

Identification of the 10Sa RNA structural gene of *Mycobacterium tuberculosis*

Jaya Sivaswami Tyagi and Asha Kiran Kinger

Department of Biotechnology, All India Institute of Medical Sciences, New Delhi, 110029, India

Submitted November 22, 1991

EMBL accession no. X60301

In the course of screening a gene bank of *Mycobacterium tuberculosis* (1) with cDNA to RNA from *M. tuberculosis*, we have isolated a clone, pAK51, which contains significant sequence homology with *Alcaligenes eutrophus* 10Sa RNA (57% identity) and *Escherichia coli* 10Sa RNA (54% identity). The 10Sa RNA has been described as a stable RNA (2, 3). The conservation of primary sequences does not end at the 3' end of the mature 10Sa RNA but extends further downstream. The 10Sa structural RNA gene is located within a 1 kb *Pst*I genomic fragment of DNA from *M. tuberculosis* H37Rv and *M. tuberculosis* H37Ra (Figure 1A). When total RNA isolated from *M. tuberculosis* H37Rv and *M. tuberculosis* H37Ra was probed with the putative 10Sa RNA-coding region from pAK51, hybridization was obtained with RNA approximately 390–400 bases in length (Figure 1B). Interestingly, the sequence located in the 3' terminal of 10Sa RNA showed strong homology (4) with the pseudouridine arm of tRNA genes from several organisms (Figure 2).

ACKNOWLEDGEMENTS

This work was funded by grants from the Council of Scientific and Industrial Research. DNA sequencing was carried out by JST during her visit to Dr. S.S.Tevethia's laboratory at the Milton S. Hershey Medical Center, Hershey, PA 17033, USA under the auspices of the Department of Biotechnology (DBT, Govt. of India) Visiting Associateship programme for which Dr.Tevethia and the DBT are sincerely acknowledged. A Research Associateship to AKK funded by the DBT is acknowledged. We thank R. Dhanasekaran for help with computer analysis.

REFERENCES

- Bhargava,S., Tyagi,A.K. and Tyagi,J.S. (1990) *J. Bacteriol.* **172**, 2930–2934.
- Chauhan,A.K. and Apirion,D. (1989) *Mol. Microbiol.* **3**, 1481–1485.
- Brown,J.W., Hunt,D.A. and Pace,N.R. (1990) *Nucl. Acids Res.* **18**, 2820.
- Pearson,W.R. and Lipman,D.J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.
- Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA* **84**, 1679–1683.

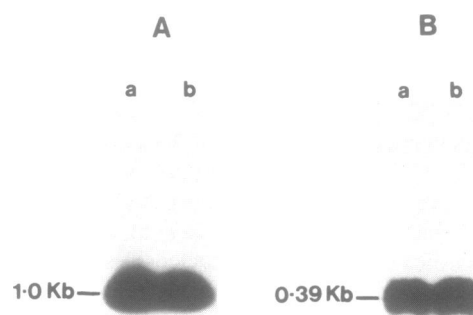


Figure 1. Southern and Northern blot hybridization of pAK51 with mycobacterial DNAs and RNAs. (A). *Pst*I digests of genomic DNA isolated from *M. tuberculosis* H37Rv (lane a) and *M. tuberculosis* H37Ra (lane b) was hybridized with ³²-P labelled 1 kb *Pst*I fragment of pAK51 (6×10^6 cpm). Lambda DNA- *Hind*III markers were run alongside. (B). Total RNA from *M. tuberculosis* H37Rv (lane a) and *M. tuberculosis* H37Ra (lane b) was hybridized (1×10^7 cpm) as mentioned above. RNA molecular weight markers (range 7.84 to 0.3 kb, Boehringer-Mannheim, Germany) were run alongside.

Gene	Alignment
<i>M. tuberculosis</i> 10Sa RNA ^a	GTAGGACCCGGTTCGATTCCCGGCAGCTCCACCG
<i>E. coli</i> 10Sa RNA ^b	T-C--GA-GC-----A-C-----C-----A
<i>A. eutrophus</i> 10Sa RNA ^c	TAC---G-----C-G-----A
Wheat germ gly-tRNA ^d	TACA-----T-G-G---A
<i>B. mori</i> gly-tRNA ^e	GGGC-G-----C-A-G---A
<i>B. mori</i> ala-tRNA ^f	AG--TA-----A-----A-----GC-----ATC
MoMuLV gly-tRNA ^g	-GGA-G-----CAA-G---A
Human gly-tRNA ^h	-GGA-G-----CAA-G---A
Human gly-tRNA ⁱ	T-GC-----GC-GCG---A
<i>H. volcanii</i> ala-tRNA ^j	AG--C-----GAG-----A
Lupin his-tRNA ^k	T-GAA--T-----CA---A

Figure 2. Homology of the 3' terminal region of *M. tuberculosis* 10Sa structural RNA with RNAs from several organisms. Both strands of insert DNA from clone pA.K51 were sequenced by the dideoxy procedure (5). The nucleotide sequence of the 10Sa RNA gene is deposited in the EMBL database under accession number X60301. The first base, G, in the figure represents position 418 in the 10Sa RNA gene sequence. Dashes indicate identical bases while substitutions are indicated accordingly. The conserved motif of the pseudouridine loop of tRNA is overlined. The putative 3' end of the mature 10Sa RNA is indicated by an asterisk. superscript a represents EMBL accession number X 60301. Superscripts b to k represent Genbank annotations ECOSSRA, AFA10SAR, WHTRG1, BMOTRG1, BMOTRA2X, MLMTRGA, HUMTRGGCC, HUMTRGCC, HALTRA1 and LUPTRH respectively.