Characterisation of kinetoplast DNA minicircles from *Herpetomonas samuelpessoai*

Guoliang Fu 1, Bronwen Lambson, Douglas Barker *

*MRC Outstation of NIMR, Molteno Laboratories, Department of Pathology, University of Cambridge, Cambridge CB2 1QP, UK*

Received 3 October 1998; received in revised form 15 December 1998; accepted 4 January 1999

Abstract

In this study, we have sequenced more than 100 clones of minicircle DNA from *Herpetomonas samuelpessoai*. An unusual amplification approach was developed to amplify minicircle DNA by using a pair of complementary primers designed from a universal stretch of minicircle sequence. Sequence analysis shows that the kinetoplast minicircles in *Herpetomonas* with a size of 1.3 kb are organised into two conserved regions and two variable regions which are located 180° apart. The potential gRNA genes are encoded in variable regions of minicircle approximately 360 bp from CSB-3 (conserved sequence block 3). A conserved upstream sequence located 30 nt before the gRNA genes was identified and is related to the gRNA genes in sequence organisation. A potential role(s) of this sequence in gRNA transcription is discussed.

1999 Published by Elsevier Science B.V. All rights reserved.

Keywords: *Herpetomonas samuelpessoai*; Kinetoplast minicircle; Guide RNA gene; RNA editing

1. Introduction

Kinetoplast DNA (mitochondrial DNA) of kinetoplastid protozoa consists of several thousand minicircles (0.5–2.5 kb) and 25–50 maxicircles (20–37 kb) catenated in a network [1]. The maxicircles are the homologues of mitochondrial DNA of other organisms and encode ribosomal RNAs and proteins involved in mitochondrial energy transduction. The function of minicircles is related to the process of RNA editing by encoding guide RNA genes. The minicircle DNA molecules are organised into at least one conserved 100–200-nucleotide region and an equivalent number of variable regions with a different size depending on the species [2]. A 12-mer universal sequence (CSB-3) has been found in the conserved region of all trypanosomatid minicircles and functions as the origin of replication for one strand of DNA [3,4].

The uridylate residue (U) insertion/deletion type of RNA editing is unique to kinetoplastid protozoa. In recent years, RNA editing in a variety of species, *Leishmania tarentolae*, *Trypanosoma brucei*, *Crithidia fasciculata*, *Trypanosoma cruzi*, *Trypanoplasma borreli* and *Herpetomonas*, has been investigated; some species have been studied extensively [5,6]. The guide RNA (gRNA) plays a key role in mediating RNA editing by providing the edited sequence information...
and some gRNA genes have been characterised in a variety of species [7–12]. However, gRNA genes in Herpetomonas have not been identified so far, despite the description of two edited genes (COIII and A6) in this genus [13,14]. The gRNA genes are encoded by both kinetoplast maxicircle and minicircle DNA; the minicircle is the main source of gRNA gene coding. In this work we have sequenced a number of minicircle DNAs from Herpetomonas samuelpessoai. Several potential H. samuelpessoai gRNA genes and conserved upstream sequences have been identified.

2. Materials and methods

2.1. Parasite culture and DNA preparation

H. samuelpessoai was obtained from the World Health Organisation Reference Centre at the London School of Hygiene and Tropical Medicine. The parasites were grown at 26°C as described [15]. The kinetoplast DNA was isolated and purified from a culture according to a similar approach applied to Leishmania [15].

2.2. PCR amplification and cloning

Based on a universal stretch of CSB-3 sequence, a pair of complementary primers, 5′-GGGGTTGGTTAG-3′ and 5′-CAACCACTAA-3′, were designed and synthesised. PCR (referred to as complementary PCR) was performed using normal conditions (buffer, dNTP, Taq units as recommended by Perkin Elmer, the manufacturer), except the primer concentration was twice normal in order to have enough primers annealing to target DNA, despite the competition of self annealing. The cycle condition was: 35 cycles at 94°C for 30 s, 51°C for 150 s, 72°C for 2 min, followed by extension at 72°C for 9 min. PCR products were purified with a Wizard PCR preps DNA Purification System (Promega) and were directly cloned into pGEM-T vector (Promega).

For amplification and sequencing of a full-length minicircle of her3, three primers were designed for PCR and primer working. The sequences of primers and their positions are shown in Fig. 1A.

2.3. DNA sequencing and analysis

More than one hundred white clones were picked and miniprepped with a Wizard minipreps DNA Purification System (Promega). DNA sequencing was carried out using the Prism Dye terminator sequencing Kit (Perkin Elmer) and an ABI 373A automated sequencer. Computer analysis was performed using DNASTAR (DNASTAR, Madison, WI) and University of Wisconsin Genetics Computer Group (UWGCG) software.

2.4. Nucleotide accession numbers

Nucleotide sequences data reported in this paper are available in the EMBL, GenBank and DDJB data bases under the accession numbers: AF064359–AF064399.

3. Results and discussion

A digestion of the kinetoplast DNA with EcoRI shows one dominant minicircle band with a size of 1.3 kb (data not shown). The traditional methods for cloning minicircle DNAs, either by digestion of purified kinetoplast DNA with a restriction enzyme and ligation into vector or by PCR amplification using two primers designed from different sites of conserved region, suffer from severe bias in that they clone only a certain set of minicircle classes due to sequence divergence. In order to clone the different minicircles with similar proportional abundance as found in nature, we applied an unusual PCR approach using a pair of complementary primers. When we started working on this project, there was no information available about Herpetomonas minicircle sequences in the data base. The only sequence segment we know, which is universal to minicircles of all species and should be present in Herpetomonas, is CSB-3, the origin of DNA replication. Based on CSB-3 we designed a pair of complementary primers (see Section 2). By optimising the PCR conditions, we successfully amplified minicircle DNA from a variety of trypanosomatids species. The following sequencing results from randomly picked clones show that a great variety of sequence classes were ob-
Fig. 1. (A) Nucleotide sequence of one complete kDNA minicircle (her3) of Herpetomonas samuelpessoaai. The CSB-3 sequences are in italics and dot underlined. The conserved upstream sequences (promoter?) and gRNA genes are shown in bold and underlined, respectively. The primers used for PCR and sequencing are shown in doublelines, arrows indicate the orientation of primers. (B) Schematic representation of the sequence organisation of kDNA minicircles of H. samuelpessoaai. The size of the minicircle is indicated as well as the location and polarity of the conserved regions containing the CSB-3, the conserved upstream sequences (promoter?), and the gRNA genes.
Fig. 2. Alignments of gRNAs with edited COIII mRNA sequence (A) and A6 sequence (B). The initiation and termination codons are underlined. Uridines added by editing are in lower case. gRNAs predicted from minicircle sequences are shown below the corresponding mRNA sequence. The vertical bars indicate Watson–Crick base pairs; colons indicate G:U base pairs, C–W pairs are indicated by a single dot, and mismatches are indicated with an X.

Fig. 3. Alignments of *H. samuelpessoaai* gRNA-encoding minicircle-variable regions (A) and appropriate minicircle-variable regions with non-gRNA-encoding capability (B). (A) Thirty-five sequences contain conserved upstream sequences and gRNA genes. gRNA genes confirmed by Fig. 2 are underlined; the conserved upstream sequences are shown by asterisks under the alignment. The consensus sequence (cons) is at the top of the alignment. Numbers between sequences indicate omitted nucleotides, 280 in ‘cons’ represents average nucleotide distance. (B) Seven sequences contain neither conserved upstream sequences nor potential gRNA genes. The imperfect short repeats are underlined by dots and double lines.
tained, indicating that the bias of this cloning procedure could be very limited.

The PCR product from *Herpetomonas* is about 650 bp, approximately half the size of the minicircle band observed in the EcoRI digestion agarose gel, indicating that the minicircle may be organised into two conserved regions and two variable regions which are located 180° apart. This was confirmed by sequencing a full-length minicircle. Similar minicircle organisation has also been described in other trypanosomatids [16,17]. The PCR products were purified and cloned into a T-vector. More than 100 clones were randomly picked and sequenced from both ends by using the ABI Prism Dye Terminator dye cycle sequencing kit. After completion of sequencing of these hundred clones, a surprising phenomenon was immediately found that all inserts of clones are in the same direction. Theoretically, PCR products cloned into T-vector should be inserted randomly in both directions. The unidirectional insertion of PCR products may be explained by the possibility that the translated or transcribed products from another direction of minicircle are toxic to bacteria, or the sequences cause plasmid replication problems. Further investigation of this unusual phenomenon will be interesting. This also reminds us that previous study has indicated that some minicircles were very difficult to clone into plasmid vector by restriction enzyme digestion [18]. These two phenomena may share a common reason.

Homologue sequences were clustered into the same classes by the program SeqManII from DNASTAR package. Forty-one sequence classes were obtained. One of these sequences, her3 was used to obtain the other half of the minicircle by PCR. A pair of primers were designed from the variable region of her3 and the PCR product was sequenced by primer walking. The full minicircle sequence is shown in Fig. 1A. The gRNA gene search was carried out by comparison of possible gRNA encoding sequences of variable region from the 42 minicircle halves and known edited mRNA sequences COIII from *H. samuelpessoai* and A6 from *H. muscarum* using the Bestfit program of the GCG package. This search resulted in finding 12 potential gRNA genes that could guide COIII editing (Fig. 2A), and one gRNA gene which could guide A6 editing (Fig. 2B). The good match between COIII and potential gRNA genes and five places of overlap is evidence for the editing process, as this pattern is very similar to those reported in other well-studied species [8,19]. The mismatch occurred with A6 from the species of *H. muscarum* rather than *H. samuelpessoai* (Fig. 2B). This may be due to sequence polymorphism in the A6 gene in the two *Herpetomonas* species.

When we compared the gRNA coding regions of sequenced minicircles, we found an unusual conserved sequence element located 30 nt before the gRNA genes. The conserved sequence, GTTCGGGTATAGGGTACA, is present in 35 minicircle halves. The multiple sequence alignment of 35 sequences is shown in Fig. 3A. An alignment of similar regions of 7 other sequences which lack the conserved element is shown in Fig. 3B. These 7 sequences without the conserved element are unlikely to have gRNA coding capability, as the sequences of these regions are more C rich rather than TA rich, which is common for gRNA genes, and have an abundance of short repeats. According to the above information, the conserved sequence element could play some role(s) which is related to gRNA function. One possible role is that it may regulate gRNA transcription and act as a transcriptional promoter. Another possibility is that this sequence may be transcribed with the coding part of gRNA and functions as a structural element of gRNA. These explanations need to be tested experimentally. In *Leishmania* and *T. cruzi*, there are no conserved sequences flanking gRNA genes. As there is no similarity to the £anking sequences found in *T. brucei* and *C. fasciculata*, the conserved upstream sequence would appear to be unique for *Herpetomonas*.

In this paper, we show that the 1.3-kb kinetoplast minicircle molecules in *H. samuelpessoai* are organised into two conserved regions located 180° apart. The potential gRNA genes are encoded at a defined location of the variable region. Fig. 1B shows a model of the sequence organisation of a minicircle in which two potential gRNA genes and two conserved upstream sequences are illustrated. We believe that this model could represent approximately two-thirds of the *Herpetomonas* minicircles, as our random cloning provided 35 minicircle halves containing potential gRNA genes with conserved upstream sequences and 7 minicircle halves without them. This model needs to be confirmed by sequencing more
complete minicircles. As other types of sequence organisation like the one shown in Fig. 1A in which only one half has potential for gRNA coding exist, it is also possible that a minicircle with both halves without gRNA genes may also exist. Nevertheless, this report provides initial information and a starting point to a better understanding of the minicircle organisation and the RNA editing in *Herpetomonas*.

**Acknowledgments**

We thank all our laboratory colleagues for technical assistance and helpful discussions. All reviewers’ and editors’ comments are highly appreciated. This work was supported by a grant from the Medical Research Council (UK).

**References**


