

A Protein Antigen of *Mycobacterium leprae* Is Related to a Family of Small Heat Shock Proteins

AUDUN H. NERLAND,^{1†*} ABU SALIM MUSTAFA,^{1‡} DOUGLAS SWEETSER,² TORE GODAL,^{1§}
AND RICHARD A. YOUNG²

Laboratory for Immunology, Institute for Cancer Research, The Norwegian Radium Hospital, Montebello, N-0310, Oslo 3, Norway,¹ and Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139²

Received 29 March 1988/Accepted 22 August 1988

The gene encoding an immunologically important 18-kilodalton protein antigen of *Mycobacterium leprae* has been sequenced, and the amino acid sequence of the antigen has been deduced. The 18-kilodalton antigen is strikingly similar in size and sequence to a family of eucaryotic heat shock proteins.

Six protein antigens have thus far been implicated in the immune response to *Mycobacterium leprae* (2, 3). These proteins have estimated molecular masses of 70, 65, 36, 28, 18, and 12 kilodaltons (kDa). Most of these proteins have

clones were isolated from healthy vaccinated volunteers (9). The identity and cellular function of the 18-kDa protein is not known.

To obtain clues to the nature of the 18-kDa antigen, the

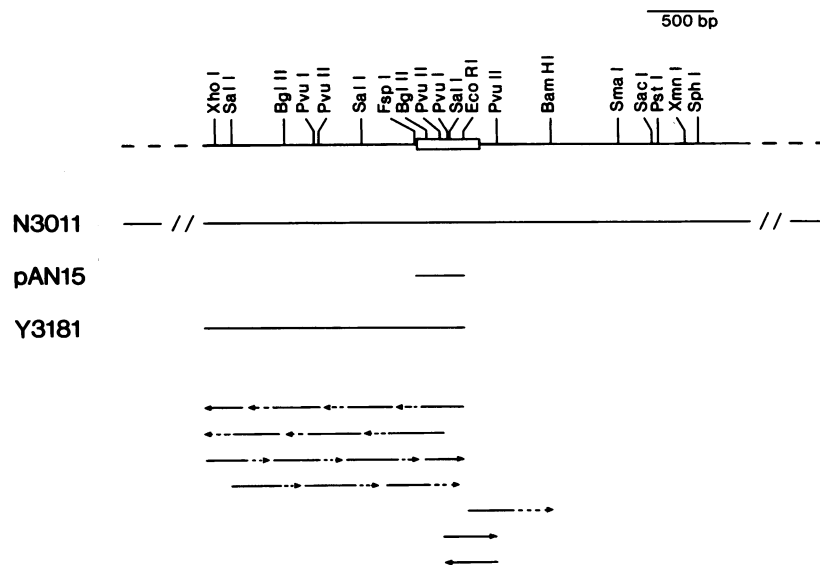


FIG. 1. Restriction maps of DNA encoding the *M. leprae* 18-kDa antigen and the DNA sequence strategy. The restriction map of *M. leprae* DNA containing the coding sequence for the 18-kDa antigen (boxed) is shown. The extents of insert DNA at clones N3011, Y3181, and pAN15, as well as the sequenced fragments, are indicated below the restriction map. bp, Base pairs.

been shown to stimulate human T cells (6, 8a, 9, 10, 12) and are therefore likely to be involved in cell-mediated immunity, an important determinant in protection against leprosy. The 18-kDa protein was the first of these antigens demonstrated to stimulate *M. leprae*-specific T-cell clones; these

gene that encodes it was isolated and sequenced. A portion of the gene encoding the 18-kDa protein was isolated previously by probing a λ gt11 recombinant DNA library with monoclonal antibodies directed against the *M. leprae* protein (17). This recombinant clone, Y3181, expressed a portion of the 18-kDa protein as a fusion protein with β -galactosidase. To obtain the sequence of the entire region encoding the 18-kDa protein, we sequenced the insert DNA from Y3181 as well as a portion of the insert DNA from a λ EMBL clone containing the entire gene encoding the 18-kDa protein (Fig. 1). The λ EMBL clone N3011 was isolated from a λ EMBL library of *M. leprae* DNA by probing with a plasmid containing a portion of the insert DNA of clone Y3181 (pAN15),

* Corresponding author.

† Present address: Center for Biotechnology, University of Bergen, P.b. 3152, Aarstad, 5001 Bergen, Norway.

‡ Present address: Whitehead Institute for Biomedical Research, Cambridge, MA 02142.

§ Present address: World Health Organization, 1211 Geneva 27, Switzerland.

```

      10      20      30      40      50
AGC AAT CTC CTC AGC TGT TCA GAC AGA AAA CTT GTC TAT CAC AAC TTG CAT CAA TAT
      *
      60      70      80      90      100     110
ATC GAC CAG TGC TAT ATC AAA TCT ATG TAG TCA GGA ACA GCT ATA TAG TTA TAG TTT
      *
      120     130     140     150     160     170
GTC ACA ACA GAT TGG AGT GCG AGG TGA CCA CAC ATG CTG ATG CGT ACT GAC COG TTC
      *
      180     190     200     210     220
CGT GAA CTG GAC CGC TTC GCC GAG CAA GTC TTA GGT AGC TCT GCC CGC CCA GCA GTA
Arg Glu Leu Asp Arg Phe Ala Glu Gln Val Leu Gly Thr Ser Ala Arg Pro Ala Val
      *
      230     240     250     260     270     280
ATG CCC ATG GAC GCT TGG CGT GAG GGC GAA GAA TTC GTC GTC GAG TTC GAC CTT CCT
Met Pro Met Asp Ala Trp Arg Glu Gly Phe Val Val Glu Phe Asp Leu Pro
      *
      290     300     310     320     330     340
GGC ATC AAA GCC GAT TCA CTG GAC ATT GAC ATC GAA CGC AAC GTA GTC ACC GTG CCG
Gly Ile Lys Ala Asp Ser Leu Asp Ile Asp Ile Glu Arg Val Val Thr Val Arg
      *
      350     360     370     380     390
GCC GAG CGC CCA GGC GTC GAC CCC GAT CGG GAA ATG CTT GCT GCC GAG CGG CCA CGC
Ala Glu Arg Pro Gly Val Asp Pro Asp Arg Glu Met Leu Ala Ala Glu Arg Pro Arg
      *
      400     410     420     430     440     450
GGT GTG TTC AAT CGT CAG CTG GTT CTC GGC GAA AAC CTC GAC ACC GAA CGG ATC TTG
Gly Val Phe Asn Arg Gln Leu Val Leu Gly Glu Asn Leu Asp Thr Glu Arg Ile Leu
      *
      460     470     480     490     500     510
GCT TCC TAC CAA GAA GGT GTC CTG AAG TTG TCG ATA CCA GTA GCC GAA AGG GCT AAA
Ala Ser Tyr Gln Glu Gly Val Leu Lys Leu Ser Ile Pro Val Ala Glu Arg Ala Lys
      *
      520     530     540     550     560     570
CCG CGC AAG ATC TCC GTT GAT CGT GGC AAC AAC GGA CAC CAG ACC ATA AAC AAA ACC
Pro Arg Lys Ile Ser Val Asp Arg Gly Asn Asn Gly His Gln Thr Ile Asn Lys Thr
      *
      580     590     600     610     620
GCA CAC GAA ATC ATA GAT GCC TAA TCG ACT GTT GTT TGC GCA ACA AGC TAA
Ala His Glu Ile Ile Asp Ala ---
    
```

FIG. 2. Sequence of the gene encoding the *M. leprae* 18-kDa antigen. Sequences were determined in both directions by the chain termination method (13). The deduced amino acid sequence is shown below the DNA sequence.

using standard procedures (7). Clone N3011 contained 17 kb of inserted DNA that included all of the DNA contained within the Y3181 insert DNA (Fig. 1).

Because Y3181 produces a fusion protein containing portions of the 18-kDa protein, the location and orientation of the gene encoding this protein could be estimated in the

overlapping restriction maps of clones Y3181 and N3011 (Fig. 1). DNA fragments isolated from clones Y3181 and N3011 were inserted into M13 vectors (8) and subjected to sequence analysis (13). The entire 1.7-kb *EcoRI* insert DNA of Y3181, a 0.4-kb *PvuII* fragment of N3011 that overlaps the right end of the Y3181 insert DNA, and the 0.6-kb *EcoRI-BamHI* fragment of N3011 that abuts the right end of the Y3181 insert DNA were sequenced (Fig. 1).

Figure 2 shows the DNA sequence obtained for the gene encoding the 18-kDa protein and the amino acid sequence predicted for the 18-kDa protein. The sequence predicts a protein of 148 amino acids with a molecular mass of 16,607 daltons, in good agreement with the molecular weight estimated on sodium dodecyl sulfate-polyacrylamide gels. For the sake of consistency, we will continue to refer to this protein by its apparent molecular mass of 18 kDa.

The amino acid sequence deduced for this antigen was compared with that of proteins in the National Biomedical Research Foundation protein sequence data base. The *M. leprae* antigen exhibited striking sequence similarity to a family of 17-kDa soybean heat shock proteins (11) (Fig. 3). A comparison of the *M. leprae* protein sequence with that of the soybean 17.5E protein revealed 31% amino acid sequence identity in 127 overlapping amino acids. A substantial fraction of the nonidentical amino acids were conserved amino acid replacements. Thus, the 18-kDa *M. leprae* protein antigen is significantly similar in sequence and size to the 17-kDa soybean heat shock proteins, which are representative of a class of small plant stress proteins (11). Moreover, these small plant heat shock proteins exhibit sequence similarities with a class of stress proteins present in animals, known as the small heat shock proteins (5, 11). We postulate that the 18-kDa *M. leprae* protein antigen is a stress protein.

It would be useful to determine directly whether specific stresses would induce increased synthesis of the *M. leprae* 18-kDa protein. However, *M. leprae* has not yet been successfully cultivated in vitro, which precludes direct analysis of the stress response in this organism.

Stress proteins have been identified as targets of the immune response for a variety of pathogens, including

```

                                                                 60
M. leprae                MLMRTDPFRELDLRRASVSTARSAPVAMPMDAWREG - EEFVV
Soybean                  MSLIPGFFGRRSNVDFPFLDMWDPFKDFHVPTSSVSAENSAFVSTRVDWKETPEAHVF
                        -  - - - * * * * * - - - - * * * * * - * * * * * - * * *
                                                                 120
M. leprae                EFDLPGIKADSLDIDIER - NVVTVRAERPGVDPDR - - EMLAAERPRGVFNRLVLGENLD
Soybean                  KADIPGLKKEEVKVEIEDDRVLQISGERNVEKEDKNDTWHRVERSSGKFTRRFRLENK
                        * - * - * - * * * * * - - * * * * * - * - * * * * * - - * * * * *
M. leprae                TERILASYQEGVLKLSIPVAERAKPRKISVDRGNNGHQTINKTAHEIIDA
Soybean                  VNEVKASMENGVLTVTPKEEVKPPDVLAIIEISG
                        * * * * * - - * * * * * * * * * * * * * * * * * * * * * * * * *
    
```

FIG. 3. Comparison of the amino acid sequences of the *M. leprae* 18-kDa antigen and the soybean 17-kDa heat shock protein 17.5E (11). The sequence of the 18-kDa *M. leprae* antigen (top line) is aligned with that of the 17-kDa soybean heat shock protein (bottom line). Symbols: -, identical residues; *, conserved amino acid replacements.

mycobacteria (14, 16), *Coxiella* spp. (15), plasmodia (1), schistosomes (4), and filaria (13a). This fact suggests that the stress response may be a natural response of the infectious pathogen to the hostile environment of the host. The observation that stress proteins are common immune targets of pathogens may reflect the fact that the stress response and the proteins induced by stress are highly conserved through procaryotes and eucaryotes.

This research was supported by grants from the National Institutes of Health (Public Health Service AI23545), the World Health Organization Programme for Vaccine Development, and the World Health Organization/World Bank/United Nations Development Program Special Program for Research and Training in Tropical Diseases. R. A. Y. is a Burroughs Wellcome scholar in molecular parasitology.

LITERATURE CITED

1. Bianco, A. E., J. M. Favaloro, T. R. Burkot, J. G. Culvenor, P. E. Crewther, G. V. Brown, R. F. Anders, R. L. Coppel, and J. Kemp. 1986. A repetitive antigen of *Plasmodium falciparum* that is homologous to heat shock protein 70 of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA 83:8713-8717.
2. Britton, W. J., R. J. Garsia, L. Hellqvist, J. D. Watson, and A. Basten. 1986. The characterization and immunoreactivity of a 70 Kd protein common to *Mycobacterium leprae* and *Mycobacterium bovis* (BCG). Lepr. Rev. 57(Suppl. 2):67-75.
3. Engers, H. D., B. R. Bloom, and T. Godal. 1984. Monoclonal antibodies against mycobacterial antigens. Immunol. Today 6: 342-348.
4. Hedstrom, R., J. Culpepper, R. A. Harrison, N. Agabian, and G. Newport. 1987. A major immunogen in *Schistosoma mansoni* infections is homologous to the heat shock protein HSP70. J. Exp. Med. 165:1430-1435.
5. Ingolia, T. D., and E. A. Craig. 1982. Four small *Drosophila* heat shock proteins are related to each other and to mammalian α -crystallin. Proc. Natl. Acad. Sci. USA 79:2360-2364.
6. Lamb, J. R., J. Ivanyi, A. D. M. Rees, J. R. Rothbard, K. Howland, R. A. Young, and D. B. Young. 1987. Mapping of T cell epitopes using recombinant antigens and synthetic peptides. EMBO J. 6:1245-1249.
7. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
8. Messing, J. 1983. New M13 vectors for cloning. Methods Enzymol. 101:20-78.
- 8a. Mustafa, A. S. 1988. Identification of T cell activating recombinant antigens shared between three candidate anti-leprosy vaccines, killed *M. leprae*, *Mycobacterium bovis* BCG and *Mycobacterium w.* Int. J. Leprosy 56:265-273.
9. Mustafa, A. S., H. K. Gill, A. H. Nerland, W. J. Britton, V. Mehra, B. R. Bloom, R. A. Young, and T. Godal. 1986. Human T cell clones recognize a major *M. leprae* protein antigen expressed in *E. coli*. Nature (London) 319:63-66.
10. Mustafa, A. S., F. Oftung, H. K. Gill, and I. Natvig. 1986. Characteristics of human T-cell clones from BCG and killed *M. leprae* vaccinated subjects and tuberculosis patients. Recognition of recombinant mycobacterial antigens. Lepr. Rev. 57(Suppl. 2):123-130.
11. Nagao, R. T., E. Czarnecka, W. B. Gurley, F. Schoffl, and J. L. Key. 1985. Genes for low-molecular-weight heat shock proteins of soybeans: sequence analysis of a multigene family. Mol. Cell. Biol. 5:3417-3428.
12. Offenhoff, T. H. M., P. R. Klaster, J. Ivanyi, B. G. Elferink, M. Y. L. DeWit, and R. R. P. DeVries. 1986. *Mycobacterium leprae* specific protein antigens defined by cloned human helper T cells. Nature (London) 319:66-68.
13. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain termination inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
- 13a. Selkirk, M. E., P. J. Rutherford, D. A. Danham, F. Partono, and R. M. Maizels. 1987. Cloned antigen genes of *Brugia filarial* parasite. Biochem. Soc. Symp. 53:91-102.
14. Shinnick, T. M., M. H. Vodkin, and J. C. Williams. 1988. The *Mycobacterium tuberculosis* 65-kilodalton antigen is a heat shock protein which corresponds to common antigen and to the *Escherichia coli* GroEL protein. Infect. Immun. 56:446-451.
15. Vodkin, M. H., and J. C. Williams. 1988. A heat shock operon in *Coxiella burnetii* produces a major antigen homologous to a protein both in mycobacteria and *Escherichia coli*. J. Bacteriol. 170:1227-1234.
16. Young, D., R. Lathigra, R. Hendrix, D. Sweetser, and R. A. Young. 1988. Stress proteins are common targets in leprosy and tuberculosis. Proc. Natl. Acad. Sci. USA 85:4267-4270.
17. Young, R. A., V. Mehra, D. Sweetser, T. M. Buchanan, J. Clark-Curtiss, R. W. Davis, and B. R. Bloom. 1985. Genes for the major protein antigens of the leprosy parasite *Mycobacterium leprae*. Nature (London) 316:450-452.