Complete nucleotide sequence of a gene encoding the 70 kd heat shock protein of *Mycobacterium paratuberculosis*

Karen Stevenson, Neil F.Inglis, Barbara Rae, William Donachie and J.Michael Sharp Moredun Research Institute, 408 Gilmerton Road, Edinburgh EH17 7JH, UK

Submitted July 10, 1991

EMBL accession no. X59437

Nine λ gt11 recombinant clones, containing overlapping fragments of a gene encoding a 70 kilodalton protein antigen, were isolated from a Mycobacterium paratuberculosis genomic library using serum from a rabbit that had been inoculated with a lysate of M. paratuberculosis. Two of the recombinant clones, λ R5 and λ R6, were lysogenised in Escherichia coli strain Y1089 and produced a free protein of 70 kd after induction. The 70 kd protein reacted with monoclonal CosII raised against the 71 kd heat shock protein of Mycobacterium tuberculosis. The DNA inserts from the λ gt11 recombinant clones hybridised to clone Y3272, a \(\lambda \) gt11 recombinant clone containing the 3' end of the gene encoding the M. tuberculosis 71 kd heat shock protein. The nucleotide sequence of the DNA insert in clone λ R5 was determined on both strands by the dideoxy-chain termination method (1) using M13-40 primer and a series of 17 mer oligonucleotides.

The gene encoding the M. paratuberculosis 70 kd antigen has an ORF of 1869 bp and codes for a protein of 623 amino acids. The predicted ATG initiation codon is preceded by a Shine-Dalgarno sequence from nucleotide 181 to 186. The amino acid sequence translated from the nucleotide sequence shows 58% homology with E. coli DnaK, 50% homology with the 70 kd heat shock protein of *Drosophila melanogaster* and 49% homology with the human 70 kd heat shock protein over the entire coding region. The 70 kd antigen of M. paratuberculosis shows a higher degree of homology with the corresponding heat shock proteins of Mycobacterium leprae and M.tuberculosis. There is 90% homology between amino acids 407 and 540 of the M. paratuberculosis 70 kd protein and amino acids 131 to 264 of the M.leprae 70 kd heat shock protein (2) and 3 to 134 of the M. tuberculosis 71 kd heat shock protein (3). The homology declines substantially towards the carboxyl termini of the proteins as can be seen in the figure below.

ACKNOWLEDGEMENTS

The authors thank Dr. D. Young for providing monoclonal CosII, and Dr. R.A. Young for providing clone Y3272. We acknowledge Oswel DNA Service for synthesising the oligonucleotides for sequencing and SERC Daresbury Laboratory for the use of the SEQNET molecular biology computing facility.

REFERENCES

- Sanger, F., Nicklen, S. and Coulsen, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
- Garsia, R.J., Hellqvist, L., Booth, R.J., Radford, A.J., Britton, W.J., Astbury, L., Trent, R.J. and Basten, A. (1989) Infect. Immun. 57, 204-212.
- Young, D., Lathigra, R., Hendrix, R., Sweetser, D. and Young, R.A. (1988) Proc. Natl. Acad. Sci. USA 85, 4267-4270.
- 4. Higgins, D.G. and Sharp, P.M. (1989) CABIOS 5, 151-153.

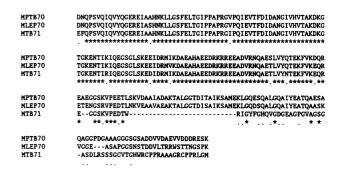


Figure 1. ALIEN alignment of the predicted protein sequences of the *M.paratuberculosis* 70 kd heatshock protein (MPTB70; amino acids 405–623), *M.leprae* 70 kd heat shock protein (MLEP70; amino acids 129–344), and *M.tuberculosis* 71 kd heat shock protein (MTB71; amino acids 1–189) showing identical (*) and conserved (.) residues. ALIEN is a multiple alignment programme by A.J. Bleasby based on the method of Higgins and Sharp (4).