splicing and SL RNA trans-splicing pathways are expected to share common snRNAs (28, 29), including the U2 and U4 snRNAs detected in this study (Fig. 1; also see ref. 54). At present, we cannot rule out the possibility that *Euglena* U1 snRNA may also participate in trans-splicing (Fig. 5).

Phylogenetic analyses provide strong evidence that *Euglena* shared a common ancestor with trypanosomatid protozoa and that together these organisms represent an early branch of the eukaryotic lineage (55, 56). It has been argued that the existence of both trans-splicing and spliceosomal cis-splicing in *Euglena* would suggest that the two processes evolved concurrently (23). Therefore, one explanation for the apparent lack of cis-splicing in trypanosomes is that these organisms lost their cis-introns at some point in time (23). Alternatively, it is possible that cis-spliced introns, and perhaps U1 snRNA, do in fact exist in trypanosomes, albeit in low abundance, but have not yet been discovered.

We thank D.F. Spencer for a gift of purified *E. gracilis* nuclear DNA and members of the Gray lab and John M. Logsdon, Jr. for helpful comments. M.W.G., who is a Fellow in the Program in Evolutionary Biology, Canadian Institute for Advanced Research, gratefully acknowledges salary and interaction support from the Canadian Institute for Advanced Research. This work was supported by Grant MT-11212 from the Medical Research Council of Canada to M.W.G.

1. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50,** 349–383.
2. Moore, M. J., Query, C. C. & Sharp, P. A. (1993) in *The RNA World*, eds. Gesteland, R. F. & Atkins, J. F. (Cold Spring Harbor Lab. Press, Plainview, New York), pp. 303–357.
3. Lamond, A. I. (1993) *BioEssays* **15,** 595–603.
4. Lührmann, R., Kastner, B. & Bach, M. (1990) *Biochim. Biophys. Acta.* **1087,** 265–292.
5. Baserga, S. J. & Steitz, J. A. (1993) in *The RNA World*, eds. Gesteland, R. F. & Atkins, J. F. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 359–381.
6. Krämer, A. (1996) *Annu. Rev. Biochem.* **65,** 367–409.
7. Mattaj, I. W., Tollervey, D. & Séraphin, B. (1993) *FASEB J.* **7,** 47–53.
8. Madhani, H. D. & Guthrie, C. (1994) *Annu. Rev. Genet.* **28,** 1–26.
9. Nilsen, T. W. (1994) *Cell* **78,** 1–4.
10. Ares, M., Jr. & Weiser, B. (1995) *Prog. Nucleic Acid Res. Mol. Biol.* **50,** 131–159.
11. Rosbash, M. & Séraphin, B. (1991) *Trends Biochem. Sci.* **16,** 187–190.
12. Muchhal, U. S. & Schwartzbach, S. D. (1992) *Plant Mol. Biol.* **18,** 287–299.
13. Muchhal, U. S. & Schwartzbach, S. D. (1994) *Nucleic Acids Res.* **22,** 5737–5744.
14. Tessier, L. H., Paulus, F., Keller, M., Vial, C. & Imbault, P. (1995) *J. Mol. Biol.* **245,** 22–33.
15. Tessier, L.-H., Chan, R. L., Keller, M., Weil, J.-H. & Imbault, P. (1992) *FEBS Lett.* **304,** 252–255.
16. Henze, K., Badr, A., Wettern, M., Cerff, R. & Martin, W. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 9122–9126.
17. Tessier, L.-H., Keller, M., Chan, R. L., Fournier, R., Weil, J.-H. & Imbault, P. (1991) *EMBO J.* **10,** 2621–2625.
18. Keller, M., Tessier, L. H., Chan, R. L., Weil, J. H. & Imbault, P. (1992) *Nucleic Acids Res.* **20,** 1711–1715.
19. Laird, P. W. (1989) *Trends Genet.* **5,** 204–208.
20. Agabian, N. (1990) *Cell* **61,** 1157–1160.
21. Nilsen, T. W. (1993) *Annu. Rev. Microbiol.* **47,** 413–440.
22. Bonen, L. (1993) *FASEB J.* **7,** 40–46.
23. Nilsen, T. W. (1994) *Science* **264,** 1868–1869.
24. Blumenthal, T. (1995) *Trends Genet.* **11,** 132–136.
25. Davis, R. E., Singh, H., Botka, C., Hardwick, C., el Meanawy, M. A. & Villanueva, J. (1994) *J. Biol. Chem.* **269,** 20026–20030.
26. Davis, R. E., Hardwick, C., Tavernier, P., Hodgson, S. & Singh, H. (1995) *J. Biol. Chem.* **270,** 21813–21819.
27. Davis, R. E. (1997) *Mol. Biochem. Parasitol.* **87,** 29–48.
28. Hannon, G. J., Maroney, P. A. & Nilsen, T. W. (1991) *J. Biol. Chem.* **266,** 22792–22795.
29. Maroney, P. A., Yu. Y.-T., Jankowska, M. & Nilsen, T. W. (1996) *RNA* **2,** 735–745.
30. Dungan, J. M., Watkins, K. P. & Agabian, N. (1996) *EMBO J.* **15,** 4016–4029.
31. Thomas, J., Lea, K., Zucker-Aprison, E. & Blumenthal, T. (1990) *Nucleic Acids Res.* **18,** 2633–2642.
32. Shambaugh, J. D., Hannon, G. E. & Nilsen, T. W. (1994) *Mol. Biochem. Parasitol.* **64,** 349–352.
33. Schnare, M. N. & Gray, M. W. (1990) *J. Mol. Biol.* **215,** 73–83.
34. Greenwood, S. J., Schnare, M. N. & Gray, M. W. (1996) *Curr. Genet.* **30,** 338–346.
35. Greenwood, S. J. & Gray, M. W. (1998) *Biochim. Biophys. Acta* **1443,** 128–138.
36. Frank, R., Müller, D. & Wolff, C. (1981) *Nucleic Acids Res.* **9,** 4967–4979.
37. Schnare, M. N., Heinonen, T. Y. K., Young, P. G. & Gray, M. W. (1985) *Curr. Genet.* **9,** 389–393.
38. Schnare, M. N., Greenwood, S. J. & Gray, M. W. (1995) *FEBS Lett.* **362,** 24–28.
39. Inoue, H., Nojima, H. & Okayama, H. (1990) *Gene* **96,** 23–28.
40. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
41. Dante, M., Michalik, S. S., Mathew, Z., Doray, B., Ho, L. L., Wise, K. J. & Christianson, T. W. (1995) *Anal. Biochem.* **225,** 348–351.
42. Branlant, C. Krol, A., Ebel, J.-P., Lazar, E., Gallinaro, H., Jacob, M., Sri-Widada, J. & Jeanteur, P. (1980) *Nucleic Acids Res.* **8,** 4143–4154.
43. Mount, S. M. & Steitz, J. A. (1981) *Nucleic Acids Res.* **9,** 6351–6368.
44. Krainer, A. R. (1988) *Nucleic Acids Res.* **16,** 9415–9429.
45. Lankat-Buttgereit, B., Gross, H. J. & Krupp, G. (1987) *Nucleic Acids Res.* **15,** 7649.
46. Branlant, C., Krol, A., Ebel, J.-P., Gallinaro, H., Lazar, E. & Jacob, M. (1981) *Nucleic Acids Res.* **9,** 841–858.
47. Guthrie, C. & Patterson, B. (1988) *Annu. Rev. Genet.* **22,** 387–419.
48. Gu, J. Chen, Y. & Reddy, R. (1998) *Nucleic Acids Res.* **26,** 160–162.
49. Zwieb, C. (1997) *Nucleic Acids Res.* **25,** 102–103.
50. Santoni, M. J., Barthels D., Vopper, G., Boned, A., Goridis, C. & Wille, W. (1989) *EMBO J.* **8,** 385–392.
51. Benoist, P., Mas, J. A., Marco, R. & Cervera, M. (1998) *J. Biol. Chem.* **273,** 7538–7546.
52. Nakari, T., Alatalo, E. & Penttilä, M. E. (1993) *Gene* **136,** 313–318.
53. Dominski, Z. & Kole, R. (1992) *Mol. Cell. Biol.* **12,** 2108–2114.
54. DiMaria, P., Palic, B., Debrunner-Vossbrinck, B. A., Lapp, J. & Vossbrinck, C. R. (1996) *Nucleic Acids Res.* **24,** 515–522.
55. Cavalier-Smith, T. (1993) *Microbiol. Rev.* **57,** 953–994.
56. Sogin, M. L. (1991) *Curr. Opin. Genet. Dev.* **1,** 457–463.