INFECTION AND IMMUNITY, Sept. 1991, p. 3205–3212
0019-9567/91/093205-08$02.00/0

Vol. 59, No. 9

The Genes Coding for the Antigen 85 Complexes of *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG Are Members of a Gene Family: Cloning, Sequence Determination, and Genomic Organization of the Gene Coding for Antigen 85-C of *M. tuberculosis*

JEAN CONTENT,[1]* ANN DE LA CUVELLERIE,[1] LUKAS DE WIT,[1] VÉRONIQUE VINCENT-LEVY-FRÉBAULT,[2] JOSETTE OOMS,[1] AND JACQUELINE DE BRUYN[1]

Department of Virology, Institut Pasteur du Brabant, 642 rue Engeland, B-1180 Brussels, Belgium,[1] and Unité de la Tuberculose et des Mycobactéries, Institut Pasteur, Paris 75015, France[2]

Received 8 April 1991/Accepted 11 June 1991

A gene encoding the 33-kDa secreted protein of *Mycobacterium tuberculosis* (antigen 85-C) was isolated and sequenced. The corresponding DNA sequence contains a 1,020-bp coding region. The deduced amino acid sequence corresponds to a 340-residue protein consisting of a 46-amino-acid signal peptide and a 294-amino-acid mature protein. Comparison with previously described genes for the 30-kDa antigen (the α antigen of *M. bovis* BCG, also called antigen 85-B) and the 32-kDa antigens from *M. bovis* BCG and *M. tuberculosis* (antigens 85-A) indicates that the three genes share considerable sequence homology (70.8 to 77.5%) but may also code for distinctive epitopes. Strong differences among the three sequences are clearly visible upstream and downstream from the region coding for the mature proteins. The three genes have been detected in the genome of *M. bovis* BCG by Southern blot hybridization with three type-specific probes. Furthermore, hybridization of large DNA fragments (100 to 1,000 kbp) from *M. tuberculosis* separated by pulsed-field gel electrophoresis showed that the three genes coding for the antigen 85 complex are not clustered within the bacterial genome.

Three antigenic proteins (85-A, 85-B, and 85-C) have been identified by using counterimmunoelectrophoresis of mycobacterial culture filtrates (5, 27). Isoelectric focusing of the purified complex antigen also indicates that the complex consists of several proteins (5a).

We have previously purified a protein with a molecular mass of 32 kDa from zinc-deficient *Mycobacterium bovis* BCG culture filtrate. This protein was identified as antigen 85-A (7). Its NH_2-terminal amino acid sequence (Phe-Ser-Arg-Pro-Gly-Leu) is identical to that reported for the α antigen (antigen 85-B) protein purified from *M. bovis* BCG substrain Tokyo (28). The antigen 85 complex is present among different strains of mycobacteria (6). It is secreted by living bacilli as a predominant protein in normal Sauton culture filtrate and could be useful in the serodiagnosis of tuberculosis (24) and leprosy (19). Furthermore, the 32-kDa protein induces specific lymphoproliferation and gamma interferon production in peripheral blood leukocytes from patients with tuberculosis (11) and those with leprosy and from purified protein derivative- and lepromin-positive healthy subjects (13a). Recent findings indicate that the amount of 32-kDa protein-induced gamma interferon in BCG-sensitized mouse spleen cells is under probable *H-2* control (10). Finally, the high affinity of mycobacteria for fibronectin is related to proteins of the antigen 85 complex (1).

Wiker et al. (29) showed recently that antigens 85-A, 85-B, and 85-C isolated from *M. bovis* BCG culture filtrate present a few amino acid replacements in their NH_2-terminal region, strongly suggesting the existence of multiple genes coding for these proteins.

The gene encoding the α antigen from *M. bovis* BCG Tokyo has been cloned by Matsuo et al. (17). On the other hand, we have described the identification of the gene encoding the 32-kDa protein of *M. tuberculosis* (3), which presented 77.5% homology at the DNA level within the coding region with the α antigen gene (17). Moreover, we have recently isolated and sequenced a corresponding 32-kDa protein genomic clone from our λgt11 BCG library (prepared from *M. bovis* BCG 1173P2). The complete sequence of this gene is identical to that from *M. tuberculosis*, except for a single silent nucleotide change (8). This confirmed our previous findings that partial protein sequences of several tryptic peptides derived from the highly purified 32-kDa protein from *M. bovis* BCG presented the characteristic 85-A sequence (3, 8) and not the 85-B sequence (9a). Thus, it was likely but not demonstrated that the genome of *M. bovis* BCG contained at least two genes coding for antigens 85-A and 85-B.

In the present study, we demonstrated the existence of a third gene coding for the so-called 33-kDa secreted protein, or antigen 85-C, of mycobacteria. We also showed that the mycobacterial genome contains at least three nonclustered genes belonging to the antigen 85 gene family.

(A preliminary report of these results was presented at the First International Conference on the Pathogenesis of Mycobacterial Infections, Stockholm, Sweden, 27 to 29 June, 1990.)

MATERIALS AND METHODS

Preparation of genomic DNA. By following the procedure described by Thole et al. (23), *M. bovis* BCG was cultivated at 37°C in Sauton medium and harvested after incubation for an additional 18 h in the presence of 1% glycine which was

* Corresponding author.

added at the end of the late-exponential growth phase. The bacteria were treated with lysozyme and proteinase K, lysed with sodium dodecyl sulfate, phenol extracted, and precipitated with ethanol.

Genomic libraries. A λgt11 recombinant library constructed from genomic DNA of *M. tuberculosis* Erdman was obtained from R. A. Young (30).

A second λgt11 recombinant library was prepared with genomic DNA from *M. bovis* BCG (8).

Oligonucleotides. Oligonucleotides were synthesized on an Applied Biosystems 381A DNA synthesizer, purified on OPC cartridges (Applied Biosystems), lyophilized, and dissolved in TE buffer (10 mM Tris HCl, 1 mM EDTA, pH 8.0). ^{32}P labelling of the oligonucleotides was done as previously described (20).

PCR. (i) Amplification and cloning of genomic DNA. A 50-ng sample of *M. bovis* BCG DNA was amplified in a 50-μl reaction containing 1× polymerase chain reaction (PCR) buffer (Amersham), each deoxynucleoside triphosphate at 200 μM, the sense oligonucleotide primer P78 (5′-CCGGAATTCATGGGCCGTGACATCAAG) and the antisense oligonucleotide primer P79 (5′-CCGGAATTCGGTCTCCCACTTGTAAGT) at 1 μM each (for the locations of these two primers, see Fig. 2; to both oligonucleotides was added an *Eco*RI sequence preceded by three additional nucleotides), and 2 U of *Taq* DNA polymerase. After denaturation for 90 s at 94°C, the reaction was submitted to 40 cycles consisting of 1 min at 93°C (denaturation), 90 s at 55°C (annealing), and 2 min at 72°C (extension), followed by a 5-min final extension at 72°C. After extraction with 150 μl of chloroform, the amplified DNA was washed three times with 0.75 ml of H_2O in a Centricon 30 device for 6 min at 6,500 rpm in a Sorvall SS 34 rotor. After digestion with *Eco*RI, the DNA was ligated into *Eco*RI-digested, phosphatase-treated BlueScribe-M13+ vector. *Escherichia coli* DH5α (Gibco-BRL) was transformed and plated on Hybond-N filters. Colonies were selected by hybridization with ^{32}P-labeled oligonucleotide probes A (5′-TCGCCCGCCCTGTACCTG) and B (5′-TCACCTGCGGTTTATCTG). The hybridization and washing conditions for the oligonucleotides were as described by Jacobs et al. (12).

(ii) Amplification of λgt11 plaque DNA. A 10-μl volume of each λgt11 plaque (one plaque was suspended in 1 ml of SM medium containing 5% chloroform) was amplified in a 100-μl reaction containing each deoxynucleoside triphosphate at 200 μM, the sense oligonucleotide B and antisense oligonucleotide P79 primers at 1 μM each, and 1× PCR buffer (Amersham). After initial denaturation for 90 s at 94°C, 2 U of *Taq* DNA polymerase was added. The reaction was submitted to 40 cycles consisting of 1 min at 94°C (denaturation), 90 s at 60°C (annealing), and 2 min at 72°C (extension), followed by a 5-min final extension at 72°C. A 10-μl volume was analyzed on a 2% agarose gel stained with ethidium bromide.

Screening of the λgt11 *M. tuberculosis* and *M. bovis* BCG recombinant DNA libraries. The two λgt11 recombinant libraries were screened by colony hybridization (20), with an 800-bp *Hin*dIII fragment which does not discriminate the 85-A gene from the 85-B gene (see Fig. 2A). Twelve positive *M. tuberculosis* plaques and 12 *M. bovis* BCG plaques were retained and screened by PCR amplification with primers B (sense) and P79 (antisense). Non-A recombinant λgt11 recombinants were further screened by hybridization with ^{32}P-labelled oligonucleotide probe C (5′-TCGCAGAGCAACGGGCCA-GAACTAC) as described above. From the *M. tuberculosis* λgt11 library, one selected bacteriophage, no. 11, was partially digested with *Eco*RI and its 5-kbp insert was subcloned in BlueScribe-M13+. From this recombinant plasmid, named 11-2, a 3,500-bp *Bam*HI-*Eco*RI fragment was subcloned in M13-mp18 and M13-mp19 (20).

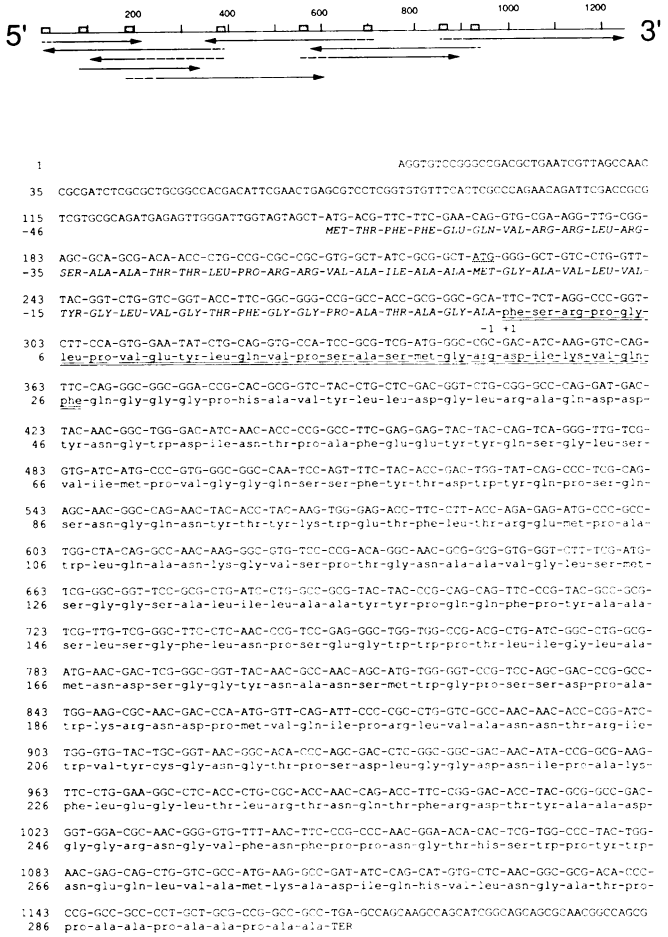

FIG. 1. Nucleotide and amino acid sequences of the 85-C antigen-containing region of *M. tuberculosis*. The previously identified 26-residue NH_2-terminal amino acid sequence of the mature protein (29) is doubly underlined. One additional ATG codon, downstream of the ATG at position 150, is underlined. Since the precise length of the signal sequence could not be determined, the option taken here represents the 46-amino-acid signal peptide corresponding to the ATG at position 150. The putative signal peptide sequence is represented in italic capitals. The top drawing represents the sequencing strategy. Arrows indicate the direction of dideoxy sequencing either in DNA subcloned as double-stranded DNA in BlueScribe-M13+ or as single-stranded DNA in the M13-mp18 vector. The entire sequence was determined by using eight synthetic oligonucleotides (gray boxes).

Recombinant DNA analysis. Recombinant DNA analysis was done as previously described (3).

Sequencing. Sequence analysis was done by the primer extension dideoxy termination method of Sanger et al. (21) after subcloning of specific fragments in BlueScribe-M13+ (4) or in M13-mp18 and M13-mp19 vectors. Sequence analysis was greatly hampered by the high G+C content of the *M. tuberculosis* DNA (65%). Sequencing reactions were therefore performed with the following DNA polymerases by manufacturers' protocols: T7 DNA polymerase (Sequenase, USB), T7 DNA polymerase (Pharmacia), and *Taq*

DNA polymerase (Promega) by using 7-deaza-dGTP instead of dGTP. Several oligodeoxynucleotides were synthesized and used to focus on ambiguous regions of the sequence. The sequencing strategy is summarized in Fig. 1.

Sequence comparison and analysis. Routine computer-aided analyses of the nucleic acid and deduced amino acid sequences were performed with the LGBC program of Bellon (2). Homology searches used the FASTA programs from Pearson and Lipman (18) and the various DNA and protein data banks of the EMBL server facilities. Multiple alignments were obtained with Align 1.01 (Scientific and Educational Software).

Southern blot analysis. Genomic DNA from *M. bovis* BCG was completely digested with *Sph*I, *Eco*RI, or *Kpn*I; electrophoresed on a 1% agarose gel; transferred to a Hybond-N filter (Amersham) after denaturation and neutralization, and either hybridized with ^{32}P-labelled oligonucleotide probes (A, B, B1, and C) under the conditions previously described (12) or randomly primed, ^{32}P-labelled DNA restriction fragments that were found to discriminate the 85-A, 85-B, and 85-C genes.

Probe 85-A was a 230-bp *Pst*I fragment from plasmid BY-5 (3) (see Fig. 2). Probe 85-B was a 400-bp *Sma*I-*Eco*RV fragment from an 85-B recombinant plasmid, named 5.1, derived from our *M. bovis* BCG λgt11 library (for a map, see Fig. 5a; see also Fig. 2). Probe 85-C was a 280-bp *Sma*I-*Kpn*I fragment from plasmid 11.2 (see Fig. 4 and 2).

These DNA fragments were prepared by gel electrophoresis on low-melting-point agarose, followed by rapid purification on Qiagen (tip 5) by the manufacturer's protocol, and labelled in the presence of [α-^{32}P]dCTP (9).

Pulsed-field gel electrophoresis DNA separation. DNA preparation, restriction enzyme digestion, and pulsed-field gel electrophoresis were performed as previously described (25). Briefly, cells from fresh cultures were mixed with 1% (vol/vol) low-melting-point agarose and submitted to successive treatments with Zymolase (Seikagaki Kogyo, Tokyo, Japan), lysozyme, and sodium dodecyl sulfate in the presence of proteinase K (Boehringer GmbH, Mannheim, Germany). After inactivation of proteinase K with phenylmethylsulfonyl fluoride (Bio-Rad Laboratories), agarose blocks were digested overnight with 50 U of *Dra*I (Bio-Rad Laboratories). Then blocks were loaded into a 1% agarose gel prepared and electrophoresed in 0.66× TBE (Tris-boric acid-EDTA) (25). Field inversion gel electrophoresis was carried out with a Dnastar Pulse apparatus. Forward and reverse pulses, respectively, were set at 0.33 and 0.11 s at the beginning of the run and 60 and 20 (or 30 and 10) s at the end of the run, depending on the molecular weight zone to be expanded. The run time was set at 36 h, the voltage used was 100 V at about 325 mA, and the temperature was maintained at 18°C. Lambda concatemers were used as molecular weight markers. At the end of the run, the gels were stained with ethidium bromide, photographed under UV light, and transferred onto nylon membranes as described by Maniatis et al. (14).

Nucleotide sequence accession number. The DNA sequence in Fig. 1 has been submitted to the EMBL data library under accession no. X57229.

RESULTS

Cloning of the 85-C gene of *M. tuberculosis*. Since no specific probe or monoclonal antibody was available to detect specifically 85-C or a related antigen which was expected to bear extensive homology to the 85-A and 85-B genes, this screening required the development of a new procedure. We used a strategy based on PCR amplification of a 245-bp DNA fragment coding for amino acids 18 to 98 of mature antigen 85-A chosen because it is surrounded at both ends by highly conserved DNA sequences when the sequences of antigens 85-A and 85-B are aligned (see primers P78 and P79 in Fig. 2). We thus speculated that equivalent homology might exist with the sequence of antigen 85-C in the same region.

From *M. bovis* BCG genomic DNA, we readily obtained a 245-bp DNA fragment. The latter was purified and subcloned in a BlueScribe-M13+ vector after digestion with *Eco*RI. About 80 recombinant plasmid-containing colonies were tested by plating on nylon filters and hybridized under stringent conditions with a labeled synthetic oligonucleotide recognizing either the 85-A (5′-TCGCCCGCCCTGTACCTG) or the 85-B (5′-TCACCTGCGGTTTATCTG) sequence within the PCR-amplified fragment (Fig. 2). Several clones that hybridized with each probe were sequenced, and the sequences were all identical to the 85-A sequence in the clones hybridizing with probe A. In those hybridizing with probe 85-B, two kinds of sequences were found, i.e., the 85-B sequence (17) and a new sequence described in Fig. 2. The latter presents a marked sequence divergence covering 24 nucleotides which is totally distinct from sequences A and B (Fig. 2). (Homology to sequence B is only 33% in this region.) Assuming that this insert might represent an amplified fragment of the 85-C gene and that this 24-nucleotide sequence is characteristic of the putative 85-C gene, we synthesized an oligonucleotide probe (probe 85-C) based on this sequence.

The latter probe was labelled with ^{32}P and used to screen a collection of 24 λgt11 recombinant phage that were selected in our *M. tuberculosis* and *M. bovis* BCG λgt11 libraries by hybridization with an 800-bp *Hin*dIII DNA fragment of the previously cloned gene for 85-A (3).

Among those phage, we further selected those containing the B oligonucleotide sequence by an analytical PCR assay using oligonucleotides probe B (sense probe) and P79 (antisense probe). This small collection of 85-B-related λgt11 recombinants was then hybridized with the described oligonucleotide 85-C, and one hybridizing λgt11-*M. tuberculosis* recombinant was retained, characterized by restriction mapping, and sequenced.

Sequence of the 85-C gene. The 1,211-nucleotide sequence derived from various sequenced fragments is represented in Fig. 1. The DNA sequence contains a 1,020-bp open reading frame, starting at position 150 and ending with a TGA codon at position 1170. The common N-terminal amino acid sequence of the antigen 85 proteins, Phe-Ser-Arg-Pro-Gly-Leu (7), was located within this open reading frame, from the nucleotide sequence beginning with a TTC codon at position 288 (Fig. 1). Therefore, the DNA region upstream of this sequence is expected to code for a signal peptide required for secretion of this antigen. The mature protein consists of 294 amino acid residues, corresponding to a calculated molecular weight of 32,021.

Interestingly, the N-terminal sequence of the mature protein contains the entire 26-amino-acid sequence (Phe-Ser-Arg-Pro-Gly-Leu-Pro-Val-Glu-Tyr-Leu-Gln-Val-Pro-Ser-Ala-Ser-Met-Gly-Arg-Asp-Ile-Lys-Val-Gln-Phe) described by Wiker et al. (29), which differs from the sequence common to 85-B and 85-A only by an alanine instead of a proline in position 16 of the mature protein. Two ATG codons were found to precede the TTC phenylalanine codon at nucleotide position 288 (Fig. 1) in the same reading frame. Use of these two ATG codons would lead to the synthesis of signal

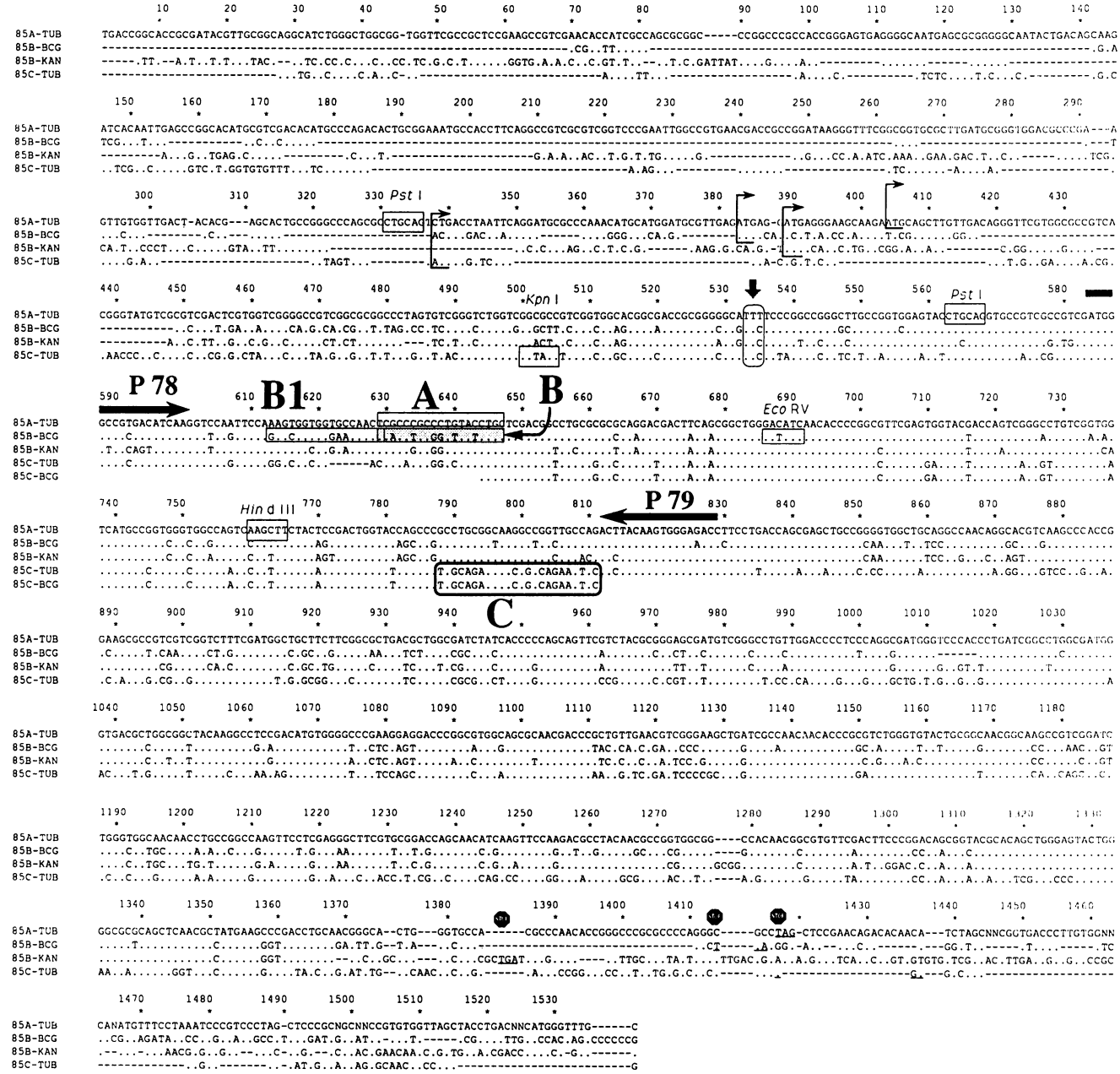

FIG. 2. Comparison of the DNA sequences of antigens 85-A, 85-B, and 85-C. All of the sequences are compared and aligned with the first line (gene 85-A). Dots indicate identical residues, dashes indicate gaps, and letters indicate substitutions. 85-A is the DNA sequence of *M. tuberculosis* (3), 85-B is the DNA sequence for the α antigen of *M. bovis* Tokyo (17), 85-C is the DNA sequence from antigen 85-C (this work), 85-B-Kan is the DNA sequence for antigen 85-B of *M. kansasii* (16), and 85-C-BCG is the partial DNA sequence of *M. bovis* BCG 1173P2 (this work). This sequence was obtained from a cloned, PCR-amplified DNA fragment. Arrowed brackets indicate the presumed initiation codons of the genes, the downward-pointing arrow indicates the first phenylalanine residue of the mature proteins, stop signs indicate the termination codons (underlined) of the genes, P78 and P79 are the sense and antisense primers used for PCR amplification, and A, B, B1, and C are the sequences used for synthesis of specific synthetic oligonucleotides probes and are framed. The indicated restriction sites were used to prepare the three type-specific probes.

peptides of either 21 or 46 amino acid residues (the latter situation is represented in Fig. 1).

The base composition of the antigen 85-C gene was identical to that of the antigen 85-A gene, with an overall G+C composition of 64. 57% and a strong preference for G or C in codon position 3 (average, 85%). In contrast to antigens 85-A and 85-B, which contain three cysteines, the sequence of antigen 85-C shows a single cysteine residue at position 209. In fact, the two substituted cysteines are located in the region of the mature 85-C protein which contains the largest divergent sequence block (Fig. 3), SQS-NGQNY (the corresponding DNA sequence was used to synthesize oligonucleotide probe C [see Materials and Methods]). Not surprisingly, this hydrophilic region is also the

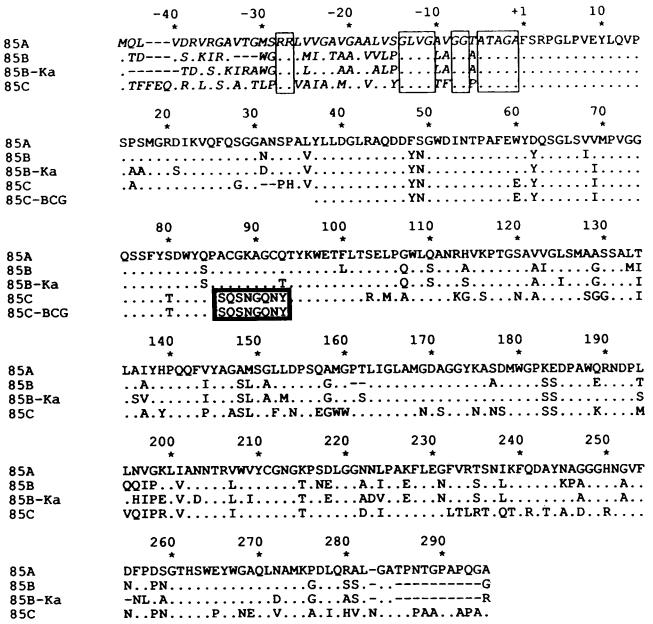

FIG. 3. Comparison of the preprotein sequences of antigens 85-A, 85-B, and 85-C. Amino acid sequences were aligned with the Align 1.01 program, which permits multiple alignments. All of the sequences are compared and aligned with the first line (the preprotein 85-A). In this presentation, sequence differences are outlined as follows: dots indicate identical residues, dashes indicate gaps, and letters indicate substitutions. 85-A is the protein sequence of *M. tuberculosis* (3), 85-B is the protein sequence of the α antigen of *M. bovis* Tokyo (17), 85-C is the protein sequence of antigen 85-C (this work), 85-B-Kans is the protein sequence of antigen 85-B of *M. kansasii* (16), and 85-C-BCG is the partial protein sequence of *M. bovis* BCG 1173P2 (this work). The C characteristic motif is framed.

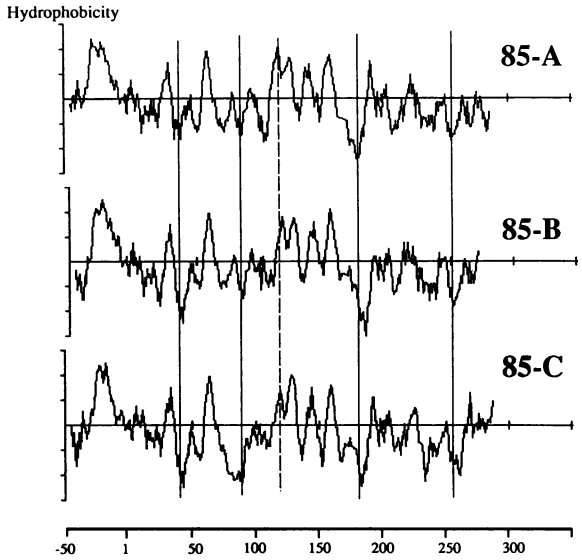

FIG. 4. Hydropathy patterns of the amino acid sequences of the *M. tuberculosis* 32-kDa protein (antigen 85-A), the α antigen of BCG (antigen 85-B), and antigen 85-C from *M. tuberculosis*. The sequences of the three preproteins (including the presumed signal peptide signals) were analyzed by using the method of Kyte and Doolittle (13), with a window of eight amino acids. Each bar on the horizontal axis represents 50 amino acids. Since the lengths of the signal sequences are slightly different (43, 40, and 46 residues for proteins 85-A, 85-B, and 85-C), the patterns are aligned to the initial residues of the three mature proteins. Solid lines are used to align hydrophilic peaks, and a dashed line is used to align hydrophobic peaks.

most divergent when the hydropathy plots of the three antigens are compared and thus could be either a variable epitope of all 85 antigens and/or a characteristic epitope of antigen 85-C, since it was also found in antigen 85-C from *M. bovis* BCG.

Another characteristic feature of antigen 85-C is the presence of the unusual hydrophobic repetitive proline-alanine motive PPAAPAAPAA at the carboxy terminus of the molecule.

Hydropathy pattern. The hydropathy pattern of *M. tuberculosis* antigen 85-C was determined by the method of Kyte and Doolittle (13). The octapeptide profiles were compared with those of antigens 85-A and 85-B (Fig. 4). As anticipated from the amino acid sequences, the patterns are roughly similar for the three antigens, except for some major differences at region 84 to 92 and in the carboxy-terminal part of the three proteins.

Sequence homologies. DNA sequences that encode antigens 85-A (3, 8), 85-B (14, 17), and 85-C were aligned (Fig. 2). The homology is maximal between the regions coding for the three mature proteins. In this region, the homology between 85-A and 85-B is 77.5%, whereas it reaches only 70.8% between the coding regions of the genes for 85-A and 85-C and 71.9% between those for 85-B and 85-C. Beyond nucleotide 1369 of the 85-A sequence and upstream nucleotide position 475 (i.e., within the signal sequence and promoter region), there is practically no homology among the three sequences. No significant homology to other DNA sequences present in the latest release of GenBank-EMBL was detected. Homologies at the amino acid level are presented in the alignment (Fig. 3), indicating higher homology between the mature 85-A and 85-B protein sequences (80.4%) than between 85-B and 85-C (68.1%) or 85-A and 85-C (70.4%).

Other comparisons between antigen 85-C and the entire SwissProt-NBRF data bank failed to detect any significant homologies to the antigen 85-C amino acid sequence. As for antigen 85-A, the 85-C sequence does not contain the RGD motif of fibronectin-binding proteins nor does it share any homology to the known fibronectin receptors or to the fibronectin-binding protein from *Staphylococcus aureus*.

Comparison of the partial, PCR-derived DNA sequence of the 85-C gene of *M. bovis* BCG 1173P2 with that of *M. tuberculosis* shows complete identity, including the characteristic region corresponding to synthetic oligonucleotide C (Fig. 2).

Upstream from nucleotide 644 of antigen 85A in Fig. 2 (corresponding to nucleotide 393 in gene C [Fig. 1]), the sequence reveals a stretch of 43 nucleotides which are perfectly identical to a homologous region in the 85B gene. It is likely that some of the amplified fragments accidentally recombined between highly homologous DNA regions, yielding a chimeric 85B-85C DNA insert. Numerous experiments, such as those presented in Fig. 5, allowed exclusion of the existence of such a hybrid structure in genomic DNA. This kind of PCR artifact has been observed previously when amplifying DNA sequences belonging to multigene families (15, 22).

Genome characterization. To confirm the existence of different genes coding for the antigen 85 complex, we digested *M. bovis* BCG genomic DNA with *Sph*I, *Eco*RI,

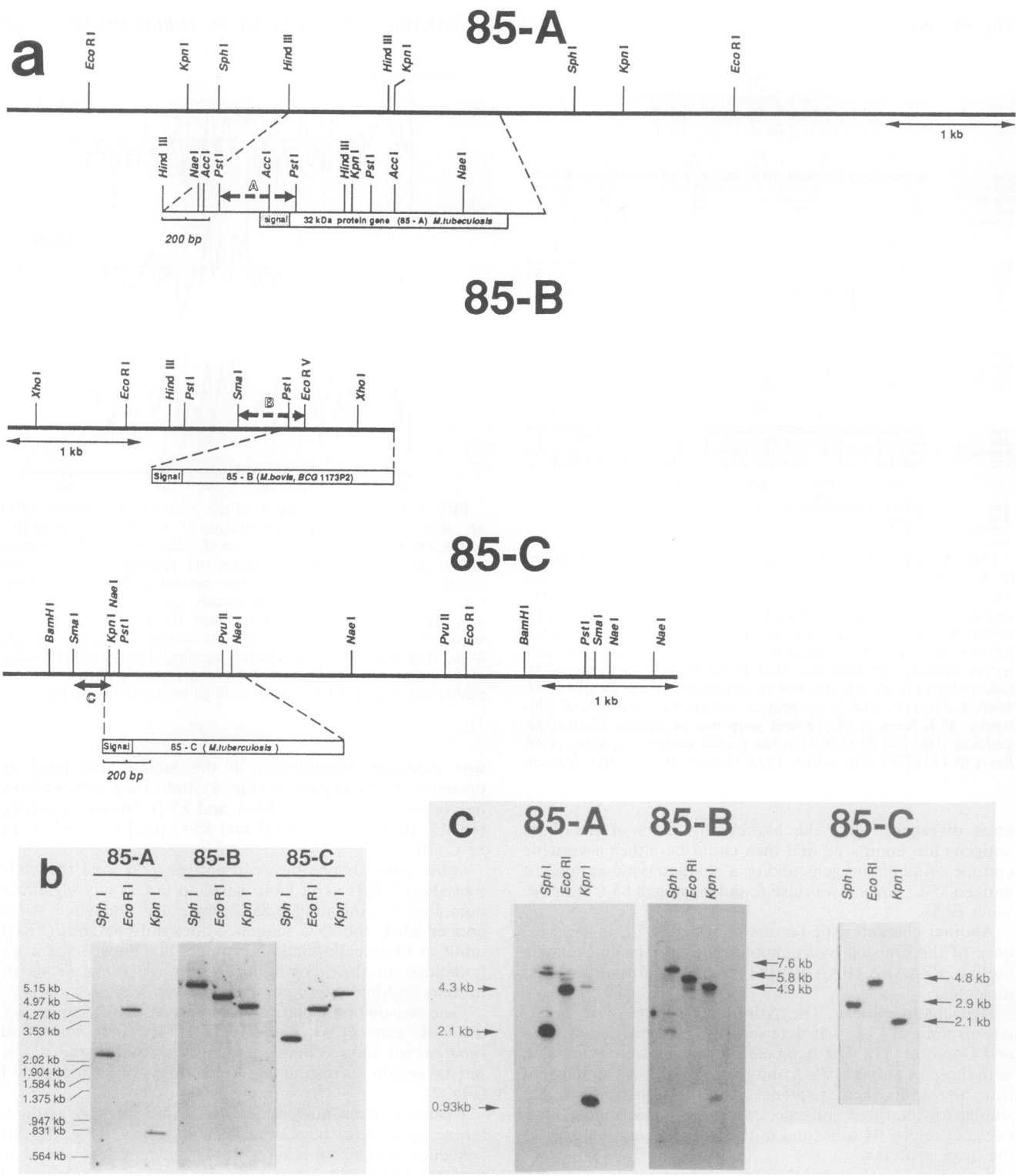

FIG. 5. (a) Restriction endonuclease maps of the genes for 85-A, 85-B, and 85-C. The map of the gene for 85-A is from Borremans et al. (3). The map of the 85-B gene was obtained from clone 5.1 derived from our *M. bovis* BCG 1173P2 λgt11 recombinant library (8). For the restriction enzymes used, this map is identical to that published for *M. bovis* BCG Tokyo (17). The coding region for the 85-B antigen is positioned as described by Matsuo et al. (17). The map of the 85-C gene corresponds to the restriction map of clone 11.2 from the *M. tuberculosis* λgt11 library obtained from R. A. Young (30) (see Materials and Methods). The position of the specific 5′ DNA restriction fragment used for Southern analysis is indicated on each map by a double arrow. (b) Southern analysis of total genomic DNA from *M. bovis* BCG 1173P2. A 15-μg sample of digested DNA was applied per lane. Hybridization was with oligonucleotide probes A, B1, and C (as described in the legend to Fig. 2) under the conditions described in Materials and Methods. Hybridization signals with oligonucleotide B were identical to those obtained with probe B1 (data not shown). Molecular weights of the hybridizing bands were calculated by comparison with standards. (c) Southern analysis of total genomic DNA from *M. bovis* BCG 1173P2. The procedure described for panel b was used. The three probes, however, were large DNA restriction fragments (as defined for panel a) which were labeled by random priming (see Materials and Methods). Parts 85-A and 85-B were obtained from a single filter, whereas 85-C was from a separate run.

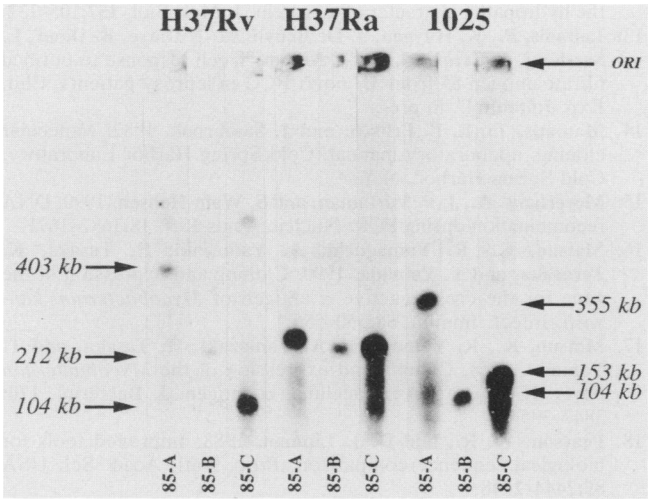

FIG. 6. Pulsed-field gel electrophoresis of *M. tuberculosis* DNA. DNAs from three strains of *M. tuberculosis* were digested with *Dra*I and separated by pulsed-field electrophoresis on an agarose gel, together with a bacteriophage λ DNA ladder, as described in Materials and Methods. After transfer to nylon filters, hybridization with probes 85-A, 85-B, and 85-C was done as described in the legend to Fig. 5. The molecular weights of the hybridizing bands were calculated by comparison with those of the λ DNA ladder. *ORI*, origin of migration.

and *Kpn*I and examined the distribution of radioactive signals in a Southern blot after hybridization with three specific oligonucleotide probes (A, B, and C) (Fig. 2; see also Materials and Methods). We obtained three clearly distinct patterns, confirming the specificity of these probes. Similar type-specific profiles were obtained with three random-priming-labeled DNA restrictions fragments (probe 85-A, 230 bp; probe 85-B, 400 bp; probe 85-C, 280 bp) which were selected within the promoter signal sequence regions of the three DNAs (Fig. 2 and 5a). With these three DNA restriction fragments, additional weak bands are also observed which clearly correspond to cross-hybridization of the probes to the two other genes. With probe 85-C, an additional *Kpn*I fragment was observed that did not hybridize to oligonucleotide probe C. This probably indicates that the corresponding *Kpn*I site is located upstream of this gene. Furthermore, the sizes of the observed restriction fragments are not always exactly as expected from the restrictions maps of the corresponding cloned genes. These discrepancies probably correspond to some minor sequence differences (restriction polymorphism), possibly in noncoding DNA regions (outside of the DNA coding for antigen 85), between our strain of *M. bovis* BCG and *M. bovis* BCG Tokyo and *M. tuberculosis*.

Pulsed-field analysis of *M. tuberculosis* genomic DNA. When we hybridized our largest available 85-A clone, BY-5 (Fig. 5a), with oligonucleotide B, we failed to detect any positive signal whereas oligonucleotide A gave positive hybridization (data not shown). This indicates that gene B is not located within 2 to 2.5 kb of the 5′ border and 4.0 kb of the 3′ border of gene A (Fig. 5a). To confirm and extend this result, we further hybridized pulsed-field-separated, *Dra*I-digested *M. tuberculosis* genomic DNA with three specific DNA probes, 85-A, 85-B, and 85-C, under stringent conditions.

Eight strains of *M. tuberculosis* were compared, showing six different patterns, three of which are illustrated in Fig. 6. For most of the strains examined, the three probes hybridized to fragments of different sizes. For instance, in *M. tuberculosis* H37Ra, the sizes of the *Dra*I fragments hybridizing with probes 85-A, 85-B, and 85-C were about 242, 212, and 225 kb for strain H37Ra; 403, 212, and 104 kb for strain H37Rv; and 355, 104, and 153 kb for strain 1025. Although various strains show some restriction fragment length polymorphism with restriction endonuclease *Dra*I, the simplest interpretation of these results is that the three antigen 85 genes are distantly located (>100 kb) within the mycobacterial genome.

DISCUSSION

The present work confirms the proposal of Wiker et al. (29) that the three secreted mycobacterial antigens 85-A, 85-B, and 85-C are the products of different genes. The three genes have now been cloned and sequenced, although in different mycobacterial species: antigen 85-A in *M. bovis* BCG 1173P2 and *M. tuberculosis* (3, 8), antigen 85-B in *M. bovis* Tokyo (17) and *M. kansasii* (16), and antigen 85-C in *M. tuberculosis* (this work). We have also obtained a partial nucleotide sequence for gene 85-B (data not shown) and a short fragment of gene 85-C of *M. bovis* BCG 1173P2 (Fig. 2).

The three genes are also detectable in genomic DNA by using specific DNA-hybridizing probes. This allowed us to show that the gene family also exists in atypical mycobacteria, such as *M. kansasii* and *M. avium* (7a). We cannot exclude the possibility that more than three genes code for members of the antigen 85 complex. In particular, isoelectric focusing of purified *M. bovis* filtrate shows the presence of several bands corresponding to proteins of 33 kDa, the presumed antigen 85-C. Further work will show whether these molecules are the result of posttranslational processing or whether they are indeed encoded by different genes. Southern blot analysis, which showed a single-band pattern for three restriction enzymes (Fig. 5b), does not allow resolution of the question, since it was done with a promotor signal sequence-specific probe to minimize cross-hybridization to genes 85-A and 85-B.

Wiker and Harboe (26) have reported that the three antigens are secreted by *M. bovis* BCG but at different rates. Their suggestion that secretion rate, which was found to be 85-A > 85-B > 85-C, could be correlated to the Gln/Glu ratios of these proteins does not seem to hold true, since we found these ratios to be 2.42 (85-A), 1.77 (85-B), and 2.0 (85-C). Possibly, signal peptide sequences (which we found to be very different; Fig. 3) or other factors play a role.

We do not know at what rate the three genes are expressed within bacterial cells. Since we show here that the three genes are not clustered within the bacterial genome, they probably function independently, each with its own individual promotor. It will be of interest to investigate the strength (and possible regulation and coordination) of these three promoters, since their sequences show no homology (Fig. 2).

The function of the antigen 85 secreted proteins is unknown. The fact that these antigens are produced and secreted abundantly may explain their immunodominant character at the T-cell level, as reflected by stimulation of specific proliferation and gamma interferon secretion in peripheral blood lymphocytes from most tuberculin- and lepromin-positive volunteers. Furthermore, these proteins may also be involved in, but not sufficient for, invasiveness and tissue destruction since they are also abundantly produced by nonpathogenic mycobacterial strains such as *M. bovis* BCG. It has also been suggested (19, 24) that antigens 85-A, 85-B, and 85-C could be very useful for the serodiag-

nosis of tuberculosis (24a) and leprosy. Availability of the cloned *M. tuberculosis* 85-C gene should now allow expression and testing of this antigen alone or in combination with antigens 85-A and 85-B, as well as exploration of potential B- or T-cell epitopes within its sequence.

ACKNOWLEDGMENTS

We are grateful to R. Young for the *M. tuberculosis* λgt11 library, J. Weckx for BCG preparations, M. Harboe (Oslo, Norway) for advice, and L. Schoofs for sequence alignments. We thank K. Huygen and L. Machtelinckx for critical reading and J. Herinckx for typing of the manuscript.

This work was partly supported by grant 3.4542.86 of the National Fund for Medical Scientific Research (Belgium) and by Innogenetics N.V.A. A. de la Cuvellerie holds a fellowship from the Institut pour l'Encouragement de la Recherche Scientifique dans l'Industrie et l'Agriculture.

REFERENCES

1. **Abou-Zeid, C., T. L. Ratliff, H. G. Wiker, M. Harboe, J. Bennedsen, and G. A. W. Rook.** 1988. Characterization of fibronectin-binding antigens released by *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG. Infect. Immun. **56**:3046–3051.
2. **Bellon, B.** 1988. Apple Macintosh programs for nucleic and protein sequence analysis. Nucleic Acids Res. **16**:1837–1846.
3. **Borremans, M., L. De Wit, G. Volckaert, J. Ooms, J. De Bruyn, K. Huygen, J. P. Van Vooren, M. Stelandre, R. Verhofstadt, and J. Content.** 1989. Cloning, sequence determination, and expression of a 32-kilodalton protein gene of *Mycobacterium tuberculosis*. Infect. Immun. **57**:3123–3130.
4. **Chen, E. J., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast simple method for sequencing plasmid DNA. DNA **4**:165–170.
5. **Closs, O., M. Harboe, N. H. Axelsen-Christensen, and M. Magnussen.** 1980. The antigens of Mycobacterium bovis, strain BCG, studied by crossed immuno-electrophoresis: a reference system. Scand. J. Immunol. **12**:249–263.

5a. **De Bruyn, J.** Unpublished data.

6. **De Bruyn, J., R. Bosmans, J. Nyabenda, and J. P. Van Vooren.** 1989. Effect of zinc deficiency on the appearance of two immunodominant protein antigens (32-kDa and 65-kDa) in culture filtrates of mycobacteria. J. Gen. Microbiol. **135**:7984.
7. **De Bruyn, J., K. Huygen, R. Bosmans, M. Fauville, R. Lippens, J. P. Van Vooren, P. Falmagne, M. Weckx, H. G. Wiker, M. Harboe, and M. Turneer.** 1987. Purification, partial characterization and identification of a 32 kDa protein antigen of Mycobacterium bovis BCG. Microb. Pathog. **2**:351–360.

7a. **de la Cuvellerie, A., and J. Content.** Unpublished data.

8. **De Wit, L., A. de la Cuvellerie, J. Ooms, and J. Content.** 1990. Nucleotide sequence of the 32 kDa-protein gene (antigen 85A) of Mycobacterium bovis BCG. Nucleic Acids Res. **18**:3995.
9. **Feinberg, A. P., and R. Vogelstein.** 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:613.

9a. **Gras, H., J. De Bruyn, and J. Content.** Unpublished data.

10. **Huygen, K., K. Palfliet, F. Jurion, J. Hilgers, R. ten Berg, J. P. Van Vooren, and J. De Bruyn.** 1989. *H-2*-linked control of in vitro gamma interferon production in response to a 32-kilodalton antigen (P32) of *Mycobacterium bovis* bacillus Calmette-Guérin. Infect. Immun. **56**:3196–3200.
11. **Huygen, K., J. P. Van Vooren, M. Turneer, R. Bosmans, P. Dierckx, and J. De Bruyn.** 1988. Specific lymphoproliferation, γ-interferon production and serum immunoglobulin G directed against a purified 32-kDa mycobacterial antigen (P32) in patients with active tuberculosis. Scand. J. Immunol. **27**:187–194.
12. **Jacobs, R. Rudersdorf, S. D. Neill, J. P. Dougherty, E. L. Brown, and E. F. Fritsch.** 1988. The thermal stability of oligonucleotide duplexes is sequence independent in tetraalkylammonium salt solutions: application to identifying recombinant DNA clones. Nucleic Acids Res. **16**:4637–4650.
13. **Kyte, J., and R. F. Doolittle.** 1982. Simple method for displaying the hydropathy character of a protein. J. Mol. Biol. **157**:105–132.

13a. **Launois, P., K. Huygen, J. De Bruyn, M. N'Diaye, B. Diouf, L. Sarthouj, J. Grimaud, and J. Millan.** T cell response to purified filtrate antigen 85 from *M. bovis* BCG in leprosy patients. Clin. Exp. Immunol., in press.

14. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. **Meyerhans, A., J. P. Vartanian, and S. Wain-Hobson.** 1990. DNA recombination during PCR. Nucleic Acids Res. **18**:1687–1691.
16. **Matsuo, K., R. Yamaguchi, A. Yamazaki, H. Tasaka, K. Terasaka, and T. Yamada.** 1990. Cloning and expression of the gene for the cross-reactive α antigen of *Mycobacterium kansasii*. Infect. Immun. **58**:550–556.
17. **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.
18. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85**:2444–2448.
19. **Rumschlag, H. S., T. S. Shinnick, and M. L. Cohen.** 1988. Serological response of patients with lepromatous and tuberculous leprosy to 30-, 31-, and 32-kilodalton antigens of *Mycobacterium tuberculosis*. J. Clin. Microbiol. **26**:2200–2202.
20. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
22. **Shuldiner, A. R., A. Nirula, and J. Roth.** 1989. Hybrid DNA artifact from PCR of closely related target sequences. Nucleic Acids Res. **17**:4409.
23. **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.
24. **Turneer, M., J. P. Van Vooren, J. De Bruyn, E. Serruys, P. Dierckx, and J. C. Yernault.** 1988. Humoral immune response in human tuberculosis: immunoglobulins G, A, and M directed against the purified P32 protein antigen of *Mycobacterium bovis* bacillus Calmette-Guérin. J. Clin. Microbiol. **26**:1714–1719.

24a. **van Vooren, J. P., A. Drowart, M. de Cock, A. van Onckelen, M. H. D'Hoope, J. C. Yernault, C. Valcke, and K. Huygen.** Humoral immune response of tuberculosis patients against the three components of the *Mycobacteriom bovis* BCG 85 complex separated by isoelectric focusing. J. Clin. Microbiol., in press.

25. **Vincent, Lévy-Frébault, V., M.-F. Thorel, A. Varnerot, and B. Gicquel.** 1989. DNA polymorphism in mycobacterium paratuberculosis, "wood pigeon mycobacteria," and related mycobacteria analyzed by field inversion gel electrophoresis. J. Clin. Microbiol. **27**:2823–2826.
26. **Wiker, H. G., M. Harboe, and S. Nagai.** 1991. Localization index for distinction between extracellular and intracellular antigens of Mycobacterium *tuberculosis*. J. Gen. Microbiol. **137**:875–884.
27. **Wiker, H. G., M. Harboe, S. Nagai, and J. Bennedsen.** 1990. Quantitative and qualitative studies on the major extracellular antigen of Mycobacterium tuberculosis H37Rv and Mycobacterium bovis BCG. Am. Rev. Respir. Dis. **141**:830–838.
28. **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.
29. **Wiker, H. G., K. Sletten, S. Nagai, and M. Harboe.** 1990. Evidence for three separate genes encoding the proteins of the mycobacterial antigen 85 complex. Infect. Immun. **58**:272–274.
30. **Young, R. A., B. R. Bloom, C. M. Grosskinsky, J. Ivanji, D. Thomas, and R. W. Davis.** 1985. Dissection of Mycobacterium tuberculosis antigens using recombinant DNA. Proc. Natl. Acad. Sci. USA **82**:2583–2587.