ELSEVIER

FEMS Microbiology Letters 204 (2001) 233–237

FEMS MICROBIOLOGY Letters

www.fems-microbiology.org

Two sequence classes of kinetoplastid 5S ribosomal RNA gene revealed among bodonid spliced leader RNA gene arrays

Débora M. Santana [a,1], Julius Lukeš [b], Nancy R. Sturm [a], David A. Campbell [a,*]

[a] *Department of Microbiology, Immunology and Molecular Genetics, University of California at Los Angeles, 10833 Le Conte Avenue, Los Angeles, CA 90095-1747, USA*

[b] *Institute of Parasitology, Czech Academy of Sciences and Department of Parasitology, Faculty of Biology, University of South Bohemia, 370 05 České Budějovice, Czech Republic*

Received 20 July 2001; accepted 18 August 2001

First published online 3 October 2001

Abstract

The spliced leader RNA genes of *Bodo saltans*, *Cryptobia helicis* and *Dimastigella trypaniformis* were analyzed as molecular markers for additional taxa within the suborder Bodonina. The non-transcribed spacer regions were distinctive for each organism, and 5S rRNA genes were present in *Bodo* and *Dimastigella* but not in *C. helicis*. Two sequence classes of 5S rRNA were evident from analysis of the bodonid genes. The two classes of 5S rRNA genes were found in other Kinetoplastids independent of co-localization with the spliced leader RNA gene.

Keywords: Kinetoplastida; Mini-exon; *Bodo*; *Cryptobia*; *Dimastigella*; *Trypanoplasma*

1. Introduction

The protozoan order Kinetoplastida is divided into two suborders, the Trypanosomatina and the Bodonina. The Bodonina consists of two families, the Bodonidae and the Cryptobiidae. The formal distinction between the two lies in the recurrent flagellum that is free from the body in the former and attached to the body in the latter. The bodonids have traditionally been regarded as phagotrophic free-living flagellates; the cryptobiids as saprotrophic commensals or parasites of fish.

The spliced leader (SL) RNA gene (also known as the 'mini-exon' gene) is present in multicopy tandem repeats in the nuclear genomes of kinetoplastid protozoa, nematode worms, and flat worms. It is a useful genetic marker to discriminate among the Kinetoplastida [1] and related organisms [2]. The only Bodonina SL RNA genes to have been studied are from *Bodo caudatus* [3] and *Trypanoplasma* [4]. To further the repertoire of molecular taxonomic markers for the Bodonina, we investigated the SL RNA genes of *Bodo saltans*, *Cryptobia helicis* and *Dimastigella trypaniformis*. The SL RNA gene repeats in *B. saltans* and *Dimastigella* were found to contain a 5S rRNA gene similar to the arrays of *B. caudatus* and *Trypanoplasma*. Comparative analysis has revealed two sequence classes of 5S rRNA in the kinetoplastid protozoa.

2. Materials and methods

2.1. Cell culture

DNA was extracted from *B. saltans* (strain K) and *Dimastigella* (strain Ulm), cultivated as described previously [5,6], and from *C. helicis*, obtained from receptacula seminis of garden snails, *Helix pomatia*, captured in South Bohemia, Czech Republic.

2.2. PCR and molecular cloning

The oligonucleotides used for SL RNA gene-repeat amplification and the thermal cycles have been described elsewhere [7]. The buffer was 100 μM of each dNTP, 1.5 mM $MgCl_2$, 10 mM Tris–HCl, pH 8.3, 25 mM KCl, 1 μg

* Corresponding author. Tel.: +1 (310) 206-5556; Fax: +1 (310) 206-3865.
E-mail address: dc@ucla.edu (D.A. Campbell).

[1] Present address: Instituto de Microbiologia-Prof. Paulo de Góes, CCS, Cidade Universitária, Ilha do Fundão, UFRJ, Rio de Janeiro, RJ, Brazil. Tel.:

0378-1097/01/$20.00
PII: S0378-1097(01)00405-0

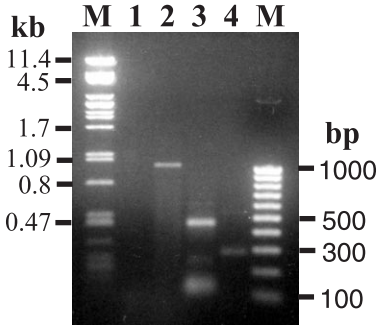

Fig. 1. Amplification of the SL RNA gene from three bodonid species. Agarose gel (1%) electrophoresis of PCR products generated from *Dimastigella* (lane 2), *B. saltans* (lane 3), and *C. helicis* (lane 4) genomic DNA samples. Lane 1 is the no-template control amplification. Molecular size markers are bacteriophage λ DNA digested with *Pst*I (left lane M) and a 100-bp ladder (right lane M; Gibco BRL)

bovine serum albumin and 2.5 U *Taq* polymerase (Sigma). Amplification products were cloned directly using the TOPO™ TA Cloning kit (Invitrogen).

2.3. Sequence analysis

The DNA sequence of the inserts was obtained in both directions using fluorescent-dye terminator methodology (Davis Sequencing). Alignments were made using the PILEUP routine (gap weight = 1.0, gap length weight = 1.0) in the University of Wisconsin Genetic Computer Group package. RNA folding was performed with 'mfold' via the Internet site (http://bioinfo.math.rpi.edu/~mfold/rna/form1.cgi).

3. Results and discussion

3.1. SL RNA gene arrays

Amplification of the SL RNA gene repeat units was performed using universal primers targeted against the 20–39 region of the SL RNA exon and the products were resolved by gel electrophoresis (Fig. 1). Discrete bands were observed of approximately 950 bp for *Dimastigella*, 470 bp for *B. saltans*, and 300 bp for *C. helicis*. The *Dimastigella* DNA sequence (GenBank # AF288761) contained a predicted 112-nt SL RNA and a 5S rRNA gene within the repeat. *B. saltans* (GenBank # AF288756-7) possessed a putative 93–94-nt SL RNA with the presence of a 5S rRNA gene; of two clones sequenced there was an additional nucleotide at position 46 in the intron that falls within a single-stranded region between stem–loop I and stem–loop II. *C. helicis* (GenBank # AF288758-60) had a 94-nt SL RNA and no associated 5S rRNA gene. Of three clones sequenced, all had identical exon sequences. Two clones had identical introns, however the third had three point mutations in the stem–loop II structure: one was in the loop and the other two were compensatory changes in the stem. In the non-transcribed spacer region, one clone had a 13-bp relative deletion and multiple additional differences. As in other apparent deletions in kinetoplastid SL RNA genes [8,9], flanking short direct repeats (GCCCT) were observed in *Cryptobia*.

New exon (SL) sequences were obtained for *B. saltans* and *Dimastigella* (Fig. 2); the *C. helicis* exon was identical to that reported for *Trypanoplasma*. The change of A to T at position 5 of the exon is the first mutation reported in the highly conserved positions 1–6 that include the site of 'cap 4' nucleotide modifications [10]. The intron sequences were different from each other and from those of *B. caudatus* and *Trypanoplasma* (Fig. 3). Although the SL RNA sequences were distinct, the sequence similarity was evident among *B. caudatus*, *C. helicis*, *B. saltans* and *Trypanoplasma*. In all cases, potential base-pairing within the intron to form stem–loop II and stem–loop III was present, as were loose consensus 'Sm'-binding sites between the two stem–loop structures (Fig. 3). Excluding the 5S rRNA sequences and short blocks of identity in repetitive sequences, the non-transcribed spacer regions were distinct.

3.2. Tandem SL RNA gene and 5S rRNA gene arrays

The gene for the 120-nt 5S rRNA is present typically at

```
...C....A--.T..A...CTA.....................  Diplonema papillatum
....T....--....A...ATA.....................  Bodo caudatus
.........--....A...ATA.....................  B. saltans
....T....--....A....CA.....................  Cryptobia helicis
.........--..A.A....TA.....................  Dimastigella
AACTAACGC--TATATAAGTATCAGTTTCTGTACTTTATTG  Leishmania spp.
.........--...TCT..ATA.....................  Phytomonas spp.
.........--...TATTAG.A.............A.....  Trypanosoma brucei
......A.A--.T.TATT..TA.............A.....  T. cyclops
......AAA-T...TAT.-.TA.............A.....  T. simiae
......AA.AA.T.T.G.AG.A.............C....  Leptomonas collosoma
                ACTT.C.G...G...A.TT...T.C.  Euglena gracilis
```

Fig. 2. Alignment of SL exon sequences from the Euglenozoa. Sequence identity is presented relative to the *Leishmania* spp. exon that has been found in over 20 different taxa; bodonids are above the consensus, trypanosomatids are below. Dots represent nucleotide identity; dashes represent gaps introduced to maximize the alignment. GenBank accession numbers are: *Diplonema papillatum* (AY007785); *B. caudatus* (X63467); *Leishmania* sp., (X73121); *Phytomonas* sp. (X87136); *T. brucei* (X00935); *T. cyclops* (AJ250743); *T. simiae* (X99907); *L. collosoma* (K02633) and *Euglena gracilis* (X63153).

```
                              Stem-loop II          Sm      Stem-loop III
    B. caudatus   GTATGAGAAGCTTTCCAGAAA------TGGAA----atttttttgATAAGATTTTCGGATC
     C. helicis   ............A.....CT.------....TTA--......aA...GC..CC.....G..GC
  Trypanoplasma   ....T.........G..TTCATTTC....---CA........a-......C........TT
     B. saltans   ........-.......TTCTA-AAC...----CA.....g-..AG..T.C.....A.
   Dimastigella   ...CATTG.CT..C.GGCTC.TAAAAA..AGCCACA....a.gAACA..A.CAG.TCTTC.GAAAGGAAGAAC
```

Fig. 3. Alignment of intron regions of the bodonid SL RNA genes. The sequences reported here are compared to *B. caudatus* (X63467) and *Trypanoplasma* (L08173). Dashes represent gaps introduced to maximize the alignment. Potential base-pairing to create stem–loop II and stem–loop III are indicated by underlining. The single-stranded consensus core RNP-binding site (Sm) is indicated by lower case letters.

high copy number per cell in a tandemly arrayed gene structure [11] similar to that of the SL RNA gene. The variable association of the 5S rRNA gene with the SL RNA gene has been noted in the Euglenozoa and the nematodes, as has its association with rRNA genes and histone genes [12]. A summary of where this association has been found in kinetoplastid protozoa is presented in Fig. 4.

Several points can be deduced from this comparison. First, the association of the two genes is frequent in organisms (*Euglena*, the Bodonidae, and the Cryptobiidae) that are considered to have branched earlier within this lineage [13]. Second, the orientation of the 5S rRNA gene relative to the SL RNA gene is not conserved, nor does it correlate with phylogenetic relationships among the Kinetoplastids [6,14,15]. Third, there are constraints on the minimal length of a repeat. The smallest repeat to contain a 5S rRNA gene is 431 bp from *B. saltans*. While the 5S rRNA gene promoter is internal, a major constraint on the minimum size of the non-transcribed spacer region is the presence of promoter elements upstream of the SL RNA gene [16].

3.3. Two classes of 5S rRNA genes

A search of the GenBank database identified 33 different 5S rRNA, or 5S rRNA-like, gene sequences from members of the Kinetoplastida. A gestalt view of a multiple alignment of these sequences suggests the presence of two distinct sequence classes (provided for reviewers; available upon request). These two clusters are exemplified by the sequences of *B. saltans* and *Leishmania tarentolae* taken from this alignment (Fig. 5A). The 5S rRNA sequence from *Dimastigella* clusters with *B. saltans*; the *B. caudatus* and *Trypanoplasma* 5S rRNA sequences cluster with *L. tarentolae* (Fig. 5A). The sequences from the two groups have 36 nt in common (32%); an additional 46 nt are shared among most members of each group (41%). A relative deletion of 12 nt in the *B. saltans* group further distinguishes the two classes. Despite the overall sequence differences, the predicted 5S rRNA molecules from *B. saltans* (Fig. 5B) and *L. tarentolae* (Fig. 5C) can adopt similar secondary structures that are consistent with established models.

The presence of two distinct groups of 5S rRNA sequences within the kinetoplastid protozoa was unexpected. Since the suborders Bodonina and Trypanosomatina were represented in both groups, there is no obvious correlation of the two sequence classes with aspects of Euglenozoa phylogeny. Nor is there a correlation with the association of the 5S rRNA genes with the SL RNA genes, since both 5S groups contained examples of associated and non-associated genes (Fig. 4). At this point we have not detected both classes within a single species, however, this can be addressed experimentally. Neither have we addressed the

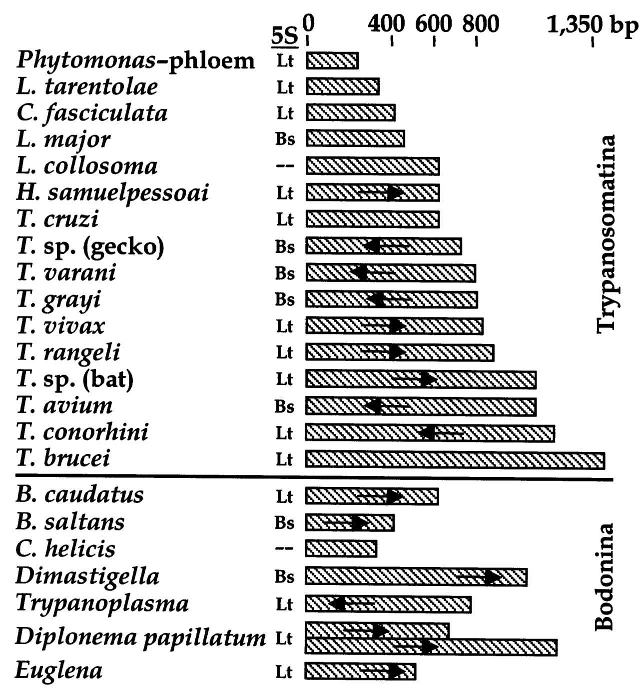

Fig. 4. Association of 5S ribosomal RNA genes and SL RNA genes within the Euglenozoa. Bars represent the length of one repeat unit and are drawn to scale relative to the +1 of the SL. Arrows represent the relative direction and approximate location of 5S rRNA genes; Lt = L. *tarentolae*-like 5S rRNA, Bs = B. *Saltans*-like 5S rRNA. Some repeats that lack 5S rRNA genes are shown to demonstrate the extremes of repeat length. GenBank accession numbers for the respective sequences are: *Phytomonas* (X87136); *L. braziliensis* (X69441); *L. tarentolae* (X73121); *Crithidia fasciculata* (U96170); *L. major* (X69449); *L. collosoma* (K02633); *H. samuelpessoai* (X62331); *T. cruzi* (X62674); *Trypanosoma* sp. (gecko) (AJ250742); *T. varani* (AJ250739); *T. grayi* (AJ250738); *T. vivax* (K00631); *T. rangeli* (X62675); *Trypanosoma* sp. (bat) (AF124146); *T. avium* (AJ250736); *T. conorhini* (AJ272600); *T. brucei* (X00935); *B. caudatus* (X63467); *Trypanoplasma* (L08172); *Diplonema papillatum* (AF329085 and AY007785); and *Euglena gracilis* (X63153).

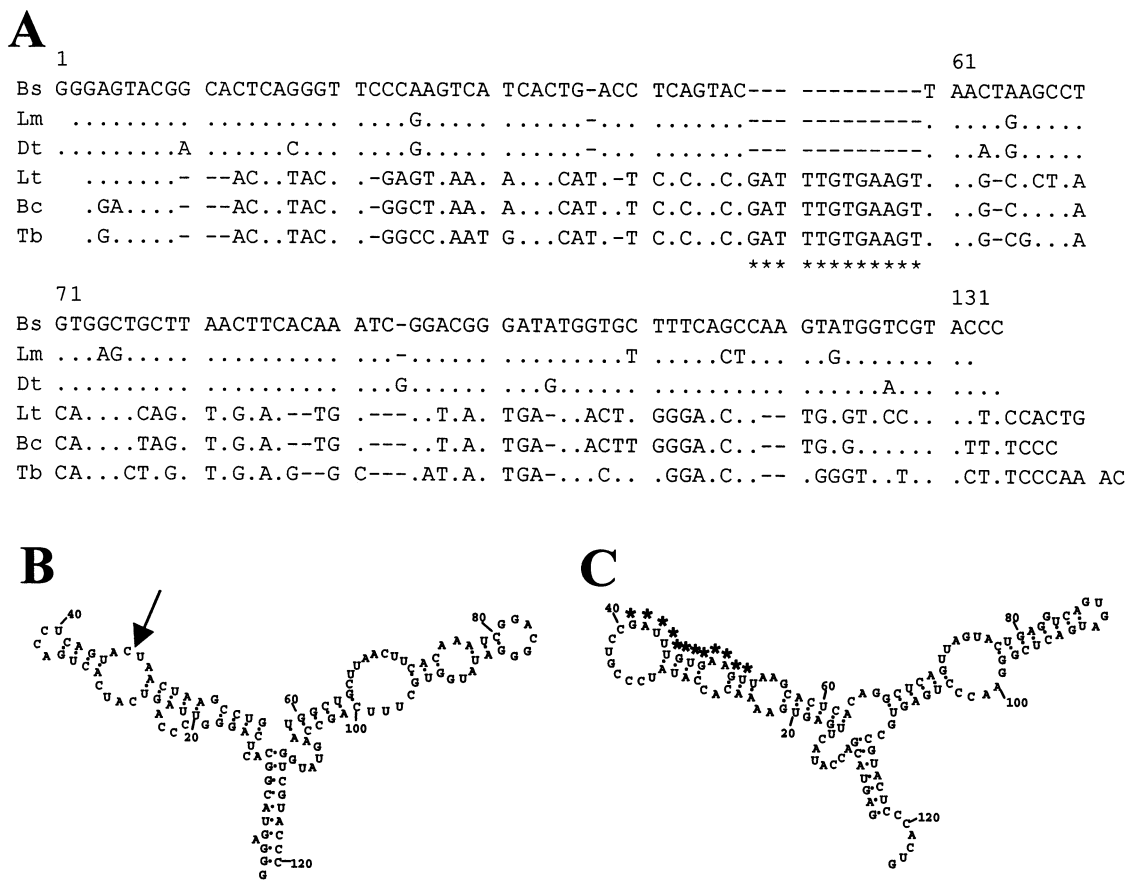

Fig. 5. Identification of two classes of 5S rRNA genes in kinetoplastid protozoa. A: Alignment of sequence from *B. saltans* (Bs), *L. major* (Lm; AL354512), *D. trypaniformis* (Dt), *L. tarentolae* (Lt; AF016249), *B. caudatus* (Bc; X63467) and *T. brucei* (Tb; M14817) numbered relative to the *L. tarentolae* sequence. The dashes represent gaps in the alignment. A 12-nt difference is highlighted by asterisks underneath the sequence. B: Predicted secondary structure of the *B. saltans* 5S rRNA. An arrow marks the position of the 12-nt difference. C: Predicted secondary structure of the *L. tarentolae* 5S rRNA. Asterisks mark the position of the 12-nt difference.

possibility that some of the sequences we have collated represent non-transcribed pseudogenes. The presence of multiple 5S sequence classes in other organisms is well-documented, and their expression has been correlated with particular biological phenomena, particularly developmental changes [17]. Future experiments will address the distribution of the two classes, and if both classes are present within a single species, whether there is differential expression during the protozoan life cycle.

Acknowledgements

This study was supported in part by a donation from Margaret H. Moffat to D.A.C., and by Grant number 204/00/1212 from the Grant Agency of the Czech Republic to J.L. D.M.S. received fellowships from Coordination for the Improvement of Higher Education Personnel (CAPES), Research Strengthening Group (RSG)-UNDP/World Bank/WHO- Special Program for Research and Training in Tropical Diseases (TDR ID number 981147) and 'José Bonifácio' University Foundation of the Federal University of Rio de Janeiro (FUJB-UFRJ), Brazil.

References

[1] Campbell, D.A., Fernandes, O. and Sturm, N.R. (1997) The mini-exon gene is a distinctive nuclear marker for distinguishing protists of the kinetoplastid/euglenoid lineage. Mem. Inst. Oswaldo Cruz 92 1 (Suppl.), 12–13.

[2] Sturm, N.R., Maslov, D.A., Grisard, E.C. and Campbell, D.A. (2001) *Diplonema* spp. possess spliced leader RNA genes similar to the Kinetoplastida. J. Eukaryot. Microbiol. 48, 325–331.

[3] Campbell, D.A. (1992) *Bodo caudatus* medRNA and 5S rRNA genes: tandem arrangement and phylogenetic analyses. Biochem. Biophys. Res. Commun. 182, 1053–1058.

[4] Maslov, D.A., Elgort, M.G., Wong, S., Peckova, H., Lom, J., Simpson, L. and Campbell, D.A. (1993) Organization of mini-exon and 5S rRNA genes in the kinetoplastid *Trypanoplasma borreli*. Mol. Biochem. Parasitol. 61, 127–135.

[5] Blom, D., de Hann, A., Van den Berg, M., Sloof, P., Jirků, M., Lukeš, J. and Benne, R. (1998) RNA editing in the free-living bodonid *Bodo saltans*. Nucleic Acids Res. 26, 1205–1213.

[6] Berchtold, M., Phillipe, H., Breunig, A., Brugerolle, G. and König, H. (1994) The phylogenetic position of *Dimastigella trypanoformis* within the parasitic kinetoplastids. Parasitol. Res. 80, 672–679.

[7] Murthy, V.K., Dibbern, K.M. and Campbell, D.A. (1992) PCR amplification of mini-exon genes differentiates *Trypanosoma cruzi* from *Trypanosoma rangeli*. Mol. Cell. Probes 6, 237–243.

[8] Fernandes, O., Degrave, W. and Campbell, D.A. (1993) The mini-

exon gene: a molecular marker for *Endotrypanum schaudinni*. Parasitology 107, 219–224.
[9] Sturm, N.R., Murthy, V.K., Garside, L. and Campbell, D.A. (1998) The mini exon gene of *Trypanosoma (Nannomonas) simiae*: sequence variation between isolates as a distinguishing molecular marker. Acta Trop. 71, 199–206.
[10] Freistadt, M.S., Cross, G.A.M. and Robertson, H.D. (1988) Discontinuously synthesized mRNA from *Trypanosoma brucei* contains the highly methylated 5′ cap structure, m7GpppA*A*C(2′-O)mU*A. J. Biol. Chem. 263, 15071–15075.
[11] Köck, J. and Cornelissen, A.W.C.A. (1990) The 5S ribosomal RNA genes of *Crithidia fasciculata*. Mol. Biochem. Parasitol. 38, 295–298.
[12] Drouin, G. and de Sá, M.M. (1995) The concerted evolution of 5S ribosomal genes linked to the repeat units of other multigene families. Mol. Biol. Evol. 12, 481–493.
[13] Vickerman, K. (1990) Phylum Zoomastigina class Kinetoplastida. In: Handbook of Protoctista, Vol. ? (Margulis, L., Corliss, J.O., Melkonian, M. and Chapman, D.J., Eds.), pp. 215–238. Jones and Bartlett, Boston, MA.
[14] Maslov, D.A. and Simpson, L. (1995) Evolution of parasitism in kinetoplastid protozoa. Parasitol. Today 11, 30–32.
[15] Wright, A.-D.G., Li, S., Feng, S., Martin, D.S. and Lynn, D.H. (1999) Phylogenetic position of the kinetoplastids, *Cryptobia bullocki*, *Cryptobia catostomi*, and *Cryptobia salmositica* and monophyly of the genus *Trypanosoma* inferred from small subunit ribosomal RNA sequences. Mol. Biochem. Parasitol. 99, 69–76.
[16] Campbell, D.A., Sturm, N.R. and Yu, M.C. (2000) Transcription of the kinetoplastid spliced leader RNA gene. Parasitol. Today 16, 78–82.
[17] Wolffe, A.P. (1994) The role of transcription factors, chromatin structure and DNA replication in 5S RNA gene regulation. J. Cell Sci. 107, 2055–2063.