JOURNAL OF BACTERIOLOGY, Feb. 1992, p. 1352–1359
0021-9193/92/041352-08$02.00/0

Vol. 174, No. 4

The 14,000-Molecular-Weight Antigen of *Mycobacterium tuberculosis* Is Related to the Alpha-Crystallin Family of Low-Molecular-Weight Heat Shock Proteins

ANNELIES VERBON,[1,2] RUDY A. HARTSKEERL,[1]* ANJA SCHUITEMA,[1] AREND H. J. KOLK,[1] DOUGLAS B. YOUNG,[3] AND RAJU LATHIGRA[3]

N. H. Swellengrebel Institute of Tropical Hygiene, Royal Tropical Institute, Meibergdreef 39,[1] and Division of Pulmonary Diseases, Academic Medical Center,[2] Amsterdam, The Netherlands, and MRC Tuberculosis and Related Infections Unit, Hammersmith Hospital, London W12 OHS, United Kingdom[3]

Received 12 August 1991/Accepted 10 December 1991

Eight monoclonal antibodies (MAbs) directed against the 14,000-molecular-weight (14K) antigen of *Mycobacterium tuberculosis* reacted specifically with mycobacteria of the *M. tuberculosis* complex. The nucleotide sequence of the gene encoding the 14K antigen was determined by using recombinant DNA clones isolated from lambda gt11 and cosmid libraries of the *M. tuberculosis* genome. The DNA sequence of the 14K protein gene coded for a polypeptide of 144 amino acids with a calculated molecular mass of 16,277 Da. The 14K antigen has a marked homology with proteins belonging to the alpha-crystallin family of low-molecular-weight heat shock proteins, which includes the 18K antigen of *M. leprae*. The eight MAbs recognized at least four distinct epitopes localized within the following three regions of the 14K protein: amino acids 10 to 92 (MAbs F67-8 and F67-16), amino acids 41 to 92 (F159-1 and F159-11), and amino acids 41 to 144 (F23-41, F24-2, F23-49, and TB68).

Mycobacterial diseases continue to be an important cause of human morbidity and mortality. It is estimated that tuberculosis is responsible for as many as 26% of all avoidable adult deaths worldwide (34). For an approach in the development of novel vaccines and diagnostic tests for mycobacterial diseases, a major effort has been invested in the identification and characterization of mycobacterial antigens involved in interactions within the host immune system (for a review, see reference 52). A set of prominent antigens from *Mycobacterium tuberculosis* and *Mycobacterium leprae* have been identified by using murine monoclonal antibodies (MAbs) (14, 15). Several of these antigens have been shown to be major targets of antibody and T-cell responses in patients with mycobacterial diseases (for reviews, see references 25 and 52). Sequence analysis revealed an intriguing relationship between a number of the prominent antigens and members of conserved heat shock protein families (53). Antigens belonging to the DnaK (hsp70), GroEL (hsp60), and GroES heat shock protein families have been identified in both *M. tuberculosis* (2, 17, 40, 53) and *M. leprae* (17, 23, 33, 53). In addition, the 18,000-molecular-weight (18K) antigen of *M. leprae* has been shown to share sequence features with alpha-crystallin and a family of low-molecular-weight heat shock proteins (35). Here, we describe the corresponding 14K antigen of *M. tuberculosis*.

This antigen is one of the prominent antigens of *M. tuberculosis* defined by MAbs (15). It was originally identified by three MAbs generated in two separate laboratories (8, 29) and has been previously assigned an apparent molecular weight of 16,000 (29, 49). The antigen has shown considerable promise as a serodiagnostic target in assay protocols based on MAb competition and direct enzyme-linked immunosorbent assay (ELISA) formats (7, 26). Recombinant clones expressing the 14K antigen have been selected from lambda gt11 expression libraries using monoclonal and polyclonal antisera (41, 55, 56), and recognition of the recombinant antigen by murine T cells has been demonstrated (27).

In this paper, we describe the molecular and immunological characterization of the 14K protein and report its link with the alpha-crystallin family of low-molecular-weight heat shock proteins.

MATERIALS AND METHODS

Bacterial strains, bacteriophages, and plasmids. Bacterial strains, bacteriophages, and plasmids used in this study are listed in Table 1. Isolated mycobacterial strains, *Nocardia asteroides* and *Escherichia coli* wild type (Table 2) were described previously (21, 39, 49). *E. coli* K-12 strain PC2495 was used as a host for plasmid pBluescript KS+ and derivatives, and strain POP2136 was used as a host for plasmid pEX2 and derivatives. M13mp18 derivatives generated for the shotgun sequencing were propagated on *E. coli* TG1. Bacteriophage lambda gt11 and derivatives were propagated on strain Y1090, and lysogens were constructed by using strain Y1089 as a host. Recombinant phage 60.1.1 and recombinant plasmid pPH6001 were kind gifts from P. Hermans, and phage Y3155 was from R. A. Young. The Lawrist4::*M. tuberculosis* library was a kind gift from N. G. Stoker.

Media and reagents. Mycobacterial strains were cultured on Löwenstein-Jensen medium at 37°C and harvested as described by Schöningh et al. (39). The mycobacteria and bacteria were sonicated and used in an immunoblot as described previously (48). Luria-Bertani (LB) medium and LB agar were used for growing *E. coli* K-12 strains (37). Strain PC2495 was grown on minimal medium (12). When appropriate, ampicillin, kanamycin sulfate, streptomycin

* Corresponding author.

TABLE 1. *E. coli* K-12 strains, bacteriophages, and plasmids

Strain, phage, or plasmid	Relevant properties	Source (reference)
Strains		
Y1089	Δ*lacU169* Δ*lon hflA150*(pMC9)	R. Young (56)
Y1090	Δ*lacU169* Δ*lon supF*(pMC9)	R. Young (56)
PC2495	JM101 *hsdR recA*	PC[a]
POP2136	*c*I857	A. Raibaud
TG1	Δ(*lac pro*) *supF hsdD5* (F *traD36 proAB lacI*^{q} *lacZ* ΔM15)	J. Sambrook (37)
Phages		
Lambda gt11		R. Young (56)
M13mp18		C. Yanisch (51)
Y3155	Lambda gt11 recombinant[b]	A. Kingston (27)
60.1.1	Lambda gt11 recombinant[b]	P. Hermans
Plasmids		
pMC9	Ap^r pBR322-*lacI*	M. Calos (6)
pEX2	Ap^r contains p_R promoter of phage lambda and *cro-lacZ* gene fusion	K. Stanley (44)
pPH6001	pEX2 recombinant[b]	P. Hermans
pTHL1060	pEX2 recombinant[b]	This study
pTHT1031	pEX2 recombinant[b]	This study
pBluescript KS+	Ap^r	Stratagene
pTHT1012	pBluescript recombinant[b]	This study
pTHT1014	pBluescript recombinant[b]	This study
pTHT1023	pBluescript recombinant[b]	This study
pTHT1027	pBluescript recombinant[b]	This study
pUC18	Ap^r	C. Yanisch (51)
pRL14.3	pUC18 recombinant[b]	This study
pLO14.1BAM	pUC18 recombinant[c]	This study

[a] PC, Phabagen Collection, Department of Molecular Cell Biology, State University of Utrecht, The Netherlands.
[b] Contains part of the 14K protein gene.
[c] Contains a 4-kb *M. tuberculosis* fragment comprising the 14K protein gene.

sulfate, and tetracycline were added as previously described (20). Restriction enzymes and T4 DNA ligase were from Boehringer GmbH (Mannheim, Germany), New England Biolabs, Inc. (Beverly, Mass.), and Pharmacia LKB (Uppsala, Sweden). Horseradish peroxidase-labeled anti-mouse immunoglobulin G (heavy and light chains) was from the Pasteur Institute, Paris, France.

Murine MAbs. MAbs TB68, F23-41, F23-49, and F24-2 have been described previously (8, 29). For production of four other MAbs, F67-8, F67-16, F159-1, and F159-11, two kinds of protocols were used for immunization of BALB/c mice. For F159-1 and F159-11, the mice were immunized with live *M. tuberculosis*, isolated from a patient. The mice were injected according to the following protocol: on day 0, 5×10^7 mycobacteria emulsified in incomplete Freund adjuvant were injected intraperitoneally; on day 117, 4 days before fusion, the mice received an intraperitoneal booster injection of 500 μg (dry weight) of sonicated *M. tuberculosis*. For F67-8 and F67-16, 100 μg (dry weight) of supernatant (100,000 × *g*) of the same sonicated strain used for F159-1 and F159-11 was injected intraperitoneally on day 0. On day 129, 4 days before fusion, 250 μg (dry weight) of supernatant of the sonicate was given intraperitoneally. Hybridomas were selected by testing the culture supernatant in an ELISA with sonicated mycobacteria and a Western blot (immunoblot) as previously described (28).

DNA technology. Standard procedures were used for the preparation of phage and plasmid DNA, restriction enzyme digestion, ligation, transformation, and transduction (37). Genomic mycobacterial DNA was isolated as described by Hartskeerl et al. (19). DNA sequencing of the pEX2 and pBluescript KS+ derivatives was done by the dideoxy chain termination method (38) with a T7 sequencing kit (Pharmacia). For DNA sequencing, plasmid DNAs were isolated by alkaline extraction (3) and further purified either by CsCl gradient centrifugation or by use of the Geneclean kit (Bio 101, Inc., La Jolla, Calif.). The DNA region comprising the coding sequence of the 14K protein was sequenced by two methods. One strategy was based on sequencing the DNA inserts of the subclones pPH6001, pTHL1060, pTHT1014, pTHT1023, and pTHT1027 with the oligonucleotide primers 86.25-2 and 17.5.88-2 for pEX2 derivatives and primers T3 and T7 (Stratagene, La Jolla, Calif.) for pBluescript KS+ derivatives. Additionally, synthetic primers were used so that the complete sequence was determined by using overlapping fragments on both strands.

The other strategy involved the shotgun sequencing of randomly generated fragments following sonication of the 2.0-kb *Eco*RI-*Pst*I fragment and ligation into the *Sma*I site of M13mp18. Sequences of the short fragments were assembled into a contiguous stretch by using the DBAUTO and DBU TIL programs in the Staden-Plus software. Sequence data were also processed with DNASIS/PROSIS software (Pharmacia), and the PIR (NBRF R25.0) and SWISS-PROT

TABLE 2. Reactivity of MAbs directed against the 14K antigen of *M. tuberculosis* by Western blot

Bacterium	Reactivity with MAbs[a]:							
	F23-41 IgG2b	F23-49 IgG2a	F24-2 IgG1	F67-8 IgG1	F67-16 IgG1	F159-1 IgG2a	F159-11 IgG2a	TB68 IgG1
M. africanum	16, 32[b]	16, 40[c]	16, 32	16, 32	16, 32	16	16	16, 32
M. bovis	16	16, 40	16, 32	16	16	16[d]	16[d]	16
M. bovis BCG	16	16, 40	16, 32	16	16	16[d]	16[d]	16
M. tuberculosis	16	16, 40	16, 32	16	16	16	16	16
M. tuberculosis H37Rv	16, 32	16, 40	16, 32	16, 32	16, 32	16	16	16, 32
M. tuberculosis H37Ra	16, 32	16, 40	16, 32	16, 32	16, 32	16	16	16, 32
M. scrofulaceum	17[d]			17[d]	17[d]			

[a] MAbs F23-41, F23-49, F24-2, and TB68 have been described previously (8, 29). MAbs F67-8, F67-16, F159-1, and F159-11 are described in Materials and Methods. Shown are molecular weights of the protein bands reacting. IgG1, IgG2, and IgG2b are immunoglobulin G subclasses. The following bacteria showed no reactivity: *M. avium*, *M. flavescens*, *M. gastri*, *M. gordonae*, *M. kansasii*, *M. leprae* 5-3, *M. marinum*, *M. nonchromogenicum*, *M. smegmatis*, *M. terrae*, *M. xenopi*, *E. coli* wild type, and *N. asteroides*.
[b] The 32K protein bands was a double band with a molecular weight of 32,000 to 33,000 and gave a weak reaction.
[c] Weak reaction with the 40,000-molecular-weight band of all strains of the *M. tuberculosis* complex.
[d] Weak reaction.

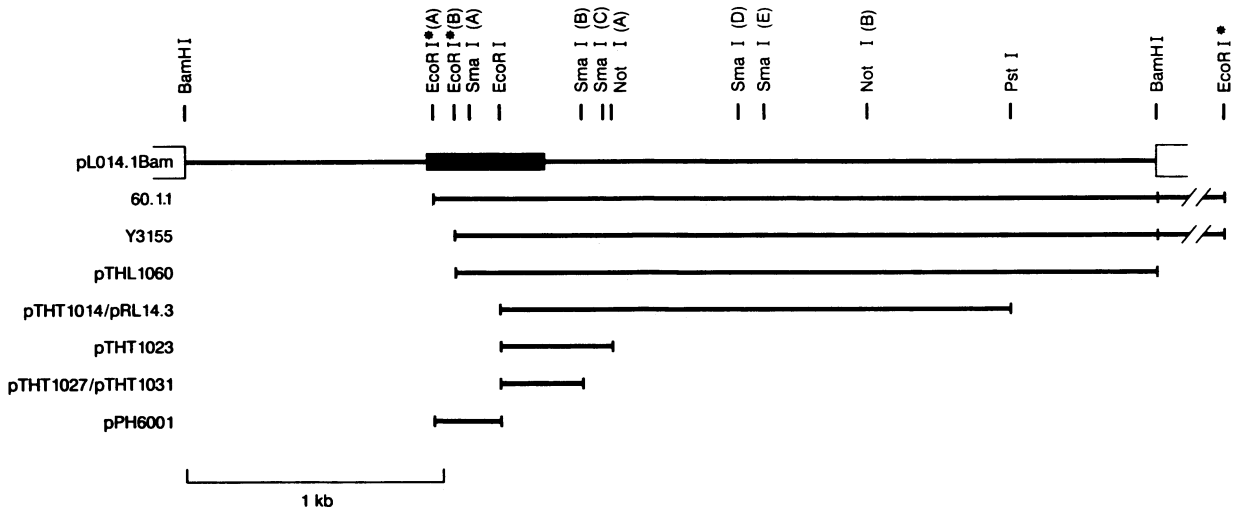

FIG. 1. Physical maps of recombinant plasmids and phages. Construction of the various recombinant plasmids has been described in Materials and Methods. The open brackets of pLO14.1 represent vector DNA. The position of the 14K protein gene is indicated by a solid bar. Transcription of the gene is from left to right. *Eco*RI* indicates the ends of the DNA inserts of recombinant phages 60.1.1 and Y3155 at the left end of the vector lambda gt11. *Eco*RI*(A) and *Eco*RI*(B) delimit the ends of the inserts at the right ends of lambda gt11 for phages 60.1.1 and Y3155, respectively. *Sma*I and *Not*I sites are distinguished by a suffix (A, B, etc.) for sake of a clear description of the construction of the various hybrid plasmids in the text.

(R14.0) data bases were used for protein homology searches. Pairwise homology percentages were calculated with UWGCG software (13). Multiple alignment was done as described by Corpet (9).

The 5′ end of the gene was obtained from cosmid clone LO14, selected from the Lawrist4::*M. tuberculosis* library, as described below. The N-terminal amino acid sequence of the 14K protein from a clinical isolate of *M. tuberculosis* (48) purified by preparative sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was kindly determined by J. Walker (MRC Laboratory of Molecular Biology, Cambridge, United Kingdom).

Southern blotting (43) and hybridization were done basically as described by Van Eys et al. (47). Preparation of ^{32}P-labeled DNA probes was done by using a random-priming DNA labeling kit (Boehringer) according to the instructions of the manufacturer. DNA amplifications were done as previously described (19) by using heat-stable *Taq* DNA polymerase (Perkin-Elmer Cetus). Oligonucleotides 1060R (5′CAAGCTTTACCGGCAGCGACACCGTGC, homologous to the region from nucleotides 314 to 333) and either the *lacZ* forward primer (5′AAAGCTTGCGACGA CTCCTGGAGCCCG) or the *lacZ* reverse primer (5′CAA GCTTGTAATGGTAGCGACCGGCGC), all containing a *Hin*dIII site, were used as the sets of primers. A reaction mixture (1 ng/ml) of recombinant phages Y3155 and 60.1.1 was used as the target DNA.

Oligonucleotides were synthesized on a 381A DNA synthesizer (Applied Biosystems) and used without further purification.

Screening of cosmid library. Cosmid clones containing the complete 14K gene were selected by screening a Lawrist4::*M. tuberculosis* H37Rv library. A total of 375 cosmid clones were screened by colony lysis and hybridization by using the ^{32}P-labeled 2.0-kb *Eco*RI-*Pst*I fragment of pRL14.3 as a probe. Two clones designated LO14.1 and LO14.9 were isolated.

Subcloning of the 14K protein gene. Recombinant plasmid pPH6001 consists of the 255-bp *Eco*RI*(A)-*Eco*RI fragment of recombinant phage 60.1.1 (Fig. 1) ligated into the *Eco*RI site of plasmid vector pEX2 in the proper orientation with respect to the direction of transcription of *cro-lacZ*. To construct hybrid plasmid pTHL1060, the recombinant phage Y3155 was digested with *Bam*HI and partially digested with *Eco*RI. The 2.9-kb *Eco*RI*(B)-*Bam*HI fragment (Fig. 1) was subcloned into the *Eco*RI-*Bam*HI site of vector pEX2. Recombinant plasmid pTHT1012 was constructed by subcloning the 156-bp *Eco*RI*(B)-*Eco*RI fragment of Y3155 into the *Eco*RI site of pBluescript. The hybrid plasmids pTHT1014 and pRL14.3 were constructed by subcloning the 2.0-kb *Eco*RI-*Pst*I fragment from phage Y3155 into the corresponding sites of pBluescript and pUC18, respectively. The 0.45-kb *Eco*RI-*Not*I(A) fragment was subcloned into the *Eco*RI-*Not*I sites of pBluescript, resulting in plasmid pTHT1023, and the 0.32-bp *Eco*RI-*Sma*I(B) fragment was subcloned into the *Eco*RI-*Sma*I site of pBluescript and pEX2, resulting in plasmids pTHT1027 and pTHT1031, respectively.

Production of recombinant 14K proteins. The expression of *lacZ-14K* hybrid genes from phages Y3155 and 60.1.1 was done by the method of Young et al. (56). Expression of the *cro-lacZ-14K* hybrid genes from plasmids pPH6001, pTHT1031, and pTHL1060 was done as described by Zabeau and Stanley (57). The production of proteins was analyzed by SDS-PAGE (30). Gels containing 8 and 13% acrylamide were used. Western blotting was done as previously described (28).

Mapping of B-cell epitopes. The reactivity of MAbs F23-41, F23-49, F24-2, F67-8, F67-16, F159-1, F159-11, and TB68 with various LacZ-14K and Cro-LacZ-14K fusion proteins was determined by Western blot analysis.

RESULTS

Specificity of MAbs directed against the 14K antigen of *M. tuberculosis*. Western blot analysis was used to determine the reactivity pattern of eight MAbs directed against the 14K antigen with several mycobacterial and nonmycobacterial

```
     SD
ATTAGGAGGCATCAAATGGCCACCACCCTTCCCGTTCAGCGCCACCCGCGGTCCCTCTTC  60
                  M  A  T  T  L  P  V  Q  R  H  P  R  S  L  F   15

CCCGAGTTTTCTGAGCTGTTCGCGGCCTTCCCGTCATTCGCCGGACTCCGGCCCACCTTC 120
 P  E  F  S  E  L  F  A  A  F  P  S  F  A  G  L  R  P  T  F   35

GACACCCGGTTGATGCGGCTGGAAGACGAGATGAAAGAGGGGCGCTACGAGGTACGCGCG 180
 D  T  R  L  M  R  L  E  D  E  M  K  E  G  R  Y  E  V  R  A   55

GAGCTTCCCGGGGTCGACCCCGACAAGGACGTCGACATTATGGTCCGCGATGGTCAGCTG 240
 E  L  P  G  V  D  P  D  K  D  V  D  I  M  V  R  D  G  Q  L   75

ACCATCAAGGCCGAGCGCACCGAGCAGAAGGACTTCGACGGTCGCTCGGAATTCGCGTAC 300
 T  I  K  A  E  R  T  E  Q  K  D  F  D  G  R  S  E  F  A  Y   95

GGTTCCTTCGTTCGCACGGTGTCGCTGCCGGTAGGTGCTGACGAGGACGACATTAAGGCC 360
 G  S  F  V  R  T  V  S  L  P  V  G  A  D  E  D  D  I  K  A  115

ACCTACGACAAGGGCATTCTTACTGTGTCGGTGGCGGTTTCGGAAGGGAAGCCAACCGAA 420
 T  Y  D  K  G  I  L  T  V  S  V  A  V  S  E  G  K  P  T  E  135

AAGCACATTCAGATCCGGTCCACCAACTGACCACTGGGTCCGTGCTGATGACCG        474
 K  H  I  Q  I  R  S  T  N  *                 *  *           144
```

FIG. 2. Nucleotide and deduced amino acid sequences of the coding region of the 14K protein gene. The putative Shine-Dalgarno (SD) sequence is underlined, and termination codons are indicated by asterisks.

strains (Table 2). All MAbs reacted with mycobacteria from the *M. tuberculosis* complex, i.e., *M. tuberculosis*, *M. bovis*, and *M. africanum*. None of the MAbs reacted with either *E. coli* or *N. asteroides*. F23-41, F67-8, and F67-16 also reacted weakly with *M. scrofulaceum*, but with an antigen of a higher apparent molecular weight (about 17,000). No reaction of the eight MAbs was seen with the other mycobacterial strains tested. In addition to the major 14K antigen band, six of the MAbs showed some weak reaction with a 40,000-molecular-weight protein band or a 32,000- to 33,000-molecular-weight double band on Western blots. This reactivity was also restricted to the *M. tuberculosis* complex. Thus, the 14K antigen contains B-cell epitopes specific for the *M. tuberculosis* complex.

Characterization of recombinant phages expressing the 14K antigen. Both recombinant phages Y3155 and 60.1.1, selected from the lambda gt11::*M. tuberculosis* library (56) by MAbs TB68 and F67-8, respectively, expressed the 14K antigen as a fusion protein with β-galactosidase. Restriction enzyme analysis of the DNA inserts revealed that both recombinant phages consisted of large and small *Eco*RI fragments, i.e., Y3155 contained 4.5- and 0.16-kb fragments, and 60.1.1 contained 4.2- and 0.26-kb fragments. DNA amplification of phages Y3155 and 60.1.1 with primers *lacZ* forward and 1060R gave bands of about 250 and 350 bp, respectively. No bands of the proper size were observed when 1060R was used in combination with the *lacZ* reverse primer. This indicated that both small *Eco*RI fragments were positioned adjacent to the 5′ part of the *lacZ* gene and thus encoded an N-terminal part of the 14K protein, whereas the C-terminal part was located on the large *Eco*RI fragments. Homology between the DNA inserts of both phages was confirmed by Southern blot analysis. The 0.26-kb fragment hybridized with the 0.16-kb fragment and, as expected, not with the large *Eco*RI fragments of both phages. Additionally, the 2.0-kb *Eco*RI-*Pst*I DNA insert from pTHT1014 hybridized to the original 4.5-kb fragment of Y3155 and to the 4.2-kb fragment of 60.1.1, but not with the small *Eco*RI fragments of both phages (data not shown). From these results, we conclude that phages Y3155 and 60.1.1 contain similar DNA fragments and express the same antigen (Fig. 1).

Selection of recombinant DNA clones containing the complete 14K protein gene. Since recombinant phages Y3155 and 60.1.1 did not comprise the complete 14K protein gene, we screened the Lawrist4::*M. tuberculosis* library to select cosmid recombinants containing the missing 5′ region of the 14K protein gene. Two identical cosmid clones were selected, and clone LO14.1 was further characterized. Southern blot analysis using the 2.0-kb *Eco*RI-*Pst*I insert of pRL14.3 indicated that an identical-size *Eco*RI-*Pst*I fragment from strain H37Rv in LO14.1 hybridized with the probe (data not shown). On the basis of restriction enzyme and Southern blot analyses, a partial map of the cosmid LO14.1 around the region of the 14K gene was constructed (Fig. 1). The 4-kb *Bam*HI fragment from LO14.1 was isolated and subcloned into pUC18 cut with the same enzyme to generate clone pLO14.1Bam.

Nucleotide sequence determination of the 14K protein gene. The composite sequence of the 14K gene depicted in Fig. 2 was obtained from the mycobacterial DNA inserts of the LO14.1Bam clone for the 5′ region and of phages Y3155 and 60.1.1 for the 3′ region. No changes at the nucleotide level of a 300-bp region of the 14K protein gene obtained either from 60.1.1 or from the clone pLO14.1Bam were observed (data not shown). The direction of transcription and the reading frame of the 14K protein gene corresponded to the reading frame of *lacZ* in phages Y3155 and 60.1.1. Translation of the sequence indicated that the initiation codon utilized was an ATG positioned at nucleotides 16 to 18 (Fig. 2). The N-terminal amino acid sequence of the purified 14K protein was established as ATTLPV, confirming the start of translation at this initiation codon. A potential Shine-Dalgarno sequence, AGGAGG, was found 7 bp upstream of the ATG initiation codon. A translation stop codon, TGA, is located at position 433. Thus, the 14K protein is composed of a polypeptide of 144 amino acids and has a calculated molecular mass of 16,277 Da, which corresponds very well with the established apparent molecular weight of 16,000, as described by Kolk et al. (29). We therefore propose that the name for this antigen should be reclassified as the 16K antigen. However, for the sake of consistency, we will continue to refer to the antigen as the 14K protein in this paper. The base composition of the gene was 61.8% guanine

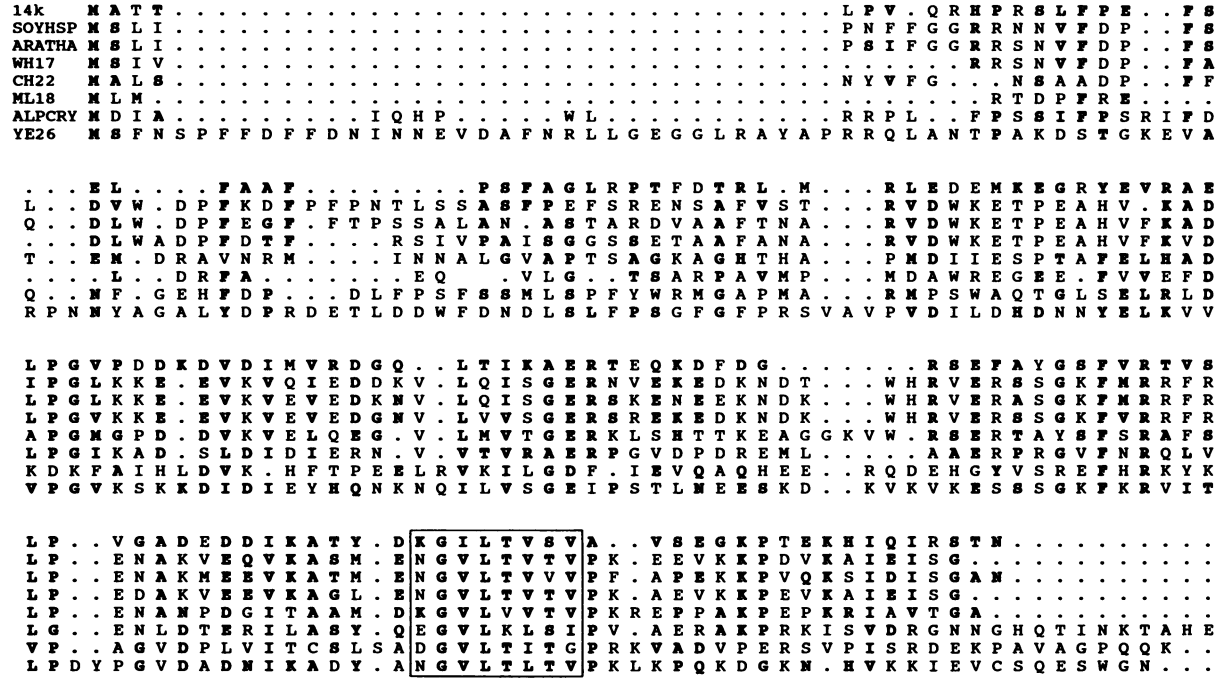

FIG. 3. Multiple sequence alignment of the 14K antigen with various low-molecular-weight heat shock proteins of the alpha-crystallin family. Protein codes shown on the left are those shown in Table 3. Residues identical with the 14K antigen and conserved amino acid replacements are shown in bold typeface. The conserved motif is boxed. Four C-terminal amino acids of ML18 are omitted.

and cytosine. Within the individual codons, however, the bases were not evenly distributed. The mean G+C content of the first two bases was 52.8%, and the mean G+C content of the third base was 79.9%.

A search for sequence homology indicated that the 14K protein had significant homology with a number of proteins belonging to the alpha-crystallin family of low-molecular-weight heat shock proteins (Fig. 3). Percentages of identity between the 14K protein and various alpha-crystallin heat shock proteins are shown in Table 3.

Mapping of B-cell epitopes on the 14K antigen. To localize the B-cell epitopes of the eight MAbs on the 14K antigen, we determined the reactivity of these MAbs with lysates of *E. coli* Y1089 lysogenized with phages Y3155 and 60.1.1 and POP2136 carrying plasmid pPH6001, pTHL1060, or pTHT1031 by Western blot analysis (Fig. 4). For unknown reasons, none of the eight MAbs reacted with the lysate of POP2136 carrying pTHT1031 (not shown). The β-galactosidase fusion protein encoded by pTHT1031 may be unstable in *E. coli*. The other four recombinant DNA clones expressed parts of the 14K protein as stable fusion proteins with β-galactosidase. MAbs F67-8 and F67-16 bound to fusion proteins which have the region from amino acids 10 to 92 (aa 10-92) in common (Fig. 4). The MAbs did not bind to fusion proteins that do not contain the region aa 10-41, indicating that this region is essential for binding. Thus, the putative epitopes are positioned within the region aa 10-92 and are likely located, partially or completely, within region aa 10-41. MAbs F23-41, F23-49, F24-2, and TB68 reacted with all fusion proteins comprising region aa 41-144. No reaction was found with the fusion protein from pPH6001 that was missing region aa 92-144. Therefore, it is concluded that the putative epitopes of MAbs F23-41, F23-49, F24-2, and TB68 are positioned within region aa 41-144 and likely, at least partially, within region aa 92-144. MAbs F159-1 and F159-11 bound to all fusion proteins that had region aa 41-92 in common, implying that their epitopes are located within this region.

TABLE 3. Sequence identity of the 14K antigen of *M. tuberculosis* with low-molecular-weight heat shock proteins belonging to the alpha-crystallin family

Sequence[a]	% Identity	Reference(s)
WH17	31.6	32
YE26	31.2	5
CH22	31.0	18
ARATHA	29.4	46
ML18	28.9	4, 35
SOYHSP	26.9	36
ALPCRY	21.4	11

[a] WH17, low-molecular-weight heat shock protein of wheat; YE26, yeast small heat shock protein hsp26; CH22, 22K heat shock protein of *Chlamydomonas* spp.; ARATHA, low-molecular-weight heat shock protein of *Arabidopsis thalania*; ML18, *M. leprae* 18K antigen; SOYHSP, soybean low-molecular-weight heat shock protein; ALPCRY, spiny dogfish alpha-crystallin B chain.

DISCUSSION

To characterize the 14K antigen of *M. tuberculosis*, we analyzed two lambda gt11 recombinant phages, Y3155 and 60.1.1, selected with MAbs TB68 and F67-8, respectively. Both recombinant phages expressed the 14K protein gene as a fusion protein with β-galactosidase with an apparent molecular weight of approximately 132,000. Restriction enzyme analysis and DNA hybridization experiments revealed that phage 60.1.1 contained an extra 5′ region of about 100 bp with respect to phage Y3155. The complete gene was obtained from a Lawrist4 cosmid library constructed with

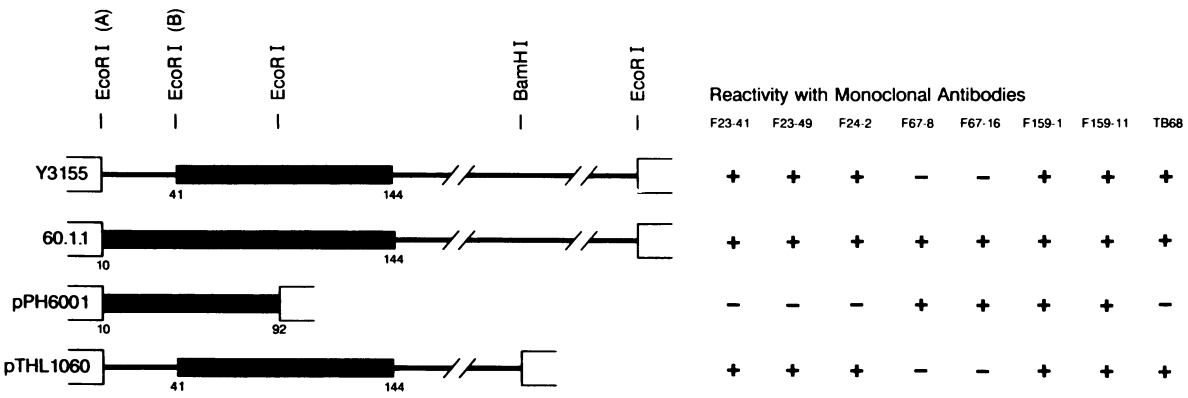

FIG. 4. Physical maps of relevant parts of recombinant phages Y3155 and 60.1.1 as well as of plasmids pPH6001 and pTHL1060 expressing various regions of the *M. tuberculosis* 14K protein gene and reactivity of MAbs to proteins expressed by these plasmids. Open brackets, vector part of recombinant DNAs; solid bars, 14K protein gene. The numbers in the 14K protein gene indicate the amino acid residues in the 14K protein sequence which delimit the deletions. The reactivity of MAbs to the protein from induced Y1089 cells harboring the phages and POP2136 harboring the plasmid was established by Western blot analysis.

genomic DNA of *M. tuberculosis* H37Rv. Since the lambda gt11 library has been constructed with DNA from strain Erdman, the sequence of the 14K protein gene presented here is composed of nucleotide sequences from two different strains. We do not expect that the genes from both strains contain major differences. First, all eight *M. tuberculosis* complex-specific MAbs directed against the 14K antigen reacted with the 14K protein of both strains. Second, the sequence of a region of about 300 bp of the 14K protein gene deduced from both strains was identical. It is likely that the 5′ part of the sequence deduced from strain H37Rv is representative of the corresponding part of the gene of strain Erdman, although the presence of one or more minor differences cannot be excluded.

The 14K protein gene encodes a protein of 144 amino acids with a calculated molecular mass of 16,277 Da. The overall G+C content and the distribution of these residues in the codons as well as the codon usage correspond very well with those described for *M. tuberculosis* (10).

A number of other lambda gt11 clones selected by MAbs to the 14K antigen were reported to express recombinant products with molecular weights in the range of 25,000 to 30,000 (41, 55, 56). A comparison of physical maps of such clones (41, 55) with those of phages Y3155 and 60.1.1 shown in this paper indicates that they should have expressed the antigen as β-galactosidase fusion proteins. It might be that the anomalous products reported correspond to partially degraded protein fragments.

A search for sequence homology revealed that the 14K protein belongs to the alpha-crystallin family of low-molecular-weight heat shock proteins. However, analysis of protein synthesis by metabolic labeling of *M. tuberculosis* provided no evidence for increased synthesis of the 14K antigen during heat shock (54). While it remains possible that stresses other than heat may induce its expression, it should be noted that membership of a heat shock family does not necessarily imply participation in the heat shock response.

Like many other heat shock proteins, members of the alpha-crystallin family can be found in a variety of eukaryotes, i.e., in plant cells (18, 32, 36), various animal and insect cells (humans, hsp27; *Drosophila* sp., hsp26; spiny dogfish, alpha-crystallin) (11, 22, 24), and in yeast cells (hsp26) (5). However, in contrast to the 70K (DnaK and hsp70) and 60K (GroEL and hsp60) families, members of the alpha-crystallin family have not been reported as major proteins in prokaryotes (50). It is striking that the only two clear examples of bacterial members of this family are the 18K antigen of *M. leprae* (4, 35) and the 14K protein described here. An inspection of sequence similarities indicates that the *M. tuberculosis* and *M. leprae* antigens are no more related to each other (29% amino acid identity) than each is to, for example, the corresponding soybean protein (27% identity with the *M. tuberculosis* 14K protein and 31% identity with the *M. leprae* 18K protein). This is quite different from the situation with other mycobacterial heat shock proteins, DnaK, GroEL, and GroES, which share more than 90% identity between species (17, 23, 42). This leads us to conclude that these two antigens are probably not closely related and should be considered as quite distinct proteins.

A common feature of the low-molecular-weight heat shock proteins is a well-conserved short sequence near the C terminus with a derived consensus sequence, D/N-G-V-L-T-I/V-X-V/A (18), corresponding to a sequence in the alpha-crystallin protein. Both the 14K protein of *M. tuberculosis* and the 18K protein of *M. leprae* contain this region, i.e., KGILTVSV, aa 118-125, and EGVLKLSI, aa 108-115, respectively (Fig. 3). This might indicate that the mycobacterial proteins have a function in common with the other alpha-crystallin heat shock proteins or that they are evolved from one ancestral molecule. Although extensively studied, the function of the small heat shock proteins has not yet been clarified (45). The 14K protein is not secreted (1, 48). However, the antigen is detectable with MAbs F23-49 and F24-2 by immunofluorescence (49) and ELISA analyses with whole mycobacteria (39), implying that it is exposed at the outside of the cell wall. Usually a signal peptide is present when proteins are transported across a membrane. Sequence analysis of the 14K protein gene, however, did not reveal the presence of such a signal peptide, suggesting that the antigen is translocated without processing. Membrane translocation without processing is not unique to the 14K protein but has also been described for another alpha-crystallin heat shock protein, the chloroplast 22-kDa protein of *Chlamydomonas* sp. (18).

We have been interested in proteins which distinguish mycobacteria of the *M. tuberculosis* complex from other mycobacteria. Such proteins could be used as antigens in a serological test either for the detection of antibodies in sera

from patients with tuberculosis (25) or for the identification of early cultures (40). Falla et al. (16) have recently described another 14K protein which appears to be *M. tuberculosis* specific. The amino acid sequence of this protein is completely different from the protein sequence reported in this paper.

Western blot analysis of the eight MAbs directed against the 14K antigen of *M. tuberculosis* revealed that these MAbs reacted almost exclusively with mycobacteria of the *M. tuberculosis* complex. Thus, although the 14K antigen is homologous to various antigens belonging to the alpha-crystallin family of heat shock proteins, it contains several *M. tuberculosis* complex-specific B-cell regions. On the basis of three different reactivity patterns with the fusion proteins, the putative epitopes of the eight MAbs were mapped at three distinct regions. F24-2 reacts with a different epitope on the 14K antigen than do F23-49 and TB68 (15). This indicates that at least four distinct murine B-cell epitopes are present on the 14K antigen. Also, in the human antibody response to the 14K antigen, more than one specific immunodominant epitope seems to be involved (26). The 14K protein is also recognized by T-helper cell clones (31). Synthesis of peptides of the 14K antigen may be useful for mapping B- and T-cell epitopes more precisely and for studying their value for diagnostic assays and vaccine development.

ACKNOWLEDGMENTS

We thank R. A. Young for phage Y3155 and P. Hermans for phage 60.1.1 and plasmid pPH6001. We thank N. G. Stoker, J. Thompson, and K. de Smet for the construction and the kind gift of the Lawrist4::*M. tuberculosis* H37Rv cosmid library and J. Walker for the N-terminal amino acid determination. We are grateful to C. Hermans, J. van Leeuwen, and S. Kuijper for excellent technical assistance and to I. Struiksma for assistance in preparing the manuscript.

A.V. was supported by a grant from the Amsterdam Society and Research Fund for Prevention and Cure of Tuberculosis. This work was supported by the Commission of the European Communities Directorate General for Science Research and Development (project no. TS2-0111), by the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases, and by the Royal Netherlands Tuberculosis Association.

REFERENCES

1. **Abou-Zeid, C., I. Smith, J. M. Grange, T. L. Ratliff, J. Steele, and G. A. W. Rook.** 1988. The secreted antigens of *M. tuberculosis* and their relationship to those recognized by the available antibodies. J. Gen. Microbiol. **134:**531–538.
2. **Baird, P. N., L. M. C. Hall, and A. R. M. Coates.** 1988. A major antigen from *M. tuberculosis* which is homologous to the heat shock proteins GroES from *E. coli* and the htpA gene product of *C. burneti*. Nucleic Acids Res. **16:**9047.
3. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7:**1513–1524.
4. **Booth, R. J., D. P. Harris, J. M. Loe, and J. D. Watson.** 1988. Antigenic proteins of *M. leprae*: complete sequence of the gene for the 18-kDa protein. J. Immunol. **140:**597–601.
5. **Bossier, P., I. T. Fitch, H. Boucherie, and M. F. Tuite.** 1989. Structure and expression of a yeast gene encoding the small heat-shock protein Hsp26. Gene **78:**323–330.
6. **Calos, M. P., T. S. Lebkowkski, and M. R. Botchan.** 1983. High mutation frequency in DNA transfected into mammalian cells. Proc. Natl. Acad. Sci. USA **80:**3015–3019.
7. **Chandramuki, A., G. H. Bothamley, P. J. Brennan, and J. Ivanyi.** 1989. Levels of antibody to defined antigens of *Mycobacterium tuberculosis* in tuberculous meningitis. J. Clin. Microbiol. **25:**821–825.
8. **Coates, A. R. M., B. W. Allen, J. Hewitt, J. Ivanyi, and D. A. Mitchison.** 1981. Antigen diversity of *M. tuberculosis* and *M. bovis* detected by means of monoclonal antibodies. Lancet **ii:**167–169.
9. **Corpet, F.** 1988. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. **16:**10881–10890.
10. **Dale, J., and A. Patki.** 1990. Mycobacterial gene expression and regulation, p. 173–198. *In* J. McFadden (ed.), Molecular biology of the mycobacteria. Surrey University Press, London.
11. **De Jong, W. W., J. A. M. Leunissen, P. J. M. Leenen, A. Zweers, and M. Versteeg.** 1988. Dogfish alpha-crystallin sequences; comparison with small heat shock proteins and *Schistosoma* egg antigen. J. Biol. Chem. **263:**5141–5149.
12. **Dente, L., M. Sollazzo, C. Baldari, G. Cesareni, and R. Cortese.** 1985. The pEMBL family of single stranded vectors, p. 101–107. *In* D. M. Glover (ed.), DNA cloning, vol. 1. IRL Press, Oxford.
13. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
14. **Engers, H. D., M. Abe, B. R. Bloom, V. Mehra, W. Britton, T. M. Buchanan, S. R. Khanolkar, D. B. Young, O. Closs, T. Gilis, M. Harboe, J. Ivanyi, A. H. J. Kolk, and C. C. Shepard.** 1985. Results of a World Health Organization-sponsored workshop on monoclonal antibodies to *Mycobacterium leprae*. Infect. Immun. **48:**603–605.
15. **Engers, H. D., V. Houba, J. Bennedsen, T. M. Buchanan, S. D. Chaparas, G. Kadival, O. Closs, J. R. David, J. D. A. van Embden, T. Godal, S. A. Mustafa, J. Ivanyi, D. B. Young, S. H. E. Kaufmann, A. G. Khomenko, A. H. J. Kolk, M. Kubin, J. A. Louis, P. Minden, T. M. Shinnick, L. Trnka, and R. A. Young.** 1986. Results of a World Health Organization-sponsored workshop to characterize antigens recognized by mycobacterium-specific monoclonal antibodies. Infect. Immun. **51:**718–720.
16. **Falla, J. C., C. A. Parra, M. Mendoza, I. C. Franco, F. Guman, J. Forero, O. Orozco, and M. E. Patarroyo.** 1991. Identification of B- and T-cell epitopes within the MPT40 protein of *Mycobacterium tuberculosis* and their correlation with the disease course. Infect. Immun. **59:**2265–2273.
17. **Garsia, R. J., L. Hellqvist, R. J. Booth, A. J. Radford, W. J. Britton, L. Astbury, R. J. Trent, and A. Basten.** 1989. Homology of the 70-kilodalton antigens from *Mycobacterium leprae* and *Mycobacterium bovis* with the *Mycobacterium tuberculosis* 71-kilodalton antigen and with the conserved heat shock protein 70 of eucaryotes. Infect. Immun. **57:**204–212.
18. **Grimm, B., D. Ish-Shalom, D. Even, H. Glaczinski, P. Ottersbach, I. Ohad, and K. Kloppstech.** 1989. The nuclear-coded chloroplast 22-kDa heat-shock protein of *Chlamydomonas*. Evidence for translocation into the organelle without a processing step. Eur. J. Biochem. **182:**539–546.
19. **Hartskeerl, R. A., M. Y. L. de Wit, and P. R. Klatser.** 1989. Polymerase chain reaction for the detection of *M. leprae*. J. Gen. Microbiol. **135:**2357–2364.
20. **Hartskeerl, R. A., J. Tommassen, and W. Hoekstra.** 1985. Relationship between proteins encoded by the exclusion determining locus of the IncI plasmid R144 and the cellular localization of these proteins in *E. coli* K-12. Mol. Gen. Genet. **200:**138–144.
21. **Hartskeerl, R. A., R. M. van Rens, F. E. M. Stabel, M. Y. L. de Wit, and P. R. Klatser.** 1990. Selection and characterization of recombinant clones that produce *M. leprae* antigens recognized by antibodies in sera from household contacts from leprosy patients. Infect. Immun. **58:**2821–2827.
22. **Hickey, E., S. E. Brandon, R. Potter, G. Stein, J. Stein, and L. A. Weber.** 1986. Sequence and organization of genes encoding the human 27 kDa heat shock protein. Nucleic Acids Res. **14:**4127–4145.
23. **Hunter, S. W., B. Rivoiret, V. Mehra, B. R. Bloom, and P. J. Brennan.** 1990. The major native proteins of the leprosy bacillus. J. Biol. Chem. **265:**14065–14068.
24. **Ignolia, T. D., and E. A. Graig.** 1982. Four small *Drosophila* heat shock proteins are related to each other and to mammalian alpha-crystallin. Proc. Natl. Acad. Sci. USA **79:**2360–2364.
25. **Ivanyi, J., G. H. Bothamley, and P. S. Jackett.** 1988. Immuno-

diagnostic assays for tuberculosis and leprosy. Br. Med. Bull. **44:**635–649.
26. **Jackett, P. S., G. H. Bothamley, H. V. Batra, A. Mistry, D. B. Young, and J. Ivanyi.** 1988. Specificity of antibodies to immunodominant mycobacterial antigens in pulmonary tuberculosis. J. Clin. Microbiol. **26:**2313–2318.
27. **Kingston, A. E., P. R. Salgame, N. A. Mitchison, and M. J. Colston.** 1987. Immunological activity of a 14-kilodalton recombinant protein of *Mycobacterium tuberculosis* H37Rv. Infect. Immun. **55:**3149–3154.
28. **Kolk, A. H. J., R. Evers, D. G. Groothuis, H. Gilis, and S. Kuijper.** 1989. Production and characterization of monoclonal antibodies against specific serotypes of *Mycobacterium avium* and the *Mycobacterium avium-Mycobacterium intracellulare-Mycobacterium scrofulaceum* complex. Infect. Immun. **57:** 2514–2521.
29. **Kolk, A. H. J., M. L. Y. Ho, P. R. Klatser, T. A. Eggelte, S. Kuijper, S. de Jonge, and J. van Leeuwen.** 1984. Production and characterization of monoclonal antibodies to *M. tuberculosis*, *M. bovis* BCG and *M. leprae*. Clin. Exp. Immunol. **58:**511–521.
30. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **226:**680–685.
31. **Matthews, R., A. Scoging, and A. D. M. Rees.** 1985. Mycobacterial antigen-specific human T-cell clones secreting macrophage activating factors. Immunology **54:**17–23.
32. **McElwain, E. F., and S. Spiker.** 1989. A wheat cDNA clone which is homologous to the 17 kd heat-shock protein gene family of soybean. Nucleic Acids Res. **17:**1764.
33. **Mehra, V., D. Sweetser, and R. A. Young.** 1986. Efficient mapping of protein antigenic determinants. Proc. Natl. Acad. Sci. USA **83:**7013–7017.
34. **Murray, C. J., K. Styblo, and A. Rouillon.** 1990. Tuberculosis in developing countries: burden, intervention and cost. Bull. Int. Union Tuberc. Lung Dis. **65:**6–24.
35. **Nerland, A. H., A. Salim Mustafa, D. Sweetser, T. Godal, and R. A. Young.** 1988. A protein antigen of *M. leprae* is related to a family of small heat shock proteins. J. Bacteriol. **170:**5919–5921.
36. **Raschke, E., G. Baumann, and F. Schoeffl.** 1988. Nucleotide sequence analysis of soybean small heat shock protein genes belonging to two different multigene families. J. Mol. Biol. **199:**549–557.
37. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
38. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
39. **Schöningh, R., C. P. H. J. Verstijnen, S. Kuijper, and A. H. J. Kolk.** 1990. Enzyme immunoassay for identification of heat-killed mycobacteria belonging to the *Mycobacterium tuberculosis* and *Mycobacterium avium* complex and derived from early cultures. J. Clin. Microbiol. **28:**708–713.
40. **Shinnick, T. M.** 1987. The 65-kilodalton antigen of *Mycobacterium tuberculosis*. J. Bacteriol. **169:**1080–1088.
41. **Shinnick, T. M., C. Krat, and S. Schadow.** 1987. Isolation and restriction site maps of the genes encoding five *Mycobacterium tuberculosis* proteins. Infect. Immun. **55:**1718–1721.
42. **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.
43. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:**503–517.
44. **Stanley, K. K., and J. P. Luzio.** 1984. Construction of a new family of high efficiency bacterial expression vectors: identification of cDNA clones coding for human liver proteins. EMBO J. **6:**1429–1434.
45. **Susek, R. E., and S. L. Lindquist.** 1989. Hsp26 of *Saccharomyces cerevisiae* is related to the superfamily of small heat shock proteins but is without a demonstrable function. Mol. Cell. Biol. **9:**5265–5271.
46. **Takahashi, T., and Y. Komeda.** 1989. Characterization of two genes encoding small heat-shock proteins in *Arabidopsis thalania*. Mol. Gen. Genet. **219:**365–372.
47. **Van Eys, G. J. J. M., J. Zaal, G. J. Schoone, and W. J. Terpstra.** 1988. DNA hybridization with hardjobovis-specific recombinant probes as a method for type discrimination of *Leptospira interrogans* serovar hardjo. J. Gen. Microbiol. **134:**567–574.
48. **Verbon, A., S. Kuijper, H. M. Jansen, P. Speelman, and A. H. J. Kolk.** 1990. Antigens in culture supernatant of *M. tuberculosis*: epitopes defined by monoclonal and human antibodies. J. Gen. Microbiol. **136:**955–964.
49. **Verstijnen, C. P. H. J., R. Schöningh, S. Kuijper, J. Bruins, R. J. van Ketel, D. G. Groothuis, and A. H. J. Kolk.** 1989. Rapid identification of cultured *M. tuberculosis* with a panel of monoclonal antibodies in Western blot and immunofluorescence. Res. Microbiol. **140:**653–666.
50. **Welch, W. J.** 1990. The mammalian stress response: cell physiology and biochemistry of stress proteins, p. 223–278. *In* R. I. Morimoto, A. Tissieres, and C. Georgopoulos (ed.), Stress proteins in biology and medicine, monograph 19. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
51. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of M13mp18 and pUC19 vectors. Gene **33:**103–119.
52. **Young, D., T. Garbe, R. Lathigra, and C. Abou-Zeid.** 1990. Protein antigens: structure, function and regulation, p. 1–36. *In* J. McFadden (ed.), Molecular biology of the mycobacteria. Surrey University Press, London.
53. **Young, D., R. Lathigra, R. Hendrix, D. Sweetser, and R. A. Young.** 1988. Stress proteins are targets in leprosy and tuberculosis. Proc. Natl. Acad. Sci. USA **85:**4267–4270.
54. **Young, D. B., and T. Garbe.** 1991. Heat shock proteins and antigens of *Mycobacterium tuberculosis*. Infect. Immun. **59:** 3086–3093.
55. **Young, D. B., L. Kent, and R. A. Young.** 1987. Screening of a recombinant mycobacterial DNA library with polyclonal antiserum and molecular weight analysis of expressed antigens. Infect. Immun. **55:**1421–1425.
56. **Young, R. A., B. R. Bloom, C. N. M. Grosskinsky, J. Ivanyi, and R. W. Davis.** 1985. Dissecting *M. tuberculosis* antigens using recombinant DNA. Proc. Natl. Acad. Sci. USA **83:**2583–2587.
57. **Zabeau, M., and K. K. Stanley.** 1982. Enhanced expression of cro-beta-galactosidase fusion proteins under the control of the p_r promoter of bacteriophage lambda. EMBO J. **1:**1217–1224.