JOURNAL OF BACTERIOLOGY, Sept. 1988, p. 3847–3854
0021-9193/88/093847-08$02.00/0

Vol. 170, No. 9

Cloning and Expression of the *Mycobacterium bovis* BCG Gene for Extracellular α Antigen

KAZUHIRO MATSUO,[1]* RYUJI YAMAGUCHI,[1] AKIHIRO YAMAZAKI,[1] HIROMICHI TASAKA,[2] AND TAKESHI YAMADA[3]

Central Research Laboratories of Ajinomoto Co. Inc., Suzuki-cho 1-1, Kawasaki-ku, Kawasaki City 210,[1] Hiroshima University School of Medicine, Kasumi 1-2-3, Minami-ku, Hiroshima City 734,[2] and Research Institute for Microbial Diseases, Osaka University, Yamada-oka 3-1, Suita City, Osaka 565,[3] Japan

Received 29 February 1988/Accepted 27 May 1988

The gene for the extracellular α antigen of *Mycobacterium bovis* BCG was cloned by using a single probe restricted to G or C in the third position. This technique should have great potential for the isolation of mycobacterial antigen genes. The gene analysis revealed that the α antigen gene encoded 323 amino acid residues, including 40 amino acids for signal peptide followed by 283 amino acids for mature protein. This is the first report on the structure of the mycobacterial signal peptide. The promoter-like sequence and ribosome-binding site were observed upstream of the open reading frame. In the coding region, the third position of the codon showed high G+C content (86%). The gene was expressed as an unfused protein in *Escherichia coli* by using an *E. coli* expression vector. This protein, which reacted with polyclonal antibody raised against α antigen from *Mycobacterium tuberculosis*, would be applicable to the immunodiagnosis of tuberculosis.

Tuberculosis and leprosy, caused by *Mycobacterium tuberculosis* and *Mycobacterium leprae*, respectively, are the two major mycobacterial diseases. It is well known that the purified protein derivatives derived from heat-killed cultures of *M. tuberculosis* elicit delayed-type hypersensitivity (DTH), while its lipid-containing cell wall contributes to the enhancement of immunization. The purification and characterization of individual antigenic proteins are essential in understanding the fundamental mechanism of the DTH reaction on the molecular level and also expected to contribute to the possible protection of not only mycobacteria-caused diseases but also other infectious diseases in which cellular immunity is involved.

So far, numerous groups have reported that common or species-specific antigens are found in mycobacteria by the use of physicochemical and monoclonal or polyclonal antibody-based techniques. However, there have been few antigens with a well-defined function. In the last decade, the development of recombinant DNA techniques have enabled us to clone and analyze genes for these antigens. This genetic information should not only provide more insight into the molecular process involved in the DTH reaction but also be extremely useful in clinical applications. For instance, important uses may be found for the immunodiagnosis of tuberculosis by genetically engineered antigens and oligonucleotide probes based on DNA sequence. Moreover, studies on antigenic regions might lead to the structure of epitope that would have additional diagnostic potential.

Recently, considerable attention has been directed to gene cloning for antigens from *M. tuberculosis* (5, 12, 40), *Mycobacterium bovis* BCG (16), and *M. leprae* (14). Detailed research has been done on the cloning, sequencing, and characterization of the gene for the 65-kilodalton (kDa) antigen found in *M. leprae* (4, 19), *M. tuberculosis* (12, 23, 39), and *M. bovis* BCG (24, 31, 32).

In this report, we describe α antigen, which is one of the major proteins secreted from *M. bovis* BCG. The protein, first reported in 1965 (37, 38), is of particular interest since it shows a DTH reaction and is a cross-reacting material which is widely distributed among the slow-growing mycobacteria, namely, *M. tuberculosis*, *M. leprae*, *M. bovis*, and nontuberculous mycobacteria. The gene for the α antigen was then cloned by using synthetic oligonucleotide probes based on the amino acid sequence of the N-terminus and, after sequencing, was expressed in *Escherichia coli*.

MATERIALS AND METHODS

Chemicals. All restriction endonucleases and modifying enzymes were purchased from Takara Shuzo Co., Ltd., Kyoto, Japan. [γ-^{32}P]ATP and [α-^{32}P]dCTP were from Amersham Japan Co., Ltd., Tokyo, Japan.

Bacterial strains and culture. *M. bovis* BCG (Tokyo) was grown at 37°C in Sauton medium (26). The culture medium for *E. coli* JM109 (20, 36) was L-broth (21).

^{32}P-labeled oligonucleotide probes. Antigen α was purified from *M. tuberculosis* as described previously (28). About 30 μg of purified α antigen was analyzed by a 470A gas-phase protein sequencer (Applied Biosystems Inc.), and the sequence of the N-terminal 30 amino acids was determined (Fig. 1). Oligonucleotide probes homologous to the N-terminal amino acid sequence were chemically synthesized by a 370A automated DNA synthesizer (Applied Biosystems Inc.) and purified by two cycles of reverse-phase liquid chromatography. The sequences for probes I and II were 5′-TTCTCGCGCCCGGGCCTGCCGGTCGAGTACCTGC AGGTCCCGTCGCCGTCGATGGG-3′ (corresponding to Phe-1 through Gly-19) and 5′-GGCCGCGACATCAA GGTGCAGTTCCAGTCGGG-3′ (corresponding to Gly-19 through Gly-29), respectively. About 100 pmol of purified oligonucleotide probes were labeled with ^{32}P by incubation at 37°C for 2 h in 30 μl of kination mixture containing 50 mM Tris hydrochloride (Tris-HCl) (pH 7.5), 10 mM $MgCl_2$, 10 mM dithiothreitol, 0.66 μM ATP, 100 μCi of [γ-^{32}P]ATP (3,000 Ci/mmol), and 15 U of T4 polynucleotide kinase. After incubation, the reaction mixtures were extracted once with phenol and labeled oligonucleotide probes were precipitated

* Corresponding author.

```
                5                  10
Phe Ser Arg Pro Gly Leu Pro Val Glu Tyr

               15                  20
Leu Gln Val Pro Ser Pro Ser Met Gly Arg

               25                  30
Asp Ile Lys Val Gln Phe Gln Ser Gly Gly
```

FIG. 1. N-terminal amino acid sequence of α antigen. The experimental procedures are described in Materials and Methods.

with 50 μg of carrier calf thymus DNA by ethanol precipitation. Precipitated probes were used for the hybridization test without any further purification steps.

Genomic Southern hybridization. *M. bovis* BCG total cellular DNA was extracted and purified as described previously (26). About 3 μg of purified *M. bovis* BCG DNA was completely digested with *Bam*HI, *Kpn*I, and *Pst*I and fractionated by 0.8% agarose gel electrophoresis (18). Fractionated DNA fragments were denatured by 0.5 N NaOH and 1.5 N NaCl and transferred to GeneScreen Plus nylon membrane filters (NEN Research Products, Boston) by the method of Southern (25). The filters were dried at 37°C for 16 h and wetted with hybridization solution (5× SSC [1× SSC is 0.15 M NaCl, 0.015 M sodium citrate], 5× Denhardt solution [18], 0.1% sodium dodecyl sulfate [SDS]) containing about 10^6 cpm of ^{32}P-labeled oligonucleotide probes (27). After hybridization at 55°C for 16 h, the filters were washed four times with 2× SSC–0.1% SDS for 15 min at 50°C and, for probe I, three times with 0.1× SSC–0.1% SDS for 15 min at 48°C. They were dried at room temperature and autoradiographed.

DNA cloning. *M. bovis* BCG DNA was digested with *Kpn*I and fractionated by 0.8% agarose gel electrophoresis. DNA fragments with a length of 4.5 to 6.0 kilobase pairs (kbp) were trapped on DE81 paper, washed with 0.6 ml of 50 mM Tris-HCl (pH 7.5), and eluted with 0.3 ml of 50 mM Tris-HCl (pH 7.5) and 2 M NaCl. Eluted DNA fragments were extracted twice with phenol and then chloroform and precipitated with ethanol. Plasmid vector pUC18 (20) was digested with *Kpn*I and dephosphorylated with bacterial alkaline phosphatase. Ligation and transformation were carried out with the use of a DNA ligation kit and *E. coli* JM109 competent cells purchased from Takara Shuzo Co., Ltd. The ampicillin-resistant (Amp^r) and white colonies on the L-broth plates containing 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and 0.004% 5-bromo-4-chloro-3-indolyl-β-D-galactoside were screened by the colony hybridization technique for the presence of DNA sequences which hybridized to ^{32}P-labeled probes (8).

DNA sequencing. The sequencing strategy is represented in Fig. 2. *Pst*I fragments from the mycobacterial DNA insert were subcloned into the *Pst*I site of pUC18. In the same way, *Sau*3AI fragments were subcloned into the *Bam*HI site, *Taq*I fragments into the *Acc*I site, *Pst*I-*Xho*I fragments into the *Pst*I-*Sal*I site, and *Kpn*I-*Sca*I fragments into the *Kpn*I-*Hinc*II site of pUC18. The nucleotide sequences of these subcloned DNA fragments were determined by the dideoxy chain termination method with intact pUC plasmids (9).

Construction of α antigen expression vector. A *Pst*I fragment (0.8 kbp) subcloned to pUC18 was isolated and recloned into the *Nco*I-*Pst*I site of pKK233-2 (1) by using a synthetic adaptor containing sequences for Phe-1 to Gln-12 since the *Pst*I fragment lacked a part of the α antigen gene for these N-terminal amino acids. The sequences for the adaptor were 5′-CATGTTCTCCCGGCCGGGGCTGCCG

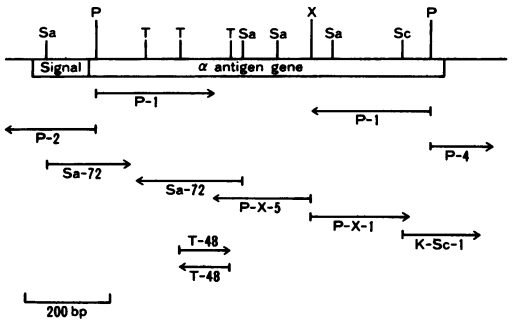

FIG. 2. Sequencing strategy. Restriction sites: Sa, *Sau*3AI; Sc, *Sca*I; T, *Taq*I; P, *Pst*I; X, *Xho*I. Arrows indicate direction and length of sequence determined in this experiment.

GTCGAGTACCTGCA-3′ (for the upper strand) and 5′-GGTACTCGACCGGCAGCCCCGGCCGGGAGAA-3′ (for the lower strand). The plasmid which had the inserted *Pst*I fragment in the correct orientation relative to the *trc* promoter (1) was named pKKα-1. This plasmid still lacks sequences for Ser-278 to Gly-283 just before the stop codon. Therefore, a *Xho*I-*Pvu*I fragment (372 bp) isolated from the cloned mycobacterial DNA was inserted into the *Xho*I-*Hin*dIII site of pKKα-1 by using a synthetic adaptor connected between the *Pvu*I and the *Hin*dIII sites as shown in Fig. 7. The sequences for the adaptor were 5′-CGGTACC TAA-3′ (for the upper strand) and 5′-AGCTTTAGGTACC GAT-3′ (for the lower strand). The resulting plasmid was named pKKα-2.

Immunological techniques. The polyclonal antibodies specific for the α antigen were prepared as described by Tasaka et al. (28). *E. coli* JM109 cells containing plasmids in which expression of the α antigen was under the control of the *trc* promoter were cultured in 50 ml of L-broth at 37°C for 4.5 h and induced to synthesize the α antigen by incubating the cells in 0.1 mM IPTG for 5 h. The cell extract was prepared in 5 ml of TE (10 mM Tris-HCl [pH 8.0], 1 mM EDTA) buffer by sonication at 100 W for 10 min followed by centrifugation at 12,000 rpm for 15 min. The extracted protein preparations (10 μl) were electrophoresed on a 15% polyacrylamide-SDS Laemmli gel (17) and electroblotted to a nitrocellulose membrane filter. The filter was shaken in 30 mM Tris-HCl (pH 7.4)–75 mM NaCl–1% bovine serum albumin at room temperature in order to block nonspecific binding. The immobilized proteins were reacted with 1:1,000-diluted antiserum at 8°C for 3 h. After the incubation, the filter was washed four times with TBSN (25 mM Tris-HCl [pH 7.5], 150 mM NaCl, 0.05% Nonidet P-40) at room temperature for 10 min, reacted with peroxidase-conjugated goat anti-rabbit immunoglobulin G (IgG) (DAKO Laboratories, Copenhagen, Denmark), and then washed with TBSN. Finally, the filter was reacted with a substrate solution containing 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 0.96 mg of 4-chloro-1-naphthol per ml, and 0.02% hydrogen peroxide and developed.

RESULTS

Southern hybridization analysis. Genomic Southern hybridization either with probe I, corresponding to Phe-1–Gly-19 or probe II, Gly-19–Gly-29 of the N-terminal amino acid sequence, was undertaken, and the results are presented in Fig. 3. Two DNA fragments of 5.2 kbp and 1.0 kbp were found to hybridize to both probes I and II when the DNA was digested with *Kpn*I and two fragments of 10 kbp

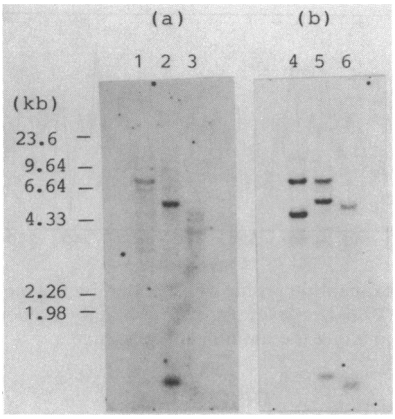

FIG. 3. (a) Hybridization pattern of 5'-end-labeled synthetic oligonucleotide probe I to *Bam*HI- (lane 1), *Kpn*I- (lane 2), and *Pst*I- (lane 3) digested *M. bovis* BCG DNA. (b) Hybridization of 5'-end-labeled synthetic oligonucleotide probe II to *Bam*HI- (lane 4), *Kpn*I- (lane 5), and *Pst*I- (lane 6) digested *M. bovis* BCG DNA. Experimental procedures are described in Materials and Methods.

and 7 kbp when it was digested with *Bam*HI. Although two radioactive bands were observed on *Pst*I-digested gels, one was less dense than in *Kpn*I- and *Bam*HI-digested gels when probe I was used. The observation suggests that *Pst*I may restrict genomic DNA in the region encoding the N-terminal portion of α antigen, resulting in low efficiency of hybridization with this probe. Indeed CTGCAG, which is the restriction site for *Pst*I, can be deduced from Leu-11 and Gln-12. The 5.2-kbp *Kpn*I fragment was cloned into pUC18 and screened by the colony hybridization method as described in Materials and Methods. Three positive colonies were found after the screening of 3,000 Amp^r white colonies. All positive colonies contained 5.2-kbp *Kpn*I fragments on agarose gel electrophoresis. One of them was designated pαL-1 and used for further study.

Nucleotide sequence. The physical map of pαL-1 is shown in Fig. 4. In order to clarify the structure of pαL-1, hybridization experiments with probes I and II were carried out. Probe II hybridized to a 0.8-kbp *Pst*I fragment, but probe I hybridized weakly to both 0.7-kbp and 0.8-kbp *Pst*I fragments (data not shown). These results were consistent with the data of genomic Southern hybridization, and *Pst*I was expected to digest a gene for the α antigen in the N-terminal region; then the 0.7-kbp and 0.8-kbp *Pst*I fragments were subcloned to pUC18, and each terminal nucleotide sequence was analyzed. The nucleotide sequence deduced from Leu-11 to Gly-29 of α antigen was found out at one terminal region of the 0.8-kbp fragment. The 0.9-kbp *Pst*I fragment was also subcloned into pUC18. The sequencing strategy for these fragments is shown in Fig. 2. The nucleotide sequence of the 1,165 bp and amino acid sequence deduced from that are shown in Fig. 5. The DNA sequence contains an open reading frame (ORF) beginning with ATG at position 91 and ending with a TGA stop codon at position 1060. The N-terminal amino acid sequence of α antigen (shown in Fig. 1) was found to be completely identical with that deduced from the nucleotide sequence beginning with the TTC codon at position 211. Therefore, the DNA sequence from position 91 to 210 is expected to encode the signal peptide which is necessary to secrete α antigen, and the mature protein must consist of 283 amino acid residues from the N-terminal Phe to the C-terminal Gly. Its theoretical molecular weight is calculated as 30,483 and is approximately identical with the

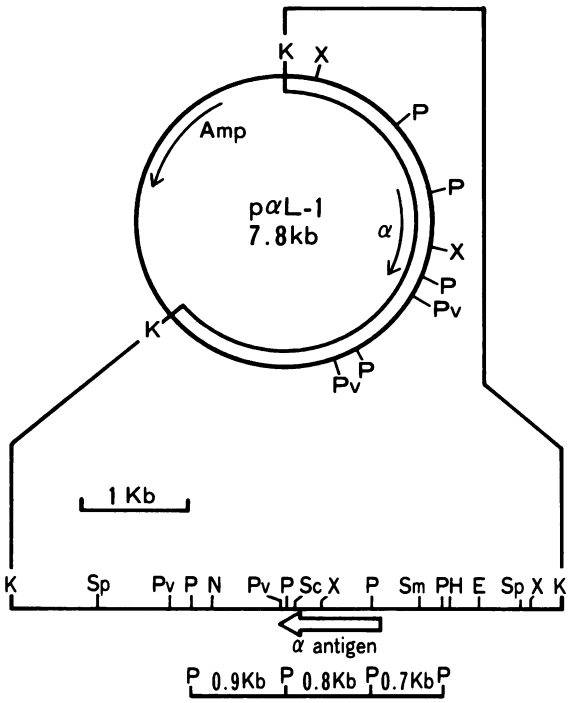

FIG. 4. Restriction endonuclease map of cloned *Kpn*I fragment of 5.2 kbp. The fragment for α antigen is shown. Arrows indicate direction of transcription. Restriction sites: E, *Eco*RI; K, *Kpn*I; H, *Hin*dIII; N, *Nco*I; P, *Pst*I; Pv, *Pvu*I; Sc, *Sca*I; Sm, *Sma*I; Sp, *Sph*I; X, *Xho*I. Amp indicates the β-lactamase gene.

reported value. The promoter sequences −35 (TCGACA) and −10 (TATGTT) with 18-bp spacing, similar to the *E. coli* consensus sequence, were located at positions 17 to 22 and 41 to 46, respectively. The Shine-Dalgarno sequence (SD sequence, ribosome-binding site) AAAGG was seen 9 bases upstream of the initiation codon. The search for hydrophilicity was carried out by the method of Hopp and Woods (10), and the hexapeptide profile is shown in Fig. 6.

Expression of α antigen in *E. coli*. By the use of the *E. coli* expression vector pKK233-2, we examined whether the cloned gene product reacts with antibody raised against α antigen. pKK233-2 contains the highly expressed *trc* promoter (a *trp-lac* fusion promoter [6] with the consensus 17-bp spacing between the *trp* −35 region and the *lacUV5* −10 region), the *lacZ* ribosome-binding site, and the ATG initiation codon located within a unique *Nco*I site (CCATGG). Digestion with *Nco*I generates a 5' protruding end containing the ATG codon, allowing direct ligation and expression of a foreign protein gene as an unfused protein. A part of the gene for *rrnB* 5S rRNA and its terminators T1 and T2 (3) is inserted downstream of the *Nco*I cloning site to enhance the turnover rate of *E. coli* RNA polymerase. The strategy for construction of the α antigen expression vector is shown in Fig. 7. *E. coli* JM109 was transformed with either pKKα-1 or pKKα-2, and transformants were designated AJα-1 and AJα-2, respectively. They were cultured in L-broth containing ampicillin and IPTG. The crude lysates of the cells were analyzed by SDS-polyacrylamide gel electrophoresis and Western blot (immunoblot) assay. The results are presented in Fig. 8. In AJα-1 and AJα-2, the polyclonal antibodies for α antigen, detected a single reactive protein migrating with a molecular weight similar to that of the

```
                                                   1                            30
                                                   ACGACTTTCGCCCGAATCGACATTTGGCCT
                                                                  - 35
                                                  60                            90
CCACACACGGTATGTTCTGGCCCGAGCACACGACGACATACAGGACAAAGGGGCACAGGT
             -10                                   SD
                                                 120                           150
ATGACAGACGTGAGCCGAAAGATTCGAGCTTGGGGACGCCGATTGATGATCGGCACGGCA
MetThrAspValSerArgLysIleArgAlaTrpGlyArgArgLeuMetIleGlyThrAla

                                                 180                           210
GCGGCTGTAGTCCTTCCGGGCCTGGTGGGGCTTGCCGGCGGAGCGGCAACCGCGGGCGCG
AlaAlaValValLeuProGlyLeuValGlyLeuAlaGlyGlyAlaAlaThrAlaGlyAla

                                                 240                           270
TTCTCCCGGCCGGGGCTGCCGGTCGAGTACCTGCAGGTGCCGTCGCCGTCGATGGGCCGC
PheSerArgProGlyLeuProValGluTyrLeuGlnValProSerProSerMetGlyArg

                                                 300                           330
GACATCAAGGTTCAGTTCCAGAGCGGTGGGAACAACTCACCTGCGGTTTATCTGCTCGAC
AspIleLysValGlnPheGlnSerGlyGlyAsnAsnSerProAlaValTyrLeuLeuAsp

                                                 360                           390
GGCCTGCGCGCCCAAGACGACTACAACGGCTGGGATATCAACACCCCGGCGTTCGAGTGG
GlyLeuArgAlaGlnAspAspTyrAsnGlyTrpAspIleAsnThrProAlaPheGluTrp

                                                 420                           450
TACTACCAGTCGGGACTGTCGATAGTCATGCCGGTCGGCGGGCAGTCCAGCTTCTACAGC
TyrTyrGlnSerGlyLeuSerIleValMetProValGlyGlyGlnSerSerPheTyrSer

                                                 480                           510
GACTGGTACAGCCCGGCCTGCGGTAAGGCTGGCTGCCAGACTTACAAGTGGGAAACCCTC
AspTrpTyrSerProAlaCysGlyLysAlaGlyCysGlnThrTyrLysTrpGluThrLeu

                                                 540                           570
CTGACCAGCGAGCTGCCGCAATGGTTGTCCGCCAACAGGGCCGTGAAGCCCACCGGCAGC
LeuThrSerGluLeuProGlnTrpLeuSerAlaAsnArgAlaValLysProThrGlySer

                                                 600                           630
GCTGCAATCGGCTTGTCGATGGCCGGCTCGTCGGCAATGATCTTGGCCGCCTACCACCCC
AlaAlaIleGlyLeuSerMetAlaGlySerSerAlaMetIleLeuAlaAlaTyrHisPro

                                                 660                           690
CAGCAGTTCATCTACGCCGGCTCGCTGTCGGCCCTGCTGGACCCCTCTCAGGGGATGGGC
GlnGlnPheIleTyrAlaGlySerLeuSerAlaLeuLeuAspProSerGlnGlyMetGly

                                                 720                           750
CTGATCGGCCTCGCGATGGGTGACGCCGGCGGTTACAAGGCCGCAGACATGTGGGGTCCC
LeuIleGlyLeuAlaMetGlyAspAlaGlyGlyTyrLysAlaAlaAspMetTrpGlyPro

                                                 780                           810
TCGAGTGACCCGGCATGGGAGCGCAACGACCCTACGCAGCAGATCCCCAAGCTGGTCGCA
SerSerAspProAlaTrpGluArgAsnAspProThrGlnGlnIleProLysLeuValAla

                                                 840                           870
AACAACACCCGGCTATGGGTTTATTGCGGGAACGGCACCCCGAACGAGTTGGGCGGTGCC
AsnAsnThrArgLeuTrpValTyrCysGlyAsnGlyThrProAsnGluLeuGlyGlyAla

                                                 900                           930
AACATACCCGCCGAGTTCTTGGAGAACTTCGTTCGTAGCAGCAACCTGAAGTTCCAGGAT
AsnIleProAlaGluPheLeuGluAsnPheValArgSerSerAsnLeuLysPheGlnAsp

                                                 960                           990
GCGTACAAGCCCGCGGGCGGGCACAACGCCGTGTTCAACTTCCCGCCCAACGGCACGCAC
AlaTyrLysProAlaGlyGlyHisAsnAlaValPheAsnPheProProAsnGlyThrHis

                                                1020                          1050
AGCTGGGAGTACTGGGGCGCTCAGCTCAACGCCATGAAGGGTGACCTGCAGAGTTCGTTA
SerTrpGluTyrTrpGlyAlaGlnLeuAsnAlaMetLysGlyAspLeuGlnSerSerLeu

                                                1080                          1110
GGCGCCGGCTGACGGGATCAACCGAAGGTTGCTTACCCGTCCACGTGAGATATACCTCGC
GlyAlaGly***

                                                1140                   1165
GACCGCCGTCTCGATCGGCATCCTGTTGTCGCTGATTGCACCACTAGGCCCCCCG
```

FIG. 5. DNA and deduced amino acid sequence of α antigen. Sequences similar to *E. coli* consensus sequences are labeled −35 and −10. SD marks a likely ribosome-binding site.

purified α antigen (Fig. 8, lanes 11, 12, 14, and 15). The reacting protein from AJα-1 was slightly larger than that from AJα-2, which was identical to the purified α antigen, because pKKα-1 lacked the nucleotide sequence for the C-terminal 6 amino acids and TGA termination codon but possessed the sequence for an additional 14 amino acid residues derived from the *Hin*dIII linker and 5S rRNA included in pKK233-2, as shown in Fig. 9 (2, 3). The theoretical molecular weights of these antibody-reactive proteins from AJα-2 and AJα-1 were 30,483 and 31,508, respectively, which were consistent with experimental values on the Western blot assay. It can therefore be concluded that the gene for mature α antigen was expressed in AJα-2.

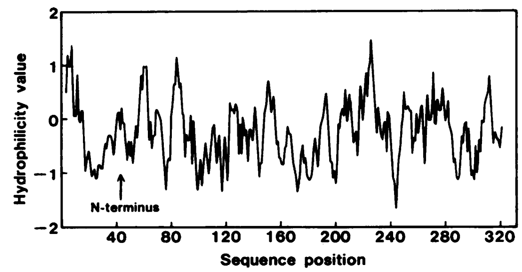

FIG. 6. Hexapeptide profile of α antigen by the method of Hopp and Woods (10). The average hydrophilicity values are plotted versus position along the amino acid sequence.

DISCUSSION

In this paper, the gene cloning of and the nucleotide sequence specific for α antigen are reported. The conclusion that the cloned DNA is specific for α antigen was drawn from the evidence that the N-terminal amino acid sequence of α antigen coincided perfectly with the deduced sequence and the polyclonal antibody raised against α antigen recognized the product in the lysate of *E. coli* harboring the recombinant plasmid. Interestingly, examination of the N-terminal amino acid sequence revealed that it was highly homologous to the MPB59 antigen which was recently reported to be widely distributed in the culture filtrate of mycobacteria, including *M. leprae*, and probably corresponded to BCG antigen 85B (22, 35). MPB59 was reported to be 28 kDa in size, which was close to the 30 kDa deduced from our theoretical calculation of molecular weight based on total sequence of the gene. We cannot answer the question of whether both proteins are identical or partially similar until we further characterize the gene for the MPB59 antigen.

Regarding the isolation of α antigen gene, many attempts with a mixed-probe method by short oligonucleotides were unsuccessful because of the high G+C content (63%) characteristic of *M. bovis* BCG (13). The method which we developed in the present work proved very promising for the isolation of such GC-rich genes. Previously, Shinnick (23) reported that the codon usage of the 65-kDa antigen gene in *M. tuberculosis* was extremely biased to the use of G+C (87%) in the third position. Ullrich et al. (33) first used the extended synthetic oligonucleotide probe for detecting the epidermal growth factor receptor gene. With these facts in mind, we designed two kinds of single long probes which were restricted to G or C in the third position. By the combination of probes I and II, the gene for α antigen was detected very efficiently in the Southern and the colony hybridization experiments. While this manuscript was being prepared, two research groups (7, 11) reported on the cloning of enzyme genes from *Streptomyces* species by the use of mixed probes including both G and C in the third position. However, our single-probe method should be much more efficient for detecting the genes of organisms with a high G+C content, such as *Mycobacterium*, *Streptomyces*, and *Thermus* species. More importantly, the present method obviates the time- and effort-consuming protein purification needed for the antibody-based techniques.

Inspection of the nucleotide sequence shown in Fig. 5 shows that the sequences TCGACA (positions 17 to 22) and TATGTT (positions 41 to 46) are similar to the consensus sequence for the −35 and −10 regions of *E. coli* promoters, respectively. There also exists the purine-rich SD sequence (AAAGG) 9 bp upstream of the ATG initiation codon for the α antigen ORF. Although the definite mycobacterial pro-

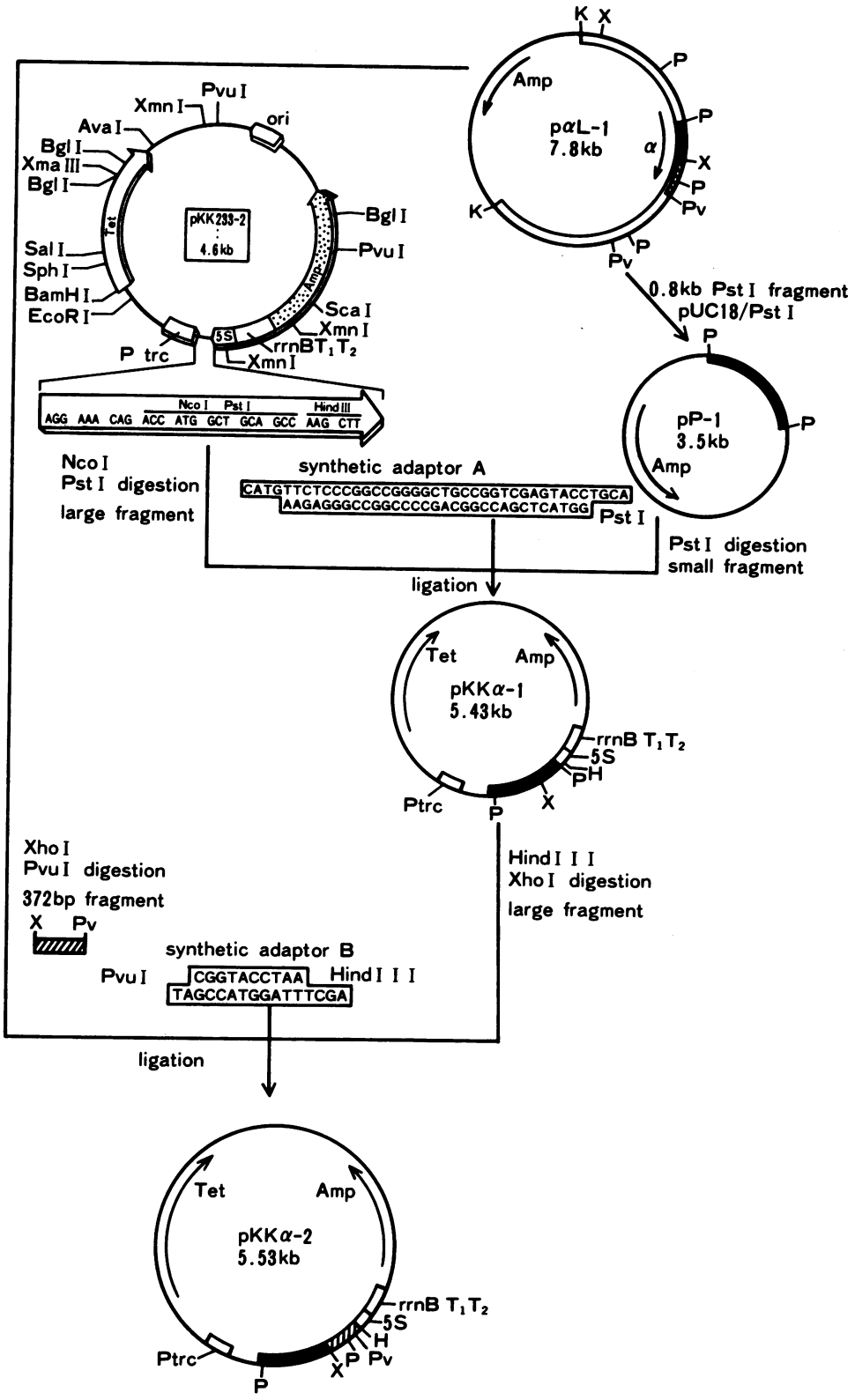

FIG. 7. Schematic outline of the construction of recombinant plasmid between *E. coli* expression vector and cloned mycobacterial DNA. The insert DNA is shown by a thick black line. The cross-hatched region of pKKα-2 shows newly inserted DNA containing sequences for the C-terminal amino acids of α antigen. See Fig. 4 legend for restriction site abbreviations.

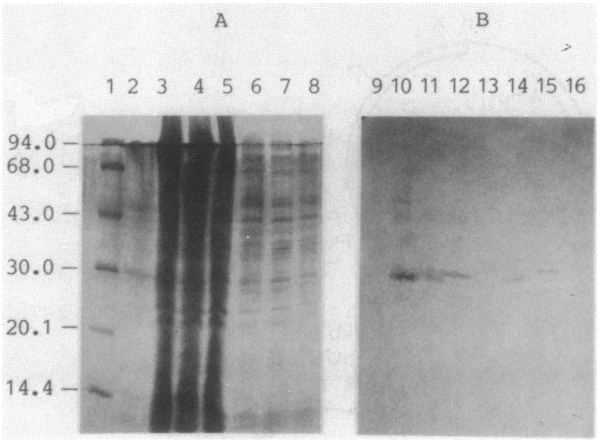

FIG. 8. Analysis of proteins from recombinant clones. Experimental procedures are described in Materials and Methods. (A) Stained with Coomassie brilliant blue for the visualization of total proteins; (B) protein which reacted with antibody against α antigen. Lanes: 1 and 9, size markers (in kilodaltons); 2 and 10, 2 μg of purified α antigen; 3 and 11, 10× concentrated lysate of clone AJα-2; 4 and 12, 10× concentrated lysate of clone AJα-1; 5 and 13, 10× concentrated control lysate of JM109(pKK233-2); 6 and 14, lysate of clone AJα-2; 7 and 15, lysate of clone AJα-1; 8 and 16, control lysate of JM109(pKK233-2). Molecular size standards (Pharmacia) are as follows: α-lactalbumin (14.4 kDa), soybean trypsin inhibitor (20.1 kDa), carbonic anhydrase (30.0 kDa), ovalbumin (43.0 kDa), bovine serum albumin (68.0 kDa), and phosphorylase B (94.0 kDa).

moter and SD sequence have not been determined experimentally, these sequences found upstream of the α antigen gene are presumed to control gene expression. The ORF of α antigen began with ATG at position 91 followed by hydrophobic amino acids upstream of mature α antigen. This hydrophobic region seemed to be a signal peptide since it had the following necessary characteristics: (i) basic amino acids exist at the N-terminal region, followed by (ii) hydrophobic domains, and (iii) a signal peptidase recognition sequence, Ala-X-Ala (34), just in front of the N-terminus of mature α antigen. The presence of signal peptide suggests that the α antigen obtained from the culture filtrate is not a degradation product but is extracellularly secreted from living cells. To our knowledge, this is the first report on the cloning of a secreting protein of mycobacteria. The codon usage of α antigen is shown in Table 1. The observation that the overall base composition of the gene for α antigen is 64% G+C and the base composition in the third position is 86% G+C is reminiscent of the 65.5% and 87% G+C, respectively, for the 65-kDa antigen (23). But in the α antigen, the repetitious nature cannot be observed, in contrast to the 65-kDa antigen.

The cloned gene was expressed by the *E. coli* expression vector pKK233-2 in which the *trc* promoter was used. Two important questions can be asked: whether the promoter for α antigen can be recognized by *E. coli* RNA polymerase, and whether the mycobacterial signal peptide is able to work in *E. coli*. These questions are projects for further study. In contrast to an expression system from a gene library, however, the system we used is a definite way to produce nonfusion proteins except Met at the N-termini with a molecular weight comparable to the same antigen in mycobacteria. On the Western blot assay, approximately 10 to 20 μg per 50 ml of culture medium was estimated to be produced in AJα-2. Further improvement of the expression level would be needed for diagnostic and medical application of this antigen.

The species specificity of α antigens from *M. tuberculosis*, *M. bovis*, *Mycobacterium kansasii*, *Mycobacterium marinum*, *Mycobacterium scrofulaceum*, *Mycobacterium gordonae*, *Mycobacterium suzlgai*, *Mycobacterium avium*, and *Mycobacterium intracellulare* has been investigated (28–30) by agar gel diffusion techniques with the respective absorbed anti-α serum, and it was concluded that species-specific antigenic determinants were observed in *M. kansasii*, *M. marinum*, *M. scrofulaceum*, *M. gordonae*, and *M. suzlgai* and that complex specific antigenic determinants were observed in the *M. tuberculosis-M. bovis-M. microti* complex and the *M. avium-M. intracellulare* complex. The gene for this antigen, therefore, can be a good candidate for the source of a defined immunodiagnostic reagent to identify these species of slowly growing mycobacteria. In this re-

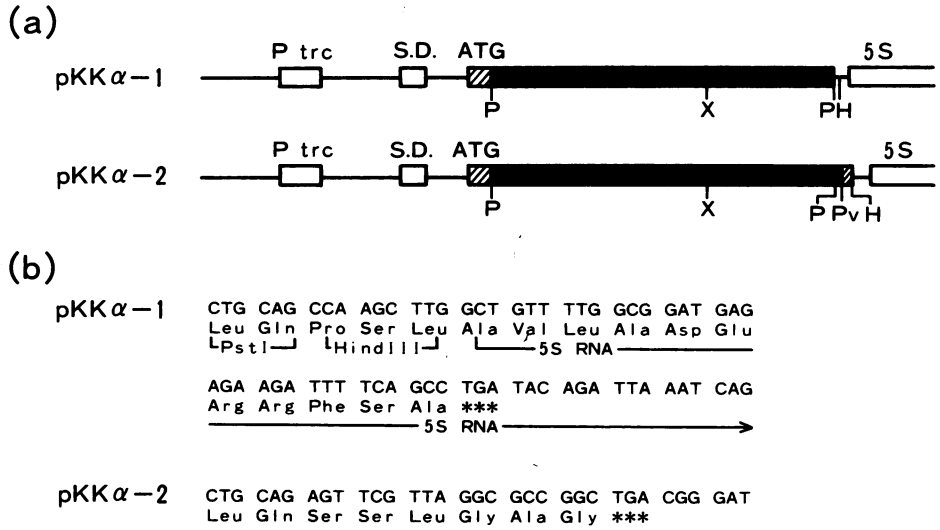

FIG. 9. (a) Structure of expression vector for α antigen. The black and cross-hatched blocks indicate mycobacterial DNA and synthetic adaptor, respectively. (b) Deduced amino acids at C-terminal region of α antigen in pKKα-1 and pKKα-2. ***, Termination codon.

TABLE 1. Codon usage in α antigen gene[a]

AA	Codon	Frequency	AA	Codon	Frequency	AA	Codon	Frequency	AA	Codon	Frequency
Phe	TTT	0	Ser	TCT	1	Cys	TGT	0	Tyr	TAT	2
Phe	TTC	10	Ser	TCC	3	Cys	TGC	3	Tyr	TAC	12
Leu	TTG	5	Ser	TCG	11	Trp	TGG	10	***	TAG	0
Leu	TTA	1	Ser	TCA	1	***	TGA	1	***	TAA	0
Leu	CTT	0	Pro	CCT	2	Arg	CGT	1	His	CAT	0
Leu	CTC	4	Pro	CCC	8	Arg	CGC	3	His	CAC	3
Leu	CTG	13	Pro	CCG	11	Arg	CGG	2	Gln	CAG	13
Leu	CTA	1	Pro	CCA	0	Arg	CGA	0	Gln	CAA	2
Val	GTT	4	Ala	GCT	3	Gly	GGT	7	Asp	GAT	2
Val	GTC	4	Ala	GCC	17	Gly	GGC	19	Asp	GAC	11
Val	GTG	3	Ala	GCG	4	Gly	GGG	6	Glu	GAG	8
Val	GTA	0	Ala	GCA	5	Gly	GGA	1	Glu	GAA	1
Ile	ATT	0	Thr	ACT	1	Ser	AGT	2	Asn	AAT	0
Ile	ATC	7	Thr	ACC	6	Ser	AGC	9	Asn	AAC	17
Met	ATG	8	Thr	ACG	2	Arg	AGG	1	Lys	AAG	9
Ile	ATA	2	Thr	ACA	0	Arg	AGA	0	Lys	AAA	0

[a] The numbers indicate the frequency of occurrence of the codons in the α antigen gene. AA, Amino acid. ***, Termination codon.

spect, research for the species-specific or complex specific epitope is important. In the hexapeptide profile (Fig. 6), some hydrophilic domains can be seen, suggesting the possible existence of some epitopes. These regions should be the first targets to determine the structures of specific epitopes experimentally.

Finally, we stress that the genetic information on a secreted protein from *M. bovis* BCG could be utilized in making a useful vaccine vehicle to produce and secrete a vaccinal protein from *M. bovis* BCG living cells. This vaccination system is expected to bring about a synergistic effect derived from the antigenicity of the secreting protein and adjuvant activity of *M. bovis* BCG cell wall and could be realized in the future since the transfection system of *M. bovis* BCG with mycobacteriophage has recently been reported (15).

ACKNOWLEDGMENTS

We thank Hirokazu Morisawa for his helpful discussion and Sachiko Mukai for synthesizing the oligonucleotide probes.

LITERATURE CITED

1. **Amann, E., and J. Brosius.** 1985. "ATG vectors" for regulated high-level expression of cloned genes in *Escherichia coli*. Gene **40**:183–190.
2. **Brosius, J.** 1984. Toxicity of an overproduced foreign gene product in *Escherichia coli* and its use in plasmid vectors for the selection of transcription terminators. Gene **27**:161–172.
3. **Brosius, J., T. J. Dull, D. D. Sleeter, and H. F. Noller.** 1981. Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. J. Mol. Biol. **148**:107–127.
4. **Buchanan, T. M., H. Nomaguchi, D. C. Anderson, R. A. Young, T. P. Gillis, W. J. Britton, J. Ivanyi, A. H. J. Kolk, O. Closs, B. R. Bloom, and V. Mehra.** 1987. Characterization of antibody-reactive epitopes on the 65-kilodalton protein of *Mycobacterium leprae*. Infect. Immun. **55**:1000–1003.
5. **Cohen, M. L., L. W. Mayer, H. S. Rumschlag, M. A. Yakrus, W. D. Jones, Jr., and R. C. Good.** 1987. Expression of proteins of *Mycobacterium tuberculosis* in *Escherichia coli* and potential of recombinant genes and proteins for development of diagnostic reagents. J. Clin. Microbiol. **25**:1176–1180.
6. **de Boer, H. A., L. J. Comstock, and M. Vasser.** 1983. The *tac* promoter: a functional hybrid derived from the *trp* and *lac* promoters. Proc. Natl. Acad. Sci. USA **80**:21–25.
7. **Fishman, S. E., K. Cox, J. L. Larson, P. A. Reynolds, E. T. Seno, W.-K. Yeh, R. Van Frank, and C. L. Hershberger.** 1987. Cloning genes for the biosynthesis of a macrolide antibiotic. Proc. Natl. Acad. Sci. USA **84**:8248–8252.
8. **Grunsteir, M., and J. Wallis.** 1979. Colony hybridization. Methods Enzymol. **68**:379–389.
9. **Hattori, M., and Y. Sakaki.** 1986. Dideoxy sequencing method using denatured plasmid templates. Anal. Biochem. **152**:232–238.
10. **Hopp, T. P., and K. R. Woods.** 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA **78**:3824–3828.
11. **Hoshiko, S., O. Makabe, C. Nojiri, K. Katsumata, E. Satoh, and K. Nagaoka.** 1987. Molecular cloning and characterization of the *Streptomyces hygroscopicus* α-amylase. J. Bacteriol. **169**:1029–1036.
12. **Husson, R. N., and R. A. Young.** 1987. Genes for the major protein antigens of *Mycobacterium tuberculosis*: the etiologic agents of tuberculosis and leprosy share an immunodominant antigen. Proc. Natl. Acad. Sci. USA **84**:1679–1683.
13. **Imaeda, T.** 1985. Deoxyribonucleic acid relatedness among selected strains of *Mycobacterium tuberculosis*, *Mycobacterium bovis*, *Mycobacterium bovis* BCG, *Mycobacterium microti*, and *Mycobacterium africanum*. Int. J. Syst. Bacteriol. **35**:147–150.
14. **Jacobs, W. R., M. A. Docherty, R. Curtiss III, 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.
15. **Jacobs, W. R., Jr., M. Tuckman, and B. R. Bloom.** 1987. Introduction of foreign DNA into mycobacteria using a shuttle phasmid. Nature (London) **327**:532–535.
16. **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.
17. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
18. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **Mehra, V., D. Sweetser, and R. A. Young.** 1986. Efficient mapping of protein antigenic determinants. Proc. Natl. Acad. Sci. USA **83**:7013–7017.
20. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
21. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. **Nagai, S., M. Harboe, and M. E. Patarroyo.** 1986. Comparative studies of MPB59, a widely cross-reacting protein of *Mycobacterium bovis* BCG, with MPB70, MPB80, and MPB64, p. 43–50. *In* Twenty-first Joint Research Conference on Tuberculosis, U.S.-Japan Cooperative Medical Science Program. Research

Institute of Tuberculosis, Japan Anti-Tuberculosis Association, Tokyo.
23. **Shinnick, T. M.** 1987. The 65-kilodalton antigen of *Mycobacterium tuberculosis*. J. Bacteriol. **169**:1080–1088.
24. **Shinnick, T. M., D. Sweetser, J. Thole, J. van Embden, and R. A. Young.** 1987. The etiologic agents of leprosy and tuberculosis share an immunoreactive protein antigen with the vaccine strain *Mycobacterium bovis* BCG. Infect. Immun. **55**:1932–1935.
25. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
26. **Suzuki, Y., K. Yoshinaga, Y. Ono, A. Nagata, and T. Yamada.** 1987. Organization of rRNA genes in *Mycobacterium bovis* BCG. J. Bacteriol. **169**:839–843.
27. **Szostak, J. W., J. I. Stiles, B.-K. Tye, P. Chiu, F. Sherman, and R. Wu.** 1979. Hybridization with synthetic oligonucleotides. Methods Enzymol. **68**:419–428.
28. **Tasaka, H., K. Kiyotani, and Y. Matsuo.** 1983. Purification and antigenic specificity of alpha protein (Yoneda and Fukui) from *Mycobacterium tuberculosis* and *Mycobacterium intracellulare*. Hirosima J. Med. Sci. **32**:1–8.
29. **Tasaka, H., and Y. Matsuo.** 1984. Specificity and distribution of alpha antigens on *Mycobacterium kansasii* and *Mycobacterium marinum*. Am. Rev. Respir. Dis. **130**:647–649.
30. **Tasaka, H., T. Nomura, and Y. Matsuo.** 1985. Specificity and distribution of alpha antigens of *Mycobacterium avium-intracellulare*, *Mycobacterium scrofulaceum*, and related species of mycobacteria. Am. Rev. Respir. Dis. **132**:173–174.
31. **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 antigen in *Escherichia coli*. Infect. Immun. **50**:800–806.
32. **Thole, J. E. R., W. J. Keulen, A. H. J. Kolk, D. G. Groothuis, L. G. Berwald, R. H. Tiesjema, and J. D. A. van Embden.** 1987. Characterization, sequence determination, and immunogenicity of a 64-kilodalton protein of *Mycobacterium bovis* BCG expressed in *Escherichia coli* K-12. Infect. Immun. **55**:1466–1475.
33. **Ullrich, A., L. Coussens, J. S. Hayflick, T. J. Dull, A. Gray, A. W. Tam, J. Lee, Y. Yarden, T. A. Libermann, J. Schlessinger, J. Downward, E. L. V. Mayes, N. Whittle, M. D. Waterfield, and P. H. Seeburg.** 1984. Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. Nature (London) **309**:418–425.
34. **von Heijne, G.** 1984. How signal sequences maintain cleavage specificity. J. Mol. Biol. **173**:243–251.
35. **Wiker, H. G., M. Harboe, S. Nagai, M. E. Patarroyo, C. Ramirez, and N. Cruz.** 1986. MPB59, a widely cross-reacting protein of *Mycobacterium bovis* BCG. Int. Arch. Allergy Appl. Immunol. **81**:307–314.
36. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
37. **Yoneda, M., Y. Fukui, and T. Yamanouchi.** 1965. Extracellular proteins of tubercle bacilli. V. Distribution of α and β antigens in various mycobacteria. Biken J. **8**:201–223.
38. **Yoneda, M., and Y. Fukui.** 1965. Isolation, purification, and characterization of extracellular antigens of *Mycobacterium tuberculosis*. Am. Rev. Respir. Dis. **92**:361–370.
39. **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.
40. **Young, R. A., B. R. Bloom, C. M. Grosskinsky, J. Ivanyi, D. Thomas, and R. W. Davis.** 1985. Dissection of *Mycobacterium tuberculosis* antigens using recombinant DNA. Proc. Natl. Acad. Sci. USA **82**:2583–2587.