JOURNAL OF BACTERIOLOGY, Dec. 1990, p. 6774–6782
0021-9193/90/126774-09$02.00/0

Vol. 172, No. 12

Cloning and Characterization of the *aroA* Gene from *Mycobacterium tuberculosis*

THOMAS GARBE,[1]* CHRISTOPHER JONES,[1]† IAN CHARLES,[2] GORDON DOUGAN,[2] AND DOUGLAS YOUNG[1]

Medical Research Council Tuberculosis and Related Infections Unit, Royal Postgraduate Medical School, Hammersmith Hospital, Ducane Road, London W12 0HS,[1] and Department of Molecular Biology, Wellcome Biotech, Beckenham, Kent BR3 3BS,[2] United Kingdom

Received 11 April 1990/Accepted 28 August 1990

The *aroA* gene from *Mycobacterium tuberculosis* has been cloned by complementation of an *aroA* mutant of *Escherichia coli* after lysogenization with a recombinant DNA library in the λgt11 vector. Detailed characterization of the *M. tuberculosis aroA* gene by nucleotide sequencing and by immunochemical analysis of the expressed product indicates that it encodes a 5-enolpyruvylshikimate-3-phosphate synthase that is structurally related to analogous enzymes from other bacterial, fungal, and plant sources. The potential use of the cloned gene in construction of genetically defined mutant strains of *M. tuberculosis* by gene replacement is proposed as a novel approach to the rational attenuation of mycobacterial pathogens and the possible development of new antimycobacterial vaccines.

Mycobacterium tuberculosis and *Mycobacterium leprae*, the etiological agents of tuberculosis and leprosy, respectively, continue to be responsible for widespread human morbidity and mortality (19). In spite of the development of effective chemotherapy regimens, it has proved difficult to control these diseases in economically deprived areas of the world; even in developed countries that are largely free of traditional mycobacterial diseases, there is concern about the increased frequency of mycobacterial infection in patients suffering from acquired immune deficiency syndrome (3). In addition to chemotherapy, prophylactic vaccination against mycobacterial infection has been widely employed with an attenuated strain of mycobacteria closely related to *M. tuberculosis* (*Mycobacterium bovis*, bacillus Calmette-Guerin [BCG]) (11). Although BCG vaccination has provided protection against tuberculosis in certain areas, it has little or no effect against the adult form of the disease in other endemic communities. The mechanisms whereby BCG induces protective immunity in some trials and the reasons for its failure in other trials remain open to speculation (11).

A major area of mycobacterial research over the last 5 years has been directed toward the identification and cloning of individual antigens involved in interactions with the immune system, with the goal of progressing toward development of "subunit" vaccines that are effective against mycobacterial disease (9, 10, 19, 36–38). The recent development of techniques for the genetic manipulation of mycobacteria (15, 18, 31), on the other hand, permits an alternative approach to the design of potential antimycobacterial vaccines by construction of mutant strains of *M. tuberculosis* that are rationally attenuated for virulence. This approach has been applied in the case of *Salmonella* pathogens, which resemble mycobacteria in their ability to replicate within cells of the host reticuloendothelial system and in the requirement for live vaccines to elicit optimal protective responses in animal models of disease (6). For *Salmonella typhimurium*, it has been found that only some attenuated mutants are able to stimulate protective immune responses (28) and that strains carrying defects in the shikimate biosynthesis pathway (with growth requirements for aromatic amino acids, *para*-aminobenzoate, and 2,3-dihydroxybenzoate) are particularly suitable as vaccine candidates (14, 28). We propose to construct corresponding mutant strains of *M. tuberculosis* that are defective in aromatic amino acid biosynthesis to evaluate their virulence and immunogenicity. Gene replacement techniques have recently been successfully applied for construction of an auxotrophic mutant strain of *Mycobacterium smegmatis* (15), and this represents an attractive approach to the generation of defined *M. tuberculosis* mutants. As an initial step in the development of *M. tuberculosis aro* mutants by gene replacement, we report here the cloning of the mycobacterial *aroA* gene, encoding a key enzyme from the shikimate biosynthesis pathway.

MATERIALS AND METHODS

Bacterial strains. Table 1 outlines the properties of the bacterial strains, bacteriophages, and plasmids used in the study.

Lysogenization and complementation of *Escherichia coli aroA*. A λgt11 library prepared from random fragments of chromosomal *M. tuberculosis* DNA was obtained from R. A. Young (Whitehead Institute, Cambridge, Mass.) (37). An overnight culture of *E. coli* AB2829 (0.1 ml) grown in lambda medium (20) supplemented with 0.2% maltose and 10 mM $MgCl_2$ was infected with an *M. tuberculosis* λgt11 library at an input ratio of approximately 2×10^{-3} phage per bacterium. After adsorption for 30 min at 30°C, the culture was diluted 100-fold with lambda medium, incubated at 30°C until the optical density at 600 nm reached 0.2, centrifuged, and suspended in 1 ml of lambda medium. This suspension (0.3 ml per plate) was spread on Middlebrook 7H9 agar plates (Difco Laboratories) enriched with 2% maltose and 3 mM isopropyl-β-D-thiogalactopyranoside (Sigma Chemical Co.). Colonies that grew up after 3 to 6 days of incubation at 30°C were used to inoculate 1 ml of Middlebrook 7H9 medium containing maltose and isopropyl-β-D-thiogalactopyranoside. When the cultures appeared to have reached the stationary phase, they were used at a dilution of 1:100 to

* Corresponding author.
† Present address: Imperial Cancer Research Fund, Clare Hall Laboratories, South Mimms, Hertfordshire, EN6 3LD, United Kingdom.

TABLE 1. Characteristics of bacterial strains, bacteriophages, and plasmids

Strain	Characteristics	Source
Bacteria		
E. coli AB2829 (*aroA*)	K-12 derivative with single mutation in *aroA*	29
E. coli JM105	Transformation recipient for plasmids and M13 clones	Pharmacia
E. coli Y1090	Propagation and indicator strain for λgt11	16
M. tuberculosis H37Rv	Virulent strain of mycobacterium originally isolated from tuberculosis patient	B. Allen, Department of Bacteriology, Royal Postgraduate Medical School, London
Bacteriophages		
M. tuberculosis λgt11 library	Genomic DNA library	37
λgt11aroA8	Prophage complements *E. coli aroA* for growth on minimal medium with supplement	This work
λgt11aroA17 λgt11aroA20	Prophage complements *E. coli aroA* for growth on minimal medium	This work
Plasmids		
pBR322	Cloning vector	1
pGEX-2T	Expression vector	Pharmacia
pBParoA	*B. pertussis aroA* gene subcloned into pUC18	25
pBRaroA8	*Eco*RI insert from λgt11aroA8 in pBR322, does not complement *E. coli aroA*	This work
pBRaroA17-9	*Eco*RI insert from λgt11aroA17 in pBR322, complements *E. coli aroA*	This work
pBRaroA17-18	Insert in reverse orientation to pBRaroA17-9, does not complement *E. coli aroA*	This work
pGEXaroA10	*Apa*I restriction fragment from λgt11aroA17 fused in frame to glutathione *S*-transferase in pGEX-2T	This work

inoculate 2 ml of Luria-Bertani (LB) medium (7). After growth at 30°C to an optical density of 0.2 at 600 nm, cultures were thermoinduced at 43°C for 15 min and further incubated at 37°C for 2 to 3 h until cleared to obtain phage lysates.

Characterization of phage clones and insert DNA. Lysates obtained after induction of complemented lysogens were amplified in *E. coli* Y1090 and then purified by cesium chloride density gradient centrifugation (7). DNA was liberated from virions by the formamide method of Davis et al. (7), and DNA inserts were excised by *Eco*RI digestion and subcloned into multicopy plasmid vectors as indicated in Table 1 by standard procedures (24). *E. coli* JM105 was transformed with recombinant plasmid vectors by using the simple transformation method of Hanahan (13), and plasmid DNA was isolated from overnight cultures (24). Mycobacterial insert DNA was characterized from phage and plasmid preparations by restriction enzyme digestion and agarose gel electrophoresis (24).

Southern blot analysis. DNA fragments for radiolabeling were separated on 0.7% low-melting-point agarose gels (BRL) in a Tris-acetate-EDTA buffer system (24) and isolated from the gel by using Geneclean (BIO 101 Inc., La Jolla, Calif.). Radiolabeling was carried out with a random-primed labeling kit (Boehringer Mannheim GmbH) with 3 μCi of [α-^{32}P]dATP (Amersham plc.), and nonincorporated label was removed by passing through a Nick Column (Pharmacia). Southern blotting and hybridization were carried out as described by Maniatis et al. (24) with nylon membranes (Hybond-N; Amersham). For hybridization of the *Bordetella pertussis* probe to *M. tuberculosis* DNA, washing was carried out under reduced stringency as follows: one wash with 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.5% sodium dodecyl sulfate (SDS) at room temperature for 5 min; two washes with 2× SSC–0.1% SDS at room temperature for 15 min, and two washes with 2× SSC–0.1% SDS at 42°C for 1 h each. Autoradiographs were prepared by exposure with X-ray film (Kodak X-Omat AR) at −70°C overnight. Chromosomal DNA from *M. tuberculosis* was prepared as described by Thole et al. (34).

Sequence analysis. Nucleotide sequence analysis was performed by using the dideoxy-chain termination method with bacteriophage T7 polymerase (33). M13 clones were constructed and isolated as described by Messing et al. (27). Labeling reactions were carried out with a Pharmacia Deaza T7 sequencing kit with 10 μCi of [α-^{35}S]dATP per reaction and universal primer or specially synthesized oligonucleotides (see Fig. 2). Oligonucleotide primers were made on an SAM1 oligonucleotide synthesizer (Biolabs). Ordering of the fragments was accomplished by electrophoresis in 6% polyacrylamide gels (27) with an IBI model STS45 gel tank and subsequent autoradiography on Kodak X-Omat AR film at −70°C for 20 h.

Overexpression of the *aroA* gene product. The expression vector pGEX-2T (Pharmacia), restricted with *Sma*I and incubated with 1 U of alkaline phosphatase (Boehringer Mannheim) with subsequent inactivation according to the recommendations of the manufacturer, was ligated with the *Apa*I fragment from λgt11aroA17 (flushed with T4 DNA polymerase from Bethesda Research Laboratories as recommended by the manufacturer) and transformed into *E. coli* JM105. Single colonies were grown up overnight at 37°C in LB containing 50 μg of ampicillin per ml, diluted 1:100 in 2 ml of LB-ampicillin, and grown to an optical density of approximately 0.2 at 600 nm. Isopropyl-β-D-thiogalactopyranoside was added to a final concentration of 0.1 mM; after incubation for 2 to 3 h at 37°C, cultures were centrifuged and suspended in 30 μl of distilled water. Samples were analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) as described below, and positive clones were identified by the presence of an overproduced fusion protein.

The glutathione *S*-transferase fusion protein containing

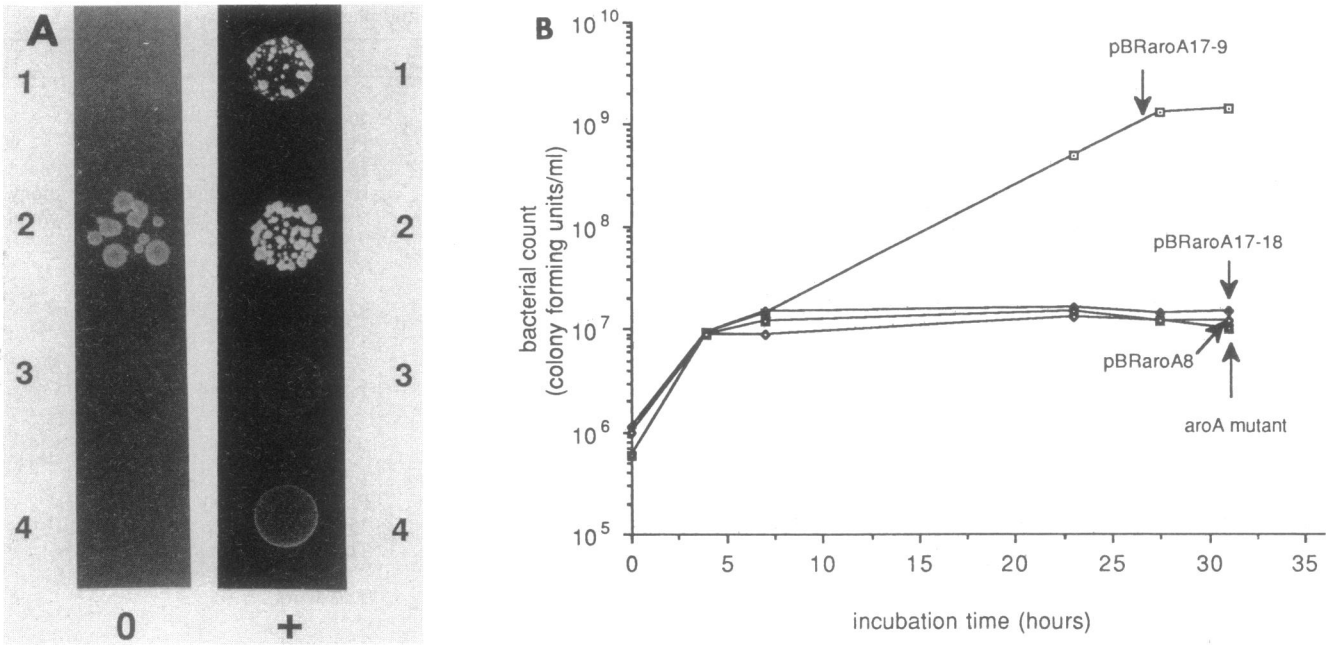

FIG. 1. Complementation of *E. coli aroA*. (A) Complementation by λgt11 prophage. An overnight culture of *E. coli aroA* was infected with phage, and samples containing 5×10^2 PFU were spotted onto agar plates prepared with M9 medium alone (0) or M9 medium supplemented with lambda medium (0.1 ml in 20 ml) (+). The extent of growth after 3 days of incubation at 30°C is shown for samples infected with λgt11aroA8 (plate 1), λgt11aroA17 (plate 2), and λgt11 containing insert DNA from an unrelated *M. tuberculosis* gene (plate 3) and for uninfected *E. coli aroA* (plate 4). Lysogenization with λgt11aroA17 allowed growth on minimal agar plates, whereas λgt11aroA8 complemented growth only on supplemented plates. (B) Complementation by plasmids. *E. coli aroA* was transformed with pBR322 constructs containing the *M. tuberculosis aroA* gene. Each culture was first grown overnight in LB medium (with or without 50 μg of ampicillin per ml as appropriate), collected by centrifugation, and suspended in the original volume of PBS. Samples were then inoculated at a dilution of 1:250 in M9 medium, and growth at 37°C was monitored over a period of 2 days by counting of bacterial colonies after plating out on LB agar. All of the cultures showed some growth over the first few hours, but only pBRaroA17-9 supported continued growth of the *aroA* mutant.

345 amino acids from the *M. tuberculosis aroA* gene product was isolated from 500 ml of an induced culture of *E. coli* JM105 transformed with pGEXaroA10 by taking advantage of the insolubility of the overproduced protein. After bacterial lysis by sonication, the pellet obtained by centrifugation at 16,000 × *g* for 20 min was washed with phosphate-buffered saline (pH 7.2) (PBS) and dissolved in 5 M guanidine hydrochloride by stirring for 1 h at room temperature. The soluble fraction was dialyzed extensively against PBS, and the precipitate was collected by centrifugation as before. Analysis by SDS-PAGE showed that the fusion protein was the major protein band present in the precipitate.

Preparation of antisera. Partially purified glutathione S-transferase–AroA fusion protein (0.2 mg) was mixed with incomplete Freund adjuvant (Difco Laboratories) and injected intramuscularly into two sites on the hind legs of a New Zealand White rabbit. Boosting injections were given subcutaneously at multiple sites after 4 weeks, and a serum sample was taken 2 weeks after the final injection. To remove antibodies to glutathione *S*-transferase and contaminating *E. coli* proteins, the serum was absorbed by three serial 1-h incubations with a crude extract of an *E. coli* strain that overexpressed an unrelated glutathione *S*-transferase fusion protein.

A rabbit antiserum raised against the purified *E. coli aroA* gene product was kindly provided by J. R. Coggins, Department of Biochemistry, University of Glasgow, Glasgow, United Kingdom.

SDS-PAGE and Western immunoblotting. *E. coli* samples for SDS-PAGE were prepared by mixing concentrated bacterial suspensions with an equal volume of sample buffer (22) containing SDS and β-mercaptoethanol and incubating for 2 min on a boiling water bath. *M. tuberculosis* H37Rv was grown with stirring at 37°C for 2 weeks in Middlebrook 7H9 medium containing ADC supplement (Difco Laboratories). Bacteria were harvested by centrifugation, washed twice with PBS, and suspended in distilled water at 1/50 of the original culture volume. An equal volume of glass beads (0.1-mm diameter) was added, and the suspension was subjected to vigorous vortex mixing for 10 min before boiling with sample buffer as above. Samples containing 10 to 50 μg of protein were separated by electrophoresis in SDS gels containing 12% (wt/vol) acrylamide (22) with a mini-gel system supplied by Hoefer Scientific.

After electrophoresis, proteins were transferred to nitrocellulose membranes as described by Towbin et al. (35). Blots were incubated by shaking at room temperature with 3% dried milk powder in PBS containing 0.2% (vol/vol) Triton X-100 (PBST-milk), then with rabbit antiserum diluted 1:2,000 in PBST-milk, and, after a wash with PBST, with affinity purified peroxidase-conjugated goat anti-rabbit immunoglobulin G (Bio-Rad Laboratories) in PBST-milk. After a wash with PBST, peroxidase activity was visualized with 3,3′-diaminobenzidine and H_2O_2 in 50 mM citrate buffer (pH 5.5).

RESULTS

Complementation of the *E. coli aroA* mutant. After lysogenization of *E. coli aroA* with the *M. tuberculosis* λgt11

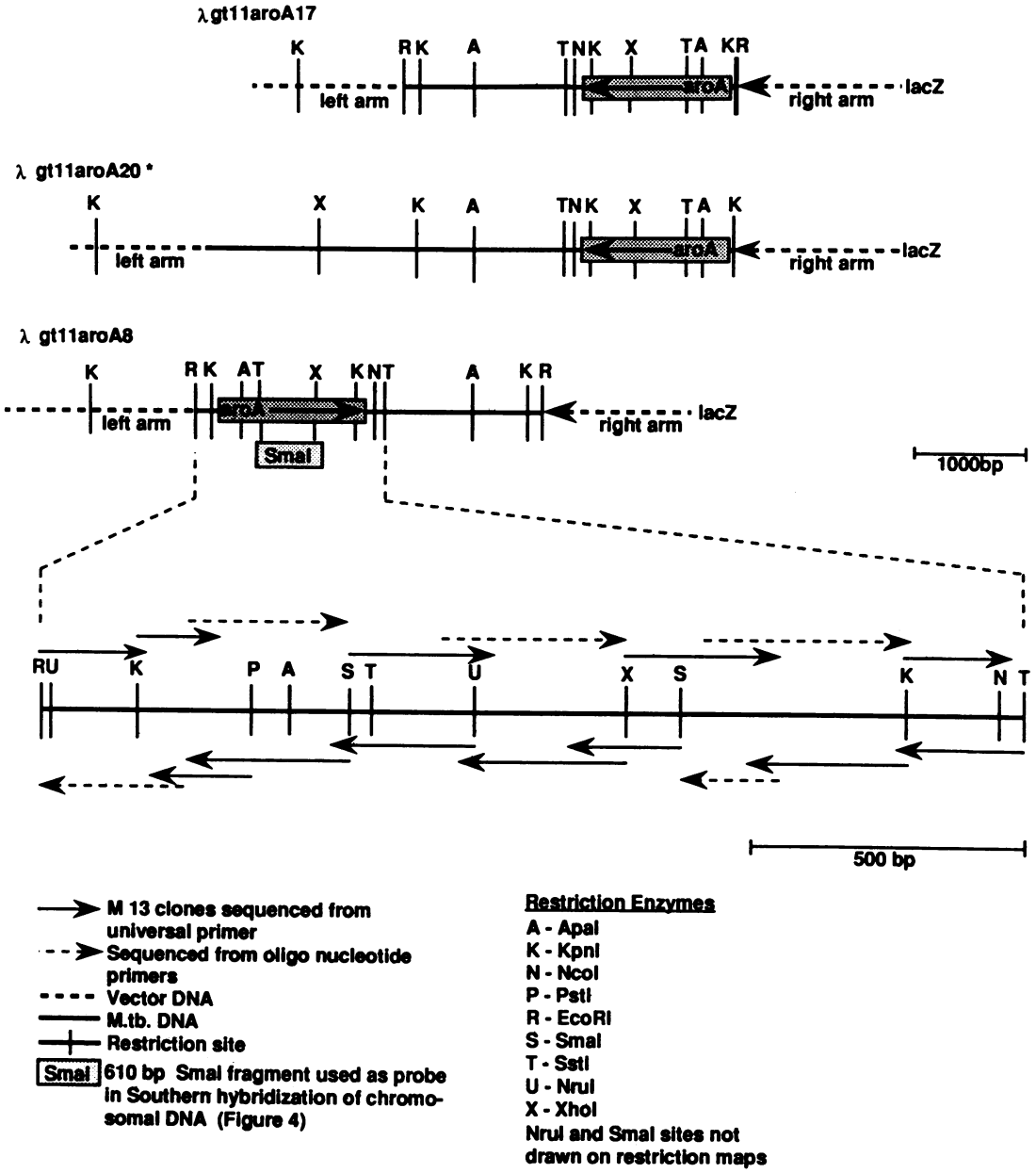

FIG. 2. Restriction maps of λgt11 clones encoding the *M. tuberculosis aroA* gene and the DNA sequencing strategy. The restriction maps of three λgt11 clones carrying *M. tuberculosis* DNA containing the coding sequence for the *aroA* gene (boxed) and its orientation with respect to *lacZ* promoter are shown. An expanded map of the *aroA* coding region outlining the sequencing strategy is shown below the restriction maps. Note that clone λgt11aroA20 (*) has lost both of its artificial *Eco*RI restriction sites.

library, slender colonies emerged on Middlebrook 7H9 plates incubated for 3 to 6 days, under conditions in which the original mutant strain showed no growth. Phage clones isolated from each of the lysogens were tested for their complementation efficiency by reinfection of the original mutant strain, and two classes of clones were identified (Fig. 1A). One set of clones (e.g., λgt11aroA17 and λgt11aroA20) showed strong complementation after lysogenization, resulting in growth on minimal agar plates prepared with Middlebrook 7H9 or M9 medium. A second set of clones (e.g., λgt11aroA8) provided only weak complementation, and it was necessary to supplement minimal agar plates with lambda medium (0.1 to 0.3 ml per plate) to allow growth of the lysogenized mutant.

The complementation efficiency of individual clones was further analyzed by subcloning of recombinant inserts in the multicopy plasmid pBR322. Recombinant plasmids were then tested for their ability to complement growth of *E. coli aroA* in M9 medium. Transformation with pBRaroA17-9 resulted in growth in M9 medium, but transformation with pBRaroA17-18 (with the insert in the reverse orientation) or pBRaroA8 failed to complement for growth (Fig. 1B).

Characterization of *aroA* clones. Phage clones λgt11aroA8, λgt11aroA17, and λgt11aroA20, isolated on the basis of their ability to complement the *E. coli aroA* mutant, were characterized with respect to their recombinant DNA inserts. Digestion of phage DNA with *Eco*RI resulted in release of a single recombinant fragment from λgt11aroA8 and

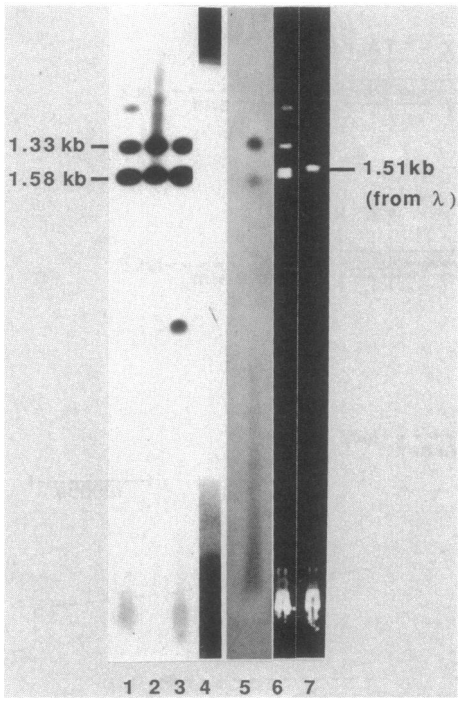

FIG. 3. Identification of the *aroA* gene by Southern blot analysis. DNA extracted from λgt11aroA8 (lanes 1, 5, and 6), λgt11aroA17 (lane 2), λgt11aroA20 (lane 3), and an unrelated λgt11 clone with deleted *Eco*RI sites (lanes 4 and 7) was digested with *Kpn*I and *Eco*RI, and restriction fragments were separated by agarose gel electrophoresis and transferred to a Hybond-N membrane. Hybridization was carried out as described in the text with radiolabeled probes prepared from the whole 3.2-kb *Eco*RI fragment derived from the pBRaroA8 subclone (lanes 1 through 4) and the *B. pertussis aroA* gene (lane 5). Two *Kpn*I fragments (1.33 and 1.58 kb) are common to all three recombinant clones, with the smaller fragment providing a strong signal on hybridization with the *B. pertussis* probe. The specificity of the hybridization is confirmed by the lack of binding to the 1.51-kb fragment derived from the λ vector which is seen with ethidium bromide staining (lane 7) but gives no hybridization signal (lane 4).

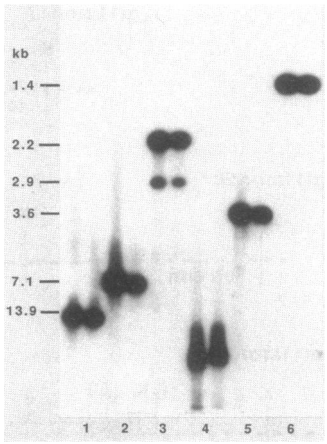

FIG. 4. Identification of the *aroA* gene on *M. tuberculosis* genomic DNA. Two independent preparations of genomic DNA from *M. tuberculosis* (2 and 1.5 μg) were restricted and loaded on a 0.7% agarose gel and electrophoresed overnight at 30 V. The gel was blotted and hybridized with the ^{32}P-labeled, pBRaroA8-derived 610-bp *Sma*I fragment from within the *M. tuberculosis aroA* gene. The two DNA preparations are shown in adjacent lanes after restriction with the following (lanes): 1, *Eco*RI; 2, *Sph*I; 3, *Xho*I; 4, *Hin*dIII; 5, *Nco*I; 6, *Not*I.

λgt11aroA17, whereas λgt11aroA20 was found to have lost its artificial *Eco*RI sites. Mapping with further restriction enzymes showed that all of the clones contained an overlapping region of *M. tuberculosis* DNA (Fig. 2), with two common *Kpn*I fragments that were readily identified after hybridization of insert DNA from each of the clones with a radiolabeled probe prepared from the 3.2-kb *Eco*RI fragment of the pBRaroA8 subclone (Fig. 3, lanes 1 through 3). The *Kpn*I fragments also hybridized to a probe prepared from a 1.4-kb *Pst*I-*Nco*I fragment carrying the *aroA* gene of *B. pertussis* (25); the smaller *Kpn*I fragment provided the stronger signal (Fig. 3, lane 5). The 1.35-kb *Kpn*I fragment was thus identified as the location of the major portion of the *M. tuberculosis aroA* gene. Southern blot analysis with genomic DNA from *M. tuberculosis* digested with a variety of restriction enzymes showed that a 610-bp *Sma*I fragment probe from this region hybridized to single fragments resulting from *Eco*RI, *Sph*I, *Nco*I, or *Not*I digestion (Fig. 4), suggesting that the mycobacterial chromosome contains only a single copy of the *aroA* gene. The signal with the *Hin*dIII-restricted DNA is blurred, probably because the *aroA* gene lies on a very large *Hin*dIII fragment, predominantly sheared by DNA manipulation. The two hybridizing *Xho*I fragments, one 2.2-kb fragment and a weaker 2.9-kb fragment, are due to the single *Xho*I site present on the *aroA* gene. The 610-bp *Sma*I fragment probe overlaps with the 2.9-kb *Xho*I fragment present in the λgt11aroA20 clone (Fig. 2) for only 101 bp and accordingly gave a weaker signal than that given by the 2.2-kb *Xho*I fragment, with which the probe shares the remaining 509 bp.

To obtain more detailed characterization of the *M. tuberculosis aroA* gene, the nucleotide sequence of the appropriate region of the DNA insert from λgt11aroA8 was determined (Fig. 5). Sequence analysis revealed an open reading frame encoding a 450-amino-acid polypeptide that shares extensive sequence identity with previously characterized bacterial, fungal, and plant 5-enolpyruvylshikimate-3-phosphate (EPSP) synthases (Fig. 6) (4, 8, 12). The *M. tuberculosis* sequence has 29% amino acid identity with *E. coli* EPSP synthase (8) and 27 and 27.5% identity with the corresponding enzymes from *Aspergillus nidulans* (4) and petunia (12), respectively. In common with other *M. tuberculosis* and *M. bovis* genes (26, 30), the *M. tuberculosis aroA* gene has a high G+C content (70%), and analysis of codon usage indicated a correspondingly strong preference for G and C in the third base position. A sequence matching an *E. coli* ribosome-binding site is present 11 nucleotides upstream from the proposed initiation codon, but no obvious promoter consensus sequence was identified.

Once the location and orientation of the *aroA* gene were determined, it was of interest to speculate on the differences in complementation efficiencies of the clones as described above. The pBR322 construct, which was functional in complementation experiments (pBRaroA17-9 in Fig. 1B), was found to have the *aroA* gene oriented correctly with respect to the P1 promoter of the plasmid (2). Complementation was not achieved by using pBRaroA8, however, even though in this construct the *aroA* gene was again oriented with the P1 promoter. Sequence analysis of the two clones showed that pBRaroA17 lacked 155 nucleotides upstream of

```
                                                                GAATTCCGCGGGACACC    17

ACTGTCGCGATCGTGTCGGTGGGTGCCACGTTGTAGCTCGTCTTCCCGCCACCGCACCCGGTGGCCTCGTCTATGGCCGT    97
                                                                ---> λgt11aroA17
GATTTTCTCGGCCAGCTGGGCCGGATCAGTGGTGACCGCAAACCGTCCGCACATGCTTCCTATGGTGCCTGGTACCCACG   177

                                                (Met)LysThrTrpProAlaProThrAlaProThrProValArgAlaThrVal
ACACCCACCGACACGGCAGGATGAAGCGGTGAAGACATGGCCAGCCCCAACGGCGCCGACGCCGGTGCGCGCTACCGTGA   257

ThrValProGlySerLysSerGlnThrAsnArgAlaLeuValLeuAlaAlaLeuAlaAlaAlaGlnGlyArgGlyAlaSer
CCGTTCCAGGCTCGAAGTCGCAGACCAACCGGGCGCTGGTGCTAGCGGCGCTGGCGGCCGCACAAGGCCGGGGCGCATCG   337

ThrIleSerGlyAlaLeuArgSerArgAspThrGluLeuMetLeuAspAlaLeuGlnThrLeuGlyLeuArgValAspGly
ACCATCTCCGGCGCGCTGCGCAGCCGCGACACCGAACTGATGCTGGACGCGCTGCAGACCCTGGGCCTGCGCGTCGACGG   417

ValGlySerGluLeuThrValSerGlyArgIleGluProGlyProGlyAlaArgValAspCysGlyLeuAlaGlyThr
TGTGGGTTCGGAACTGACGGTCAGCGGCCGAATCGAACCGGGGCCCGGCGCTCGGGTGGACTGTGGCTTGGCGGGCACGG   497

ValLeuArgPheValProProLeuAlaAlaLeuGlySerValProValThrPheAspGlyAspGlnGlnAlaArgGlyArg
TGTTGCGGTTTGTTCCGCCGCTGGCGGCGCTGGGCTCCGTCCCGGTCACCTTCGACGGCGATCAGCAAGCCCGGGGACGG   577

ProIleAlaProLeuLeuAspAlaLeuArgGluLeuGlyValAlaValAspGlyThrGlyLeuProPheArgValArgGly
CCCATCGCACCGCTGCTGGATGCGCTGCGCGAGCTCGGCGTCGCCGTCGACGGCACCGGTCTACCGTTTCGGGTTCGCGG   657

AsnGlySerLeuAlaGlyGlyThrValAlaIleAspAlaSerAlaSerSerGlnPheValSerGlyLeuLeuLeuSer
CAACGGGTCGCTCGCCGGCGGCACCGTGGCCATCGACGCGTCGGCGTCCTCACAGTTCGTGTCCGGGCTGCTGCTGTCCG   737

AlaAlaSerPheThrAspGlyLeuThrValGlnHisThrGlySerSerLeuProSerAlaProHisIleAlaMetThrAla
CGGCATCGTTCACCGATGGCCTGACCGTCCAACACACCGGTTCGTCGCTGCCGTCTGCGCCGCACATCGCGATGACGGCG   817

AlaMetLeuArgGlnAlaGlyValAspIleAspAspSerThrProAsnArgTrpGlnValArgProGlyProValAlaAla
GCGATGCTGCGGCAAGCCGGAGTCGACATCGACGACTCGACACCGAACCGTTGGCAGGTGCGCCCCGGTCCGGTGGCGGC   897

ArgArgTrpAspIleGluProAspLeuThrAsnAlaValAlaPheLeuSerAlaAlaValValSerGlyGlyThrVal
GCGGCGCTGGGACATCGAACCGGACCTGACCAACGCGGTGGCTTTCCTGTCAGCGGCCGTGGTCAGCGGCGGCACCGTGC   977

ArgIleThrGlyTrpProArgValSerValGlnProAlaAspHisIleLeuAlaIleLeuArgGlnLeuAsnAlaValVal
GCATCACCGGCTGGCCTAGAGTCAGCGTGCAACCCGCCGACCACATCTTGGCAATTTTGCGGCAGCTCAATGCCGTTGTC  1057

IleHisAlaAspSerSerLeuGluValArgGlyProThrGlyTyrAspGlyPheAspValAspLeuArgAlaValGlyGlu
ATTCATGCTGATTCATCCCTCGAGGTGCGCGGTCCAACGGGATACGACGGGTTTGACGTCGACTTGCGCGCCGTCGGCGA  1137

LeuThrProSerValAlaAlaLeuAlaAlaLeuAlaSerProGlySerValSerArgLeuSerGlyIleAlaHisLeu
GCTGACGCCATCGGTCGCGGCGCTGGCGGCGCTGGCATCCCCGGGATCGGTGTCCAGACTAAGCGGCATTGCCCATCTGC  1217

ArgGlyHisGluThrAspArgLeuAlaAlaLeuSerThrGluIleAsnArgLeuGlyGlyThrCysArgGluThrProAsp
GGGGCCACGAAACCGACCGGCTCGCCGCGCTGAGCACCGAGATCAACCGGTTGGGGGGCACCTGCCGGGAAACACCCGAC  1297

GlyLeuValIleThrAlaThrProLeuArgProGlyIleTrpArgAlaTyrAlaAspHisArgMetAlaMetAlaGlyAla
GGTCTGGTGATCACCGCGACGCCGTTGCGGCCCGGCATCTGGCGGGCATACGCGGACCATCGAATGGCGATGGCCGGCGC  1377

IleIleGlyLeuArgValAlaGlyValGluValAspAspIleAlaAlaThrThrLysThrLeuProGluProProArg
GATCATTGGGCTGCGGGTGGCCGGAGTCGAGGTCGACGACATCGCCGCCACCACCAAGACGCTGCCGGAGTTTCCGCGGC  1457

LeuTrpAlaGluMetValGlyProGlyGlnGlyTrpGlyTyrProGlnProArgSerGlyGlnArgAlaArgArgAlaThr
TGTGGGCCGAGATGGTCGGACCCGGCCAGGGGTGGGGGTACCCCCAGCCGCGCAGCGGCCAGCGGGCGAGGCGGGCAACC  1537

GlyGlnGlySerGlyGly***
GGGCAGGGGTCCGGCGGTTGAGGCCCGGCGACTACGACGAGTCCGACGTCAAGGTGCGCTCCGGCAGGAGTTCGCGGCCG  1617


CGGACCAAGACCCGTCCCGAGCACGCCGACGCCGAGGCCGCCATGG   1663
```

FIG. 5. Nucleotide sequence and translation of the *M. tuberculosis aroA* gene. The nucleotide sequence of a portion of the DNA insert from λgt11aroA8 was determined by the strategy outlined in the legend to Fig. 2. Translation of the open reading frame encoding the 450-amino-acid *aroA* gene product is shown with a potential ribosome-binding site underlined. Arrows indicate the position of a possible stem-and-loop structure between nucleotide residues 65 and 82. The start site of the DNA insert from λgt11aroA17 (residue 156) was determined by nucleotide sequence analysis and is also shown.

the open reading frame that were present in pBRaroA8 (Fig. 5). Within this region, residues 65 through 82 have the potential to form a loop structure with a stem containing an exact match of 7 bp. Formation of such a stem-and-loop structure could result in transcriptional termination, and its presence provides a possible explanation for the failure of pBRaroA8 in complementation experiments. Interestingly, in the original phage clone (λgt11aroA8), the weak comple-

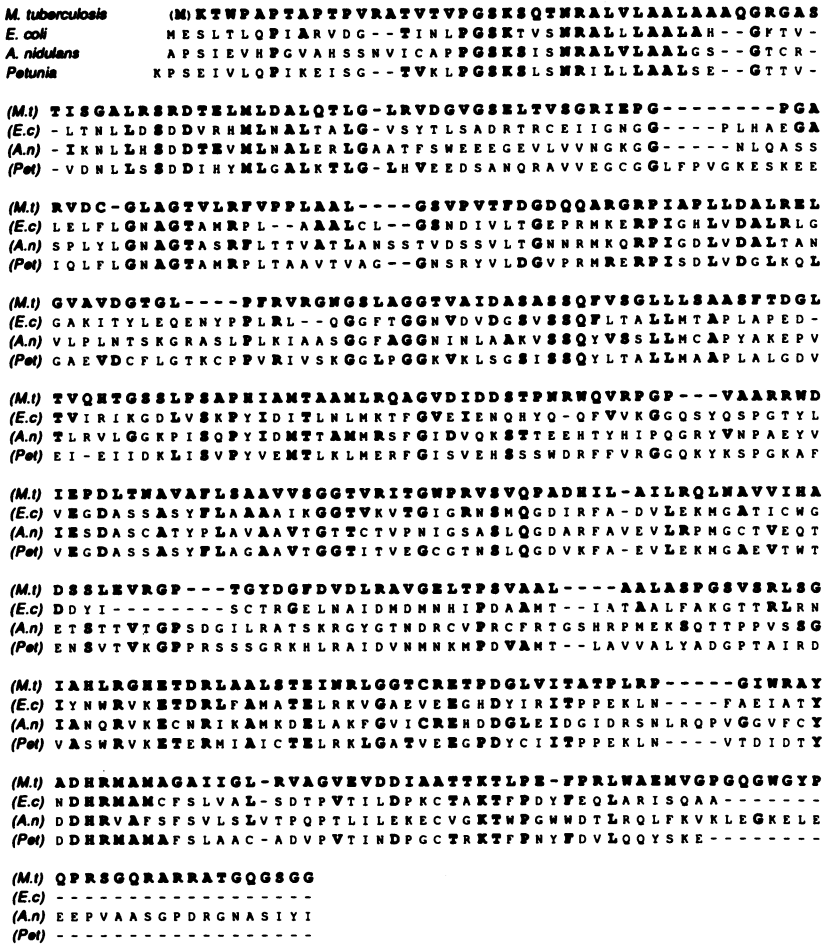

FIG. 6. Comparative sequence analysis of EPSP synthases. The amino acid sequence derived from analysis of the *M. tuberculosis aroA* gene (*M.t*) is shown in single-letter code along with the sequences derived for related enzymes from *E. coli* (8) (*E.c*), *Aspergillus nidulans* (4) (*A.n*), and petunia (12) (*Pet*). Residues that are identical to those of the *M. tuberculosis* sequence are shown in boldface type. Gaps that have been introduced to maximize sequence alignment are indicated by dashes.

mentation seen in Fig. 1A was achieved with the *aroA* gene present in the reverse orientation with respect to the *lacZ* promoter (Fig. 2).

Immunochemical analysis of *M. tuberculosis* EPSP synthase. A portion of the *aroA* open reading frame (corresponding to amino acids 105 through 450) was overexpressed as a glutathione *S*-transferase fusion protein, and the partially purified product was used to elicit a hyperimmune rabbit antiserum. Western blot analysis (Fig. 7, lane 1) showed that the absorbed rabbit antiserum recognized predominantly a single polypeptide with an apparent molecular mass of 53 kDa in *M. tuberculosis* extracts, indicating that the *aroA* gene is, in fact, expressed in *M. tuberculosis*. In common with the corresponding *E. coli* protein (23), the molecular mass of the EPSP synthase from *M. tuberculosis* found by SDS-PAGE (53 kDa) is significantly higher than that determined by analysis of the derived amino acid sequence from the cloned gene (46 kDa). The rabbit antiserum showed a prominent cross-reaction with a 50-kDa band in extracts of *E. coli* JM105 grown in M9 medium supplemented with 1 μg of thiamine per ml (Fig. 7, lane 3); the 50-kDa band comigrated with the protein recognized by an antiserum raised to the purified *E. coli* EPSP synthase (lane 3). Antiserum raised to the *E. coli* protein recognized the overexpressed fusion protein produced from the *M. tuberculosis* gene, although it

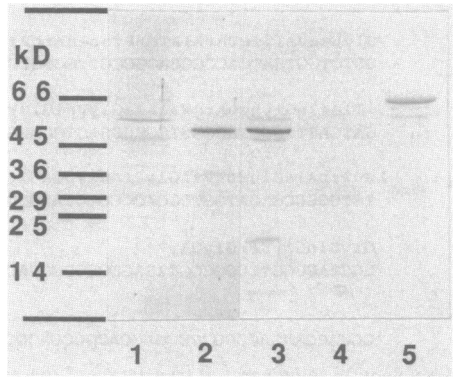

FIG. 7. Expression and antigenic cross-reactivity of the *M. tuberculosis aroA* gene product. An antiserum raised against the purified *E. coli aroA* gene product that recognized a single 50-kDa band in *E. coli* extracts (lane 3) cross-reacted in Western blot assays with a purified fusion protein constructed from amino acids 105 to 450 of the *M. tuberculosis aroA* gene (lane 5) but gave no detectable signal when tested with total cell extracts from *M. tuberculosis* (lane 4). Antiserum raised to the fusion protein reacted with a protein of the same molecular mass in *E. coli* extracts (lane 2) and also recognized a 53-kDa band in *M. tuberculosis* extracts (lane 1).

did not give a detectable signal with *M. tuberculosis* extracts in Western blot experiments (Fig. 7, lanes 4 and 5). The anti-*E. coli* EPSP synthase antiserum did not bind to free glutathione *S*-transferase or to unrelated fusion proteins.

DISCUSSION

Rapid progress has been achieved over the last few years in the cloning of genes encoding protein antigens from pathogenic mycobacteria in *E. coli* by using screening procedures based on detection by antibodies and oligonucleotide probes (26, 30, 37, 38). Although several of the antigens have subsequently been shown to fulfill important metabolic functions (36), direct attempts to clone mycobacterial genes coding for proteins with known enzyme activities have met with only limited success (5, 15, 17). It has been concluded that regulatory sequences controlling transcription and translation of mycobacterial genes differ from those in *E. coli* (21) and that the functional expression of mycobacterial enzymes in *E. coli* is therefore dependent on the presence of a strong exogenous promoter in plasmid vectors that allows multiple copies of the gene to be present (17). In the case of *aroA*, however, we have found that a single gene copy present in the form of a lambda prophage was sufficient for complementation, although correct orientation of the gene with respect to an exogenous promoter is clearly advantageous. We have used the same approach to clone the *aroB* and *aroD* genes of *M. tuberculosis* and have found that the λgt11 library is more consistent in complementation experiments than libraries constructed in multicopy pUC vectors, which show frequent rearrangements of mycobacterial DNA inserts. λgt11 libraries containing DNA from pathogenic mycobacteria are widely available (37, 38) and may prove useful in further complementation experiments for cloning of mycobacterial genes from key metabolic pathways.

EPSP synthase (the *aroA* gene product) is the site of action of the herbicide glyphosate (32), and the enzyme has been studied in detail in bacterial, fungal, and plant cells (4, 8, 12, 23). It has been pointed out that the plant enzyme is more closely related to the *E. coli* enzyme (54% amino acid identity) than to the *Aspergillus* enzyme (38% identity); this phenomenon has been discussed with regard to the possible evolution of plastids (where the plant enzyme is located) from endosymbiotic procaryotes (12). The special relationship between the *E. coli* and plant sequences is not shared, however, by the *M. tuberculosis* enzyme, which appears to be equidistant from the plant, fungal, and bacterial proteins as judged by sequence conservation. If plastids evolved from an intracellular procaryote, therefore, it must be concluded that the original bacterium was more closely related to *E. coli* than to *M. tuberculosis*.

In conjunction with the development of techniques for the genetic manipulation of mycobacteria, cloning of genes for key metabolic enzymes will allow initiation of a genetic approach to the understanding of mycobacterial physiology. In addition, enzymes, such as EPSP synthase, that are essential for bacterial survival in vivo and that have no counterpart in the mammalian host cell may represent suitable targets for the development of novel antimicrobial agents. The availability of cloned genes and the access to structural information through nucleotide sequence analysis could play an important role in this regard. Of particular interest in our own research is the potential use of cloned genes in the generation of genetically defined mutant strains of *M. tuberculosis* by gene replacement, and experiments to construct mycobacterial strains carrying deletions in the *aroA* gene are currently in progress.

ACKNOWLEDGMENTS

We are grateful to Juraj Ivanyi for discussion and encouragement, to Mark Gittens and Andrew Foey for assistance in protein purification and preparation of antisera, and to Hugh Spence for making oligonucleotides.

This investigation received financial support from the United Nations Development Program/World Bank/World Health Organization Special Programme for Research and Training in Tropical Diseases.

LITERATURE CITED

1. **Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heynecker, H. W. Boyer, J. H. Grosa, and S. Falkow.** 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2**:95–113.
2. **Brosius, J., R. L. Cate, and A. P. Perlmutter.** 1982. Precise location of two promoters for the β-lactamase gene of pBR322. J. Biol. Chem. **257**:9205–9210.
3. **Chaisson, R. E., and G. Slutkin.** 1989. Tuberculosis and human immunodeficiency virus infection. J. Infect. Dis. **159**:96–100.
4. **Charles, I., J. W. Keyte, W. J. Brammar, M. Smith, and A. R. Hawkins.** 1986. The isolation and nucleotide sequence of the complex AROM locus of *Aspergillus nidulans*. Nucleic Acids Res. **14**:2201–2213.
5. **Clark-Curtiss, J. E., W. R. Jacobs, M. A. Docherty, L. R. Ritchie, and R. Curtiss.** 1985. Molecular analysis of DNA and construction of genomic libraries of *Mycobacterium leprae*. J. Bacteriol. **161**:1093–1102.
6. **Collins, F. M.** 1974. Vaccines and cell-mediated immunity. Bacteriol. Rev. **38**:371–402.
7. **Davis, R. W., D. Botstein, and J. R. Roth.** 1980. Advanced bacterial genetics. A manual for genetic engineering. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
8. **Duncan, K., A. Lewendon, and J. R. Coggins.** 1984. The complete amino acid sequence of *Escherichia coli* 5-enolpyruvylshikimate-3-phosphate synthase. FEBS Lett. **170**:59–63.
9. **Engers, H. D., and Workshop Participants.** 1985. Results of a World Health Organization sponsored workshop on monoclonal antibodies to *Mycobacterium leprae*. Infect. Immun. **48**:603–605.
10. **Engers, H. D., and Workshop Participants.** 1986. Results of a World Health Organization sponsored workshop to characterize antigens recognized by mycobacteria-specific monoclonal antibodies. Infect. Immun. **51**:718–720.
11. **Fine, P. E. M.** 1988. BCG vaccination against tuberculosis and leprosy. Br. Med. Bull. **44**:691–703.
12. **Gasser, C. S., J. A. Winter, C. M. Hironaka, and D. M. Shah.** 1988. Structure, expression, and evolution of the 5-enolpyruvylshikimate-3-phosphate synthase genes of petunia and tomato. J. Biol. Chem. **263**:4280–4289.
13. **Hanahan, D.** 1985. Techniques for transformation of *E. coli*, p. 109–135. *In* D. M. Glover (ed.), DNA cloning, vol. 1. IRL Press, Inc., Washington, D.C.
14. **Hoiseth, S. K., and B. A. D. Stocker.** 1981. Aromatic-dependent *Salmonella typhimurium* are non-virulent and are effective live vaccines. Nature (London) **241**:238–239.
15. **Husson, R. N., B. E. James, and R. A. Young.** 1990. Gene replacement and expression of foreign DNA in mycobacteria. J. Bacteriol. **172**:519–524.
16. **Huynh, T. V., R. A. Young, and R. W. Davis.** 1985. Construction and screening of cDNA libraries in λgt10 and λgt11, p. 49–78. *In* D. M. Glover (ed.), DNA cloning, vol. 1. IRL Press, Inc., Washington, D.C.
17. **Jacobs, W. R., M. A. Docherty, R. Curtiss, and J. E. Clark-Curtiss.** 1986. Expression of *Mycobacterium leprae* genes from a *Streptococcus mutans* promoter in *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA **83**:1926–1930.
18. **Jacobs, W. R., R. Tuckman, and B. R. Bloom.** 1987. Introduction of foreign DNA into mycobacteria using a shuttle phasmid. Nature (London) **327**:532–535.
19. **Joint International Union against Tuberculosis and World Health Organisation Study Group.** 1982. Tuberculosis control. Tubercle **63**:157–169.

20. **Karn, J., S. Brenner, and L. Barnett.** 1983. New bacteriophage lambda vectors with positive selection for cloned inserts. Methods Enzymol. **101**:3–19.
21. **Kieser, T., M. T. Moss, J. W. Dale, and D. A. Hopwood.** 1986. Cloning and expression of *Mycobacterium bovis* BCG DNA in *Streptomyces lividans*. J. Bacteriol. **168**:72–80.
22. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
23. **Lewendon, A., and J. R. Coggins.** 1987. 3-Phosphoshikimate 1-carboxyvinyltransferase from *Escherichia coli*. Methods Enzymol. **142**:342–348.
24. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Maskell, D. J., P. Morrissey, and G. Dougan.** 1988. Cloning and nucleotide sequence of the *aroA* gene of *Bordetella pertussis*. J. Bacteriol. **170**:2467–2471.
26. **Matsuo, K., R. Yamaguchi, A. Yamazaki, H. Tasaka, and T. Yamada.** 1988. Cloning and expression of the *Mycobacterium bovis* BCG gene for extracellular α-antigen. J. Bacteriol. **170**: 3847–3854.
27. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–89.
28. **O'Callaghan, D., D. Maskell, F. Y. Liew, C. S. F. Easmon, and G. Dougan.** 1988. Characterization of aromatic- and purine-dependent *Salmonella typhimurium*: attenuation, persistence, and ability to induce protective immunity in BALB/c mice. Infect. Immun. **56**:419–423.
29. **Pittard, J., and B. J. Wallace.** 1966. Distribution and function of genes concerned with aromatic biosynthesis in *Escherichia coli*. J. Bacteriol. **91**:1494–1508.
30. **Shinnick, T. M.** 1987. The 65-kilodalton antigen of *Mycobacterium tuberculosis*. J. Bacteriol. **169**:1080–1088.
31. **Snapper, S. B., L. Lugosi, A. Jekkel, R. E. Melton, T. Kieser, B. R. Bloom, and W. R. Jacobs.** 1988. Lysogeny and transformation in mycobacteria: stable expression of foreign genes. Proc. Natl. Acad. Sci. USA **85**:6987–6991.
32. **Steinrucken, H. C., and N. Amrhein.** 1980. The herbicide glyphosate is a potent inhibitor of 5-enol pyruvyl shikimic acid-3-phosphate synthase. Biochem. Biophys. Res. Commun. **94**:1207–1212.
33. **Tabor, S., and C. C. Richardson.** 1987. DNA sequencing analysis with a modified bacteriophage T7 polymerase. Proc. Natl. Acad. Sci. USA **84**:4767–4771.
34. **Thole, J. E. R., H. G. Dauwerse, P. K. Das, D. G. Groothuis, L. M. Schouls, and J. D. A. van Embden.** 1985. Cloning of *Mycobacterium bovis* BCG DNA and expression of antigens in *Escherichia coli*. Infect. Immun. **50**:800–806.
35. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.
36. **Young, D. B.** 1988. Structure of mycobacterial antigens. Br. Med. Bull. **44**:562–583.
37. **Young, R. A., B. R. Bloom, C. M. Grosskinsky, J. Ivanyi, D. Thomas, and R. W. Davis.** 1985. Dissecting *Mycobacterium tuberculosis* antigens using recombinant DNA. Proc. Natl. Acad. Sci. USA **82**:2583–2587.
38. **Young, R. A., V. Mehra, D. Sweetser, T. 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.