Molecular and Biochemical Parasitology, 36 (1989) 197–200
Elsevier

MOLBIO 01199

Short Communication

Sequence of a cDNA for the ND1 gene from *Leishmania major*: potential uridine addition in the polyadenosine tail

David A. Campbell[1], Terry W. Spithill[2], Nicholas Samaras[2], Agda M. Simpson[3] and Larry Simpson[3]

[1]*Department of Microbiology and Immunology, University of California, Los Angeles, CA, U.S.A.;* [2]*The Walter and Eliza Hall Institute of Medical Research, Melbourne, Victoria, Australia; and* [3]*Department of Biology and Molecular Biology Institute, University of California, Los Angeles, CA, U.S.A.*

(Received 10 March 1989; accepted 30 April 1989)

Key words: *Leishmania major*; NADH dehydrogenase 1; Poly-A tail, Uridine addition

Members of the family Kinetoplastida represent a morphologically similar group of protozoa, which share a variety of unusual biochemical pathways, e.g., discontinuous transcription, *trans*-splicing and RNA editing. While studying genetic variability in *Leishmania*, we have, by serendipity, identified a cDNA clone which has more relevance to the question of RNA editing than genetic variability.

RNA editing is a phenomenon first described by Benne [1] in which four non-template-encoded uridine residues were added to the transcript of the *Trypanosoma brucei* cytochrome oxidase II gene, thus overcoming a reading frame shift present in the mitochondrial genome. Subsequently, RNA editing has been shown to create methionine initiation codons [2], and to occur by the deletion of template-encoded uridines [2] in a number of kinetoplastid mitochondrial genes. In an extreme example, over 50% of the *Trypanosoma brucei* cytochrome *c* oxidase III transcript is created by U addition or deletion [3].

A cDNA library (constructed in the plasmid vector pBR322) was synthesized from *Leishmania major* strain V121 (MHOM/IL/67/JERICHO II) promastigote poly(A)$^+$ RNA [4] and screened by hybridization to [^{32}P]cDNA probes synthesized from promastigotes of *L. major* V121 or *Leishmania tropica* L32 (MHOM/IQ/65/A.SINAI I) (T.W.S., unpublished result). One clone (P101/5) selected by differential hybridization with *L. major* but not *L. tropica*, was studied further. However, subsequent analysis with a *L. tropica* ND1 probe (unpublished result) demonstrated the presence of a ND1 gene and transcript in *L. tropica*, suggesting that the selection of clone P101/5 was artefactual.

Clone P101/5, which contains an incomplete cDNA sequence of 1030 bp (Fig. 1), hybridizes to a 1.25-kb poly(A)$^+$ RNA of *L. major* on Northern blots. The DNA sequence predicts a single open reading frame of 301 amino acids, which is extremely hydrophobic. The DNA sequence of P101/5 is 82% similar, and the predicted amino acid sequence is 91% similar, with the ND1 gene of *Leishmania tarentolae* encoding subunit 1 of NADH dehydrogenase 1. Comparison of the *L. major* ND1 sequence with ND1 sequences from other organisms, by the SEQDP program [5], reveals statistically significant align-

Correspondence address: David A. Campbell, Dept. of Microbiology and Immunology, University of California, Los Angeles, CA 90024, U.S.A.

Note: Nucleotide sequence data reported in this paper have been submitted to the GenBank™ Data Bank with the accession number X13751

Abbreviations: ND1, NADH dehydrogenase subunit 1; 3 bp, base pair.

```
 D  I  L  V  V  L  V  L  T  G  F  V  S  L  C  E  R  R  I  L
GACATTTTAGTTGTTCTTGTCTTGACCGGCTTTGTAAGTCTTTGTGAAAGAAGAATTTTA  60
 A  L  V  Q  I  R  I  G  P  A  L  C  F  F  G  I  L  T  P  I
GCTTTAGTTCAAATACGAATAGGTCCGGCCCTTTGTTTTTTTGGAATTTTAACACCAATA 120
 T  D  G  I  K  L  F  I  K  F  I  I  F  V  I  S  F  E  I  I
ACTGACGGAATAAAACTTTTTATTAAATTTATCATTTTTGTAATTAGTTTTGAAATCATT 180
 Y  L  I  G  A  I  L  I  T  T  C  C  I  F  I  G  W  F  Y  F
TACTTAATAGGTGCAATCTTAATTACTACATGCTGCATTTTTATTGGTTGGTTTTATTTT 240
 P  I  G  F  I  L  L  L  D  T  G  F  T  I  T  I  M  M  C  V
CCTATAGGTTTTATACTGTTATTAGACACAGGATTCACTATTACAATAATGATGTGCGTA 300
 H  V  F  S  N  M  F  S  T  F  F  V  G  C  F  L  F  S  S  C
CATGTTTTCAGTAATATGTTTAGCACATTTTTTGTAGGTTGCTTTTTATTCTCTAGTTGC 360
 F  V  Y  L  S  A  M  R  T  M  F  F  S  I  I  S  E  S  G  L
TTTGTATATCTTTCTGCAATGCGCACTATGTTTTTTAGTATTATTTCTGAAAGTGGCTTA 420
 F  L  L  Y  T  T  V  Y  S  L  D  Y  F  S  F  F  C  I  K  D
TTTTTATTATACACAACTGTATATTCGTTAGACTACTTTAGCTTTTTTTGCATTAAAGAC 480
 I  C  V  G  Q  I  Y  I  T  N  F  Y  I  A  G  I  L  F  V  C
ATTTGTGTTGGGCAGATTTATATTACAAACTTTTATATTGCCGGCATCTTATTTGTATGT 540
 I  F  W  I  S  M  L  L  D  G  L  K  L  P  F  D  Y  M  E  C
ATATTTTGAATAAGTATGCTTCTAGACGGATTAAAACTACCATTCGACTATATGGAATGT 600
 E  S  E  L  V  A  G  L  I  T  E  L  S  G  F  F  F  I  L  Y
GAAAGTGAATTAGTCGCTGGGTTGATTACAGAACTATCAGGATTCTTTTTTATATTATAT 660
 S  V  L  E  I  N  H  V  L  L  T  T  L  L  L  A  S  L  C  F
TCCGTTCTTGAAATAAATCATGTGTTACTGACAACATTATTATTAGCCAGTCTCTGTTTT 720
 G  G  L  F  I  C  F  K  A  I  I  I  L  I  I  G  F  F  Y  P
GGGGGCTTATTTATATGTTTTAAAGCTATAATTATCTTAATTATTGGGTTTTTTTATCCA 780
 R  V  I  G  Y  R  L  K  I  T  T  A  Q  A  F  I  L  I  F  L
AGGGTAATTGGTTATAGACTAAAAATTACAACAGCGCAAGCATTTATACTTATTTTTTTA 840
 F  Y  M  C  V  L  T  F  I  W  L  F  T  T  K  I  I  A  I  L
TTCTACATGTGCGTTCTTACTTTTATTTGACTGTTCACTACAAAAATAATTGCTATATTA 900
 F  *
TTTTAATATCGAAAACTGAATAATTTTTTTTTTTAAAAAAAAAAAAAAAAAAAAAAATTT 960
TTAAAATTAAATATTAAAATTAAAAAAAAAAAAAAAAAAAAAATATAAAAAAAATTAATATA  1020
ATAAAAAAAA  1030
```

Fig. 1. The DNA sequence of the P101/5 insert (GenBank Accession No. X13751) and translated amino acid sequence. The DNA sequence was determined completely on both strands by the chemical method of Maxam und Gilbert [7]. The first amino acid shown corresponds to residue 13 of ND1 from *Leishmania tarentolae* [6].

ments of 68.2, 52.2, 18.5, 11.4, 9.8 and 9.3 SD units for *L. tarentolae*, *T. brucei*, *Aspergillus nidulans*, human, *Neurospora crassa*, and *Drosophila*, respectively.

The cDNA sequence shows a putative poly(A) tail extending from position 918 to 1030, which is interrupted by 13 insertions of a total of 31 T (=U) residues. In the absence of the *L. major* maxicircle genomic sequence, we cannot be certain where the poly(A) tail actually begins, although two observations indicate position 918. First, the P101/5 sequence diverges from the *L. tarentolae* genomic sequence [6] at position 917. This position may however represent a divergence between the *L. major* and *L. tarentolae* sequences. Secondly, the cDNA sequence downstream of position 918 contains 100% A/T residues, whereas the intergenic region between ND1 and the next downstream gene, MURF1, is G+C-rich in both of the related species *L. tarentolae* and *T. brucei* [6]. We therefore interpret the 3′ A+T-rich sequences of P101/5 as representing the poly(A) tail of the ND1 transcript, which has been edited by U addition.

Acknowledgements

This investigation received financial support from the Australian National Health and Medical Research Council, NIH grant AI 19347 (T.S.), the Rockefeller Foundation Great Neglected Diseases Programme, the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases and the John D. and Catherine T. MacArthur Foundation (T.W.S.), NIH grant AI 09102 (L.S.) and Basic Sciences Research Grant RR5354 (D.A.C.).

References

1 Benne, R., Van Den Burg, J., Brakenhoff, J.P.J., Sloof, P., Van Boom, J.H.V. and Tromp, M.C. (1986) Major transcript of the frameshifted coxII gene from trypanosome mitochondria contain four nucleotides that are not encoded in the DNA. Cell 46, 819–826.

2 Shaw, J.M., Feagin, J.E., Stuart, K. and Simpson, L. (1988) Editing of kinetoplastid mitochondrial mRNAs by uridine addition and deletion generates conserved amino acid sequences and AUG codons. Cell 53, 401–411.

3 Feagin, J.E., Abraham, J.M. and Stuart, K. (1988) Extensive editing of the cytochrome *c* oxidase III transcript in *Trypanosoma brucei*. Cell 53, 413–422.

4 Kidane, G.Z., Samaras, N. and Spithill, T.W. (1989) Cloning of developmentally regulated genes from *Leishmania major* and expression following heat induction. J. Biol. Chem. 264, 4244–4250.

5 Kanehisa, M.I. (1982) Los Alamos sequence analysis package for nucleic acids and proteins. Nucleic Acids Res. 10, 183–196.

6 Simpson, L., Neckelman, N., de la Cruz, V.F., Simpson, A.M., Feagin, J.E., Jasmer, P. and Stuart, K. (1987) Comparison of the maxicircle (mitochondrial) genomes of *Leishmania tarentolae* and *Trypanosoma brucei* at the level of nucleotide sequence. J. Biol. Chem. 262, 6182–6196.

7 Maxam, A.M. and Gilbert, W. (1977) A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA 74, 560–564.