A Novel Pathogenic Taxon of the Mycobacterium tuberculosis Complex, Canetti: Characterization of an Exceptional Isolate from Africa

Dick Van SooLingen,1,* Theo HoogEndBoezem,2 Petra E. W. De Haas,1 Peter W. M. Hermans,2 MARIANNE A. KOEDAM,3 Koos S. TePPEMA,3 Patrick J. Brennan,4 GURDyal S. BesRA,4 Francoise PortaeLs,5 Janetta Top,6 Leo M. SchoULs,6 and Jan D. A. Van EMBdEN6

Diagnostic Laboratory for Infectious Diseases and Perinatal Screening, 1 Laboratory for Pathology/Immunology Biology, and Research Laboratory for Infectious Diseases, 6 National Institute oF Public Health and the Environment (RIVM), 3720 BA BILTHOVEN, and Department of Pediatrics, Sophia Children’s Hospital, Erasmus University Rotterdam, 3000 DR Rotterdam, 2 The Netherlands; Department of Microbiology, Colorado State University, Fort Collins, Colorado 80523; and Mycobacteriology Unit, Institute of Tropical Medicine Prince Leopold, Antwerp, Belgium

In an attempt to characterize an unusual mycobacterial strain isolated from a 2-year-old Somali patient with lymphadenitis, we applied various molecular methods not previously used for the taxonomic classification of mycobacteria. This isolate, designated So93, did not differ from Mycobacterium tuberculosis in the biochemical tests and in its 16S rRNA sequence, but produced smooth and glossy colonies, which is highly exceptional for this species. This smooth phenotype was unstable and switched nonreversibly to a rough colony morphology with a low frequency. The two colony types were equally virulent for the guinea pig, exhibiting characteristic tuberculous disease. Both morphotypes had shorter generation times than the M. tuberculosis reference laboratory strain H37Rv and clinical isolates of M. tuberculosis and Mycobacterium bovis. Furthermore, the So93 isolate differed from all M. tuberculosis complex strains described thus far by having only a single copy of insertion sequence IS1081, an unusual composition of the direct repeat cluster, and a characteristic phenolic glycolipid and lipooligosaccharide. This glycolipid had previously been observed only in a smooth isolate of M. tuberculosis obtained in 1969 by Canetti in France. Analysis of the Canetti strain showed that it shared virtually all genetic properties characteristic of So93, distinguishing these two strains from the known M. tuberculosis complex taxa, M. tuberculosis, Mycobacterium africanum, M. bovis, and Mycobacterium microti. The natural reservoir, host range, and mode of transmission of the group of bacteria described in this paper are presently unknown. This study, partly based on not previously used molecular criteria, supports the idea that the established members within the M. tuberculosis complex and the newly described Canetti grouping should be regarded as a single species, which likely will be designated “M. tuberculosis”.

The Mycobacterium tuberculosis complex constitutes a genetically closely related group, and its members, M. tuberculosis, Mycobacterium africanum, Mycobacterium bovis, and Mycobacterium microti, may be considered as subspecies of M. tuberculosis (39, 47). The close relatedness between M. tuberculosis complex bacteria has been established by DNA-DNA hybridization (≥95%) (1), multiple-locus enzyme electrophoresis (15), and sequencing of 16S ribosomal RNA and housekeeping genes (15). Furthermore, repetitive DNA elements, such as the insertion sequence IS6110 (43) and the direct repeat (DR) (18), have been found restricted to the M. tuberculosis complex. Nevertheless, the host range and pathogenicity of the M. tuberculosis complex species vary enormously. The natural reservoir of M. tuberculosis and M. africanum is limited to humans (47) and that of M. microti is limited to voles (48). In contrast, the host range of M. bovis is very broad, and this species causes disease among a wide range of wild and domestic mammals as well as in humans (38). Colonies of primary M. tuberculosis cultures almost invariably have a characteristic patterned texture, due to tight cording of the bacterial cells (32), and this feature is often used to distinguish M. tuberculosis from other mycobacterial species. Differences in colony morphotypes of M. tuberculosis H37 and M. bovis vaccine strain BCG have been associated with variation in virulence (10, 27, 28, 36, 49). Similar to the method used by Calmette to attenuate M. bovis, resulting in the BCG vaccine strain (10), H37 was attenuated by serial passage on potato bile medium. Both the attenuated M. bovis strain (M. bovis BCG) and the attenuated H37 strain (H37Ra) differed in colony morphology from the parental strains.

Variation in colony morphology is more commonly observed among non-M. tuberculosis complex mycobacteria, such as Mycobacterium avium. The smooth, transparent morphotype of this species is associated with virulence, since most patient-derived strains showing this morphotype in primary culture are more pathogenic than rough variants when tested in animal models (24, 34). Colony morphology conversion of mycobacteria has been shown to be associated with changes in cell wall contents, such as glycolipids (3, 5), lipooligosaccharides (LOSs), and mycosides (12, 29). However, the biochemical and genetic bases for the differences in colony morphology of M. tuberculosis complex strains are presently unknown. Daffe et al. (12) showed that the smooth Canetti M. tuberculosis strain differs from the commonly rough strains by the presence of large amounts of LOS. Furthermore, they speculated that smooth forms of M. tuberculosis may be attenuated by virtue of their ability to synthesize large amounts of phthioceranic acids instead of other phthienoic acids. Such smooth colony variants are usually not encountered among clinical M. tuberculosis isolates.

In 1993, we encountered an M. tuberculosis complex strain with a smooth and glossy colony morphology which had been obtained from a lymph node biopsy from the neck of a 2-year-old Somali patient who came to The Netherlands in 1993.
DNA typing of this strain with various repetitive DNA elements suggested that this strain was unusual and that it should be grouped differently from the previously described members of the *M. tuberculosis* complex group of bacteria. This study was undertaken to precisely classify this mycobacterial isolate and to compare it with the previously isolated Canetti *M. tuberculosis* strain, which also had a smooth morphotype. For this purpose, various molecular methods were applied, which have not previously been used to classify mycobacteria. Although these novel methods are likely more precise than the traditional, mainly numerical phenotypic methods (22), our results are consistent with the previously held view (39, 47) that all groupings within the *M. tuberculosis* complex, including the Somali isolate and the Canetti strain, should be regarded as subspecies belonging to a single bacterial species.

**MATERIALS AND METHODS**

Mycobacterial strains and phenotypic identification. All strains used in this study are listed in Table 1. Strain So93 was obtained from A. de Jong, Regional Health Laboratory, Deventer, The Netherlands. The Canetti strain was kindly provided by B. Gicquel and V. Vincent (Institute Pasteur, Paris, France). Drug susceptibility tests were done according to the recommendations of the manufacturer (GenProbe, Inc., San Diego, Calif.).

DNA typing of this strain with various repetitive DNA elements suggested that this strain was unusual and that it should be grouped differently from the previously described members of the *M. tuberculosis* complex group of bacteria. This study was undertaken to precisely classify this mycobacterial isolate and to compare it with the previously isolated Canetti *M. tuberculosis* strain, which also had a smooth morphotype. For this purpose, various molecular methods were applied, which have not previously been used to classify mycobacteria. Although these novel methods are likely more precise than the traditional, mainly numerical phenotypic methods (22), our results are consistent with the previously held view (39, 47) that all groupings within the *M. tuberculosis* complex, including the Somali isolate and the Canetti strain, should be regarded as subspecies belonging to a single bacterial species.

**MATERIALS AND METHODS**

**Mycobacterial strains and phenotypic identification.** All strains used in this study are listed in Table 1. Strain So93 was obtained from A. de Jong, Regional Health Laboratory, Deventer, The Netherlands. The Canetti strain was kindly provided by B. Gicquel and V. Vincent (Institute Pasteur, Paris, France). Drug susceptibility tests were done according to the recommendations of the manufacturer (GenProbe, Inc., San Diego, Calif.).

**Electron microscopy.** For scanning electron microscopy (SEM), smooth and rough colony variants of strain So93 on Löwenstein-Jensen plates were fixed with 4% (vol/vol) formaldehyde-1% (vol/vol) glutaraldehyde in 0.1 M sodium cacodylate-HCl buffer (pH 7.4). Fixed cells were adsorbed to glow-discharged carbon-stabilized Formvar-coated grids coated with 1% (vol/vol) OsO4 for 90 min, washed again, dehydrated in a graded series of acetone, and finally washed twice in 100% (vol/vol) acetone. The specimens were examined in a Philips EM 400T TEM at an operating voltage of 60 kV.

**Transmission electron microscopy (TEM).** Cells were fixed in 1.25% (wt/vol) glutaraldehyde-0.1 M sodium cacodylate-HCl buffer (pH 7.2). Fixed cells were adsorbed to glow-discharged carbon-stabilized Formvar-coated grids coated with 1% (vol/vol) poly-L-lysine (molecular weight, 30,000 to 70,000), and gold sputtered in a sputter coater (Polaron E500). The specimens were examined in a Philips EM 400T TEM at an operating voltage of 60 kV.

**Analysis of the cell wall glycolipids.** For transmission electron microscopy (TEM), cells were fixed in 1.25% (wt/vol) glutaraldehyde-0.1 M sodium cacodylate-HCl buffer (pH 7.2). Fixed cells were adsorbed to glow-discharged carbon-stabilized Formvar-coated grids coated with 1% (vol/vol) poly-L-lysine (molecular weight, 30,000 to 70,000), and gold sputtered in a sputter coater (Polaron E500). The specimens were examined in a Philips EM 400T TEM at an operating voltage of 60 kV.

**Characterization of housekeeping genes.** Single-strand conformation polymorphism analysis of housekeeping genes of So93, *M. tuberculosis* H37Rv, and *M. bovis* BCG vaccine strain P3 was performed as described by Orita et al. (26).

**Analysis of the cell wall glycolipids.** For transmission electron microscopy (TEM), cells were fixed in 1.25% (wt/vol) glutaraldehyde-0.1 M sodium cacodylate-HCl buffer (pH 7.2). Fixed cells were adsorbed to glow-discharged carbon-stabilized Formvar-coated grids coated with 1% (vol/vol) poly-L-lysine (molecular weight, 30,000 to 70,000), and gold sputtered in a sputter coater (Polaron E500). The specimens were examined in a Philips EM 400T TEM at an operating voltage of 60 kV.

**DNA sequencing.** The 5' 550-bp region of the 16S rRNA gene of So93 was amplified by PCR with the broad-host-range primers 16SlF (AGAGTTTGATCCTTGGAGGCAGCAAG) and 16S1R (CTTACCGGCTCAATTCGATACTG). The amplified product was sequenced by the chain termination method (30). The sequence was analyzed with the Windows version of the Gelcompar software, version 3.10 (Applied Maths, Kortrijk, Belgium), after alignment of the autoradiograms with a scanner at 190 dpi (HP Scanjet IICx/T; Hewlett-Packard, Cambridge, Mass.). Comparison of the 16S rRNA fingerprints was performed as described previously (45). The pulsed-field gel electrophoresis and RFEL fingerprints were normalized with *M. tuberculosis*-specific bands present in the fingerprints of all strains. Comparison of the fingerprints was performed by using the Dice similarity coefficient applied to peaks and the unweighted pair group mean arithmetic clustering method (30).

**Characterization of housekeeping genes.** Single-strand conformation polymorphism analysis of housekeeping genes of So93, *M. tuberculosis* H37Rv, and *M. bovis* BCG vaccine strain P3 was performed as described by Orita et al. (26).

**Analysis of the cell wall glycolipids.** For transmission electron microscopy (TEM), cells were fixed in 1.25% (wt/vol) glutaraldehyde-0.1 M sodium cacodylate-HCl buffer (pH 7.2). Fixed cells were adsorbed to glow-discharged carbon-stabilized Formvar-coated grids coated with 1% (vol/vol) poly-L-lysine (molecular weight, 30,000 to 70,000), and gold sputtered in a sputter coater (Polaron E500). The specimens were examined in a Philips EM 400T TEM at an operating voltage of 60 kV.

**DNA sequencing.** The 5' 550-bp region of the 16S rRNA gene of So93 was amplified by PCR with the broad-host-range primers 16SlF (AGAGTTTGATCCTTGGAGGCAGCAAG) and 16S1R (CTTACCGGCTCAATTCGATACTG). The amplified product was sequenced by the chain termination method (30). The sequence was analyzed with the Windows version of the Gelcompar software, version 3.10 (Applied Maths, Kortrijk, Belgium), after alignment of the autoradiograms with a scanner at 190 dpi (HP Scanjet IICx/T; Hewlett-Packard, Cambridge, Mass.). Comparison of the 16S rRNA fingerprints was performed as described previously (45). The pulsed-field gel electrophoresis and RFEL fingerprints were normalized with *M. tuberculosis*-specific bands present in the fingerprints of all strains. Comparison of the fingerprints was performed by using the Dice similarity coefficient applied to peaks and the unweighted pair group mean arithmetic clustering method (30).

**Characterization of housekeeping genes.** Single-strand conformation polymorphism analysis of housekeeping genes of So93, *M. tuberculosis* H37Rv, and *M. bovis* BCG vaccine strain P3 was performed as described by Orita et al. (26).

**Analysis of the cell wall glycolipids.** For transmission electron microscopy (TEM), cells were fixed in 1.25% (wt/vol) glutaraldehyde-0.1 M sodium cacodylate-HCl buffer (pH 7.2). Fixed cells were adsorbed to glow-discharged carbon-stabilized Formvar-coated grids coated with 1% (vol/vol) poly-L-lysine (molecular weight, 30,000 to 70,000), and gold sputtered in a sputter coater (Polaron E500). The specimens were examined in a Philips EM 400T TEM at an operating voltage of 60 kV.

**DNA sequencing.** The 5' 550-bp region of the 16S rRNA gene of So93 was amplified by PCR with the broad-host-range primers 16SlF (AGAGTTTGATCCTTGGAGGCAGCAAG) and 16S1R (CTTACCGGCTCAATTCGATACTG). The amplified product was sequenced by the chain termination method (30). The sequence was analyzed with the Windows version of the Gelcompar software, version 3.10 (Applied Maths, Kortrijk, Belgium), after alignment of the autoradiograms with a scanner at 190 dpi (HP Scanjet IICx/T; Hewlett-Packard, Cambridge, Mass.). Comparison of the 16S rRNA fingerprints was performed as described previously (45). The pulsed-field gel electrophoresis and RFEL fingerprints were normalized with *M. tuberculosis*-specific bands present in the fingerprints of all strains. Comparison of the fingerprints was performed by using the Dice similarity coefficient applied to peaks and the unweighted pair group mean arithmetic clustering method (30).

**Characterization of housekeeping genes.** Single-strand conformation polymorphism analysis of housekeeping genes of So93, *M. tuberculosis* H37Rv, and *M. bovis* BCG vaccine strain P3 was performed as described by Orita et al. (26).

**Analysis of the cell wall glycolipids.** For transmission electron microscopy (TEM), cells were fixed in 1.25% (wt/vol) glutaraldehyde-0.1 M sodium cacodylate-HCl buffer (pH 7.2). Fixed cells were adsorbed to glow-discharged carbon-stabilized Formvar-coated grids coated with 1% (vol/vol) poly-L-lysine (molecular weight, 30,000 to 70,000), and gold sputtered in a sputter coater (Polaron E500). The specimens were examined in a Philips EM 400T TEM at an operating voltage of 60 kV.

**DNA sequencing.** The 5' 550-bp region of the 16S rRNA gene of So93 was amplified by PCR with the broad-host-range primers 16SlF (AGAGTTTGATCCTTGGAGGCAGCAAG) and 16S1R (CTTACCGGCTCAATTCGATACTG). The amplified product was sequenced by the chain termination method (30). The sequence was analyzed with the Windows version of the Gelcompar software, version 3.10 (Applied Maths, Kortrijk, Belgium), after alignment of the autoradiograms with a scanner at 190 dpi (HP Scanjet IICx/T; Hewlett-Packard, Cambridge, Mass.). Comparison of the 16S rRNA fingerprints was performed as described previously (45). The pulsed-field gel electrophoresis and RFEL fingerprints were normalized with *M. tuberculosis*-specific bands present in the fingerprints of all strains. Comparison of the fingerprints was performed by using the Dice similarity coefficient applied to peaks and the unweighted pair group mean arithmetic clustering method (30).
**RESULTS**

**Phenotypic characterization.** Strain So93 showed abundant growth on Löwenstein-Jensen (LJ) medium tubes with a smooth, white, and glossy colony morphology within 6 days. Although this colony morphology and rapid growth are non-characteristic of *M. tuberculosis* complex isolates, the 16S rRNA Genprobe test for the *M. tuberculosis* complex was positive. Furthermore, So93 was positive for niacin production, nitrate reductase, nitrite accumulation, and pyrazinamidase activities and was negative for the Tween hydrolysis test. Growth of the strain was not inhibited by thiophen, nor stimulated by pyruvate on LJ medium. The isolate was susceptible to isoniazid, rifampin, ethambutol, cycloserine, ethionamide, thiacetazone, amikacin, ciprofloxacin, and rifabutin; resistant to streptomycin and clarithromycin; and moderately susceptible to p-aminosalicylic acid (2 µg/ml). All of these characteristics, except the moderate susceptibility to p-aminosalicylic acid, are specific for *M. tuberculosis*.

The generation time of So93 during exponential growth in liquid TA medium (days 4 to 9) was determined in three independent experiments and found to be 17 h, whereas values of 25.2, 21.4, 40.4, and 39.8 h were calculated for two clinical isolates of *M. tuberculosis* and two isolates of *M. bovis*, respectively (Fig. 1). This doubling time of the So93 isolate in TA medium is significantly shorter (*P < 0.0001*) than those found among the other *M. tuberculosis* complex isolates.

**Colony morphology and microscopic variation.** The So93 isolate showed a smooth, white, and glossy colony morphology on solid medium. When single colonies were replated, rough colony variants of So93 were observed with a frequency of about 1 out of 500 colonies (Fig. 2A). Single colonies of both variants were repeatedly plated on LJ medium and 7H10 plates. The progeny derived from smooth colonies produced mainly smooth colonies, whereas rough colony variants were reproducibly observed with a frequency of 1:500 to 1:1,000. No revertants from rough to smooth colonies was observed (frequency of <1:10,000). A rough So93 colony variant was selected for further analysis and was designated So93R. The generation time of So93R in TA medium was 18.0 h, which does not differ significantly from that of So93 (17.0 h).

In general, rough colonies were larger than smooth ones. By SEM, the difference in surface texture between the two variants was clearly visible at low magnification. Smooth colonies varied from small, singular, flat, cone-shaped to larger compound colonies with a smooth surface, whereas rough colonies showed a tubular surface (data not shown). At higher magnification, individual So93 cells of the smooth variant were arranged chaotically, whereas cells of the rough variant were oriented in compact parallel arrangements. This characteristic difference in cellular orientation was also found by SEM analysis of cross sections of colonies (data not shown). Negative-stained TEM preparations showed that the cells of the rough variant were usually present as large cord-like aggregates, whereas So93 bacteria were aggregated in small lumps.

Ultrathin sections revealed no differences in ultrastructure between the two variants and the control strain, H37Rv (data not shown).
**Characterization of housekeeping genes.** The sequence of the 5' 550 bp of the 16S rRNA gene of strain So93 was determined, and it was identical to published 16S rRNA gene sequences of *M. tuberculosis* strains. On the basis of single-strand conformation polymorphism analysis (26), So93 did not differ in the sequence of the housekeeping genes *rpoB, katG, rpsL*, and *gyrA* of So93 from either H37Rv or *M. bovis* BCG (data not shown).

**Genetic characterization by polymorphic genetic markers.** IS6110 probing showed the presence of two IS6110 copies in the genome of both So93 variants (Fig. 3A). The IS6110 fingerprint of So93 was compared with over 5,000 fingerprints in the database in our laboratory, which contains fingerprints of approximately 3,500 *M. tuberculosis* complex strains isolated in The Netherlands during the last 3 years (40) and 1,500 isolates from over 30 different countries. None of these isolates matched the fingerprint of So93, indicating that this *M. tuberculosis* strain had not been previously encountered among the clinical isolates during the last 4 years in The Netherlands and that this type of strain is also uncommon in many other countries.

Previously, it was shown that the insertion element IS1081 is invariably present in *M. tuberculosis* complex bacteria and that the number of IS1081 copies varies from four to seven (44). However, strain So93 and its rough variant, So93R, were found to carry only a single copy of IS1081, which in both strains was on a similar-size *Pvu*II fragment (Fig. 3B). In Bilthoven (RIVM), over 2,000 clinical *M. tuberculosis* complex strains have been analyzed so far, and none of these strains carried less than four copies of this insertion element (9). Therefore, the finding of a single IS1081 copy in So93 is highly exceptional.

A novel method of strain differentiation, spoligotyping, was recently developed (20, 45). With this method, the presence of 43 short spacer DNA sequences in the polymorphic DR locus of *M. tuberculosis* complex bacteria was investigated. Both So93 and So93R were found to share only two of the 43 spacer DNA sequences derived from the DR locus sequence of *M. tuberculosis* H37Rv and *M. bovis* BCG vaccine strain P3 (Fig. 4). The presence of only two known spacer sequences in the DR locus is highly exceptional and has previously been encountered only among four isolates of *M. microti*; however, these spacers were distinct from the ones of So93. The spoligotype of So93 was not observed previously among approxi-
FIG. 3. IS6110 (A) and IS1081 (B) hybridization patterns of PvuII-digested genomic DNA of So93 (lane 3), So93R (lane 4), and the Canetti strain (lane 5). For comparison, clinical M. tuberculosis isolates myc 96-1314 (lane 1), myc 96-166 (lane 2), and myc 96-1118 (lane 6) were included.

mately 1,500 M. tuberculosis complex strains investigated in Bilthoven (RIVM) (9).

Other polymorphic repetitive elements in M. tuberculosis complex bacteria are the PGRS (31) and the MPTR (19). Southern blotting showed the presence of multiple copies of PGRS, and MPTR DNA sequences also appeared to be present in So93. Furthermore, the RFLP patterns of the So93 variants were invariably most distantly related compared to other M. tuberculosis complex isolates tested (data not shown). No differences between So93 and other M. tuberculosis complex strains were observed in Southern blot analysis when the housekeeping genes katG, 16s rRNA, and groEL were used as probes.

In the past, the exceptional cell wall content of a M. tuberculosis isolate growing in smooth colonies and isolated by Canetti in France was studied extensively (13). Therefore, we tested the Canetti strain with all of the markers described above. Surprisingly, this strain was found to be highly similar to So93 by all of these polymorphic genetic markers, except for the IS1081 and IS6110 fingerprints. Like So93, the Canetti strain carried a single copy of IS1081. However, the apparent size of the IS1081-containing PvuII fragment of the Canetti strain was 4.4 kb, instead of 2.8 kb in the case of So93. Sequence analysis of the chromosomal region bordering IS1081 in So93 and the Canetti strain showed that IS1081 is inserted at exactly the same chromosomal position in both strains. The difference in the size of the IS1081-containing PvuII restriction fragments was shown to be due to modification of a PvuII site in the Canetti strain, as described previously for other PvuII sites (40). Because this modified PvuII restriction site is not cleaved in the Canetti strain, the IS1081-containing PvuII fragment is larger than that of So93.

The Canetti strain was found to carry three IS6110 copies, and two of the insertion sequence-containing PvuII fragments matched the IS6110-containing PvuII fragments of So93 (Fig. 3), indicating an extra copy of IS6110 in the Canetti strain.

In spoligotyping, the Canetti strain reacted with the same two spacers as So93 and So93R, indicating the same DR region composition.

Van Steenbergen et al. recently used a high-resolution RFLP technique for bacterial strain typing, RFEL (46). This technique was applied to compare So93 with 10 other M. tuberculosis complex isolates, including all (sub)species of this group of bacteria (Fig. 5). The RFEL technique allowed us to distinguish 51 restriction fragments in the molecular size range from 160 to 400 bp. So93 shared 45 of these with the other M. tuberculosis isolates. The other isolates shared at least 48 of the 51 fragments (Fig. 5). These results indicate that So93 is ge-
Variations in the recA gene. One of the IS1081 elements is located in the vicinity of the recA gene in the genome of M. tuberculosis complex strains (11). Because this IS1081 copy is missing in So93, the presence of the recA gene in So93 was examined with a PCR with primers based on the recA gene sequence of H37Rv. In restriction endonuclease analyses of the similar-size PCR product, it became clear that the recA gene of So93 differs from that of other M. tuberculosis strains. Therefore, the sequences of the entire recA gene of the So93 isolate and of the Canetti strain were determined (EMBL accession no. AJ000011 and AJ000012, respectively). The sequence was found to differ in 18 nucleotides from the published sequence of H37Rv (Fig. 6). Fourteen of these substitutions were in the intein region of the recA gene, and 4 resided in the 5' extein part. The mutations in the intein region resulted in six amino acid substitutions, whereas the rest of the mutations were synonymous. The highest degree of variation was found in the 91-bp stretch from positions 1594 to 1625 of the intein of recA. Comparison of the recA gene region of the Canetti strain with the equivalents in So93 and H37Rv revealed that all mutations observed in So93 were also present in the Canetti M. tuberculosis strain. In addition, two mutations were found at positions 16 and 1165. The first mutation was found upstream from the recA gene in a noncoding area. The second mutation appeared to be a synonymous mutation in the intein region of recA. The recA sequence was also determined in three M. bovis strains and one M. microti isolate. These were identical to recA of H37Rv. Thus, the deviations in the recA gene in So93 and the Canetti strain seem to be restricted to this particular genotype of the M. tuberculosis complex (data not shown).

Two dimensional SDS-PAGE whole-protein analysis. In order to compare the protein expressions of So93 and So93R with those of other M. tuberculosis complex strains, whole-protein extracts of these strains were analyzed by two-dimensional SDS-PAGE (data not shown). The polypeptide patterns...
were compared with those of H37Rv and \textit{M. tuberculosis} clinical isolate myc 94-2165. Computer-assisted analysis of the protein profiles identified a total of 686 distinct protein moieties, 78\% of which were shared by all four strains. The similarity coefficients between the different protein profiles are summarized in Table 2. So93 shared 86 and 84\% of its protein spots with H37Rv and the clinical \textit{M. tuberculosis} isolate, respectively, whereas the latter two strains shared 90\% of the spots. This suggests that strain So93 is more distantly related to H37Rv and the clinical \textit{M. tuberculosis} isolate than the latter two strains are related mutually. The rough variant So93R differed from its parental strain in 20 protein spots (3\%).

**Cell wall glycolipid composition.** The apolar and polar lipid populations were separately extracted from the So93 and So93R strains and compared to each other, initially by TLC. Two-dimensional TLC of the apolar lipids, first in chloroform-methanol (96:4) and then in toluene-acetone (80:20), showed no obvious difference. Both glycolipid profiles were dominated by the presence of the triglycosyl-containing phenolic glycolipid characteristic of the Canetti strain (13). However, TLC in chloroform-methanol-water (65:25:4) of the polar lipids from the smooth So93 strain showed the presence only of the multiglycosyl LOSs, also characteristic of the Canetti strain (12); these were absent from So93R. To provide unequivocal evidence of this dramatic qualitative difference between So93 and So93R and to generate further evidence of structure, the apolar and polar lipid populations from both strains were hydrolyzed, and alditol acetates were generated and analyzed by GC and GC-mass spectrometry (Table 2). The structure of the triglycosyl unit of the unique phenolic glycolipid of the Canetti strain was 2,3,4-tri-O-methyl-L-fucopyranosyl-(1\,\rightarrow\,3)-L-rhamnopyranosyl-(1\,\rightarrow\,2)-O-methyl-L-rhamnopyranosyl (113) characteristic of the Canetti strain (112); GC demonstrated the presence of these three sugars in the apolar glycolipid (Table 2). However, examination of the polar lipid fractions demonstrated the presence of the sugars characteristic of the major LOSs of the Canetti strain (2-O-methyl-L-fucopyranosyl, 2-O-methyl-rhamnopyranosyl, and 4-O-methyl-L-rhamnopyranosyl) only in the case of the smooth So93 strain (Table 3). The identities of all of these 6-deoxyhexoses were confirmed by mass spectrometry as described previously (12).

**Virulence of the strains So93, So93R, and H37Rv.** Guinea pigs were inoculated with cultures of So93, So93R, and H37Rv (two animals for each strain). None of the animals showed signs of clinical disease when autopsied 8 weeks after inoculation. Overwhelming disseminated tuberculous lesions were found at necropsy only in the guinea pigs inoculated with So93 and So93R. In addition, these animals showed severe loss of body fat depots. In the two H37Rv-inoculated animals, macroscopic lesions were confined to the inoculation site and/or draining lymph nodes, and no decrease of fat deposits was observed. Microscopic and/or bacteriologic examination showed infection of livers, spleens, and lungs of all six animals (Table 4). A remarkable gross finding in all six animals was severe enlargement of the left iliac nodes in H37Rv-inoculated animals without obvious enlargement of inguinal node groups. In the So93R-inoculated guinea pigs, involvement of spleen, lungs, iliac, and bronchial nodes (data not shown) was more extensive than in the animals infected with the parental strain So93. The average root index of virulence (25), as calculated from the macroscopic lesions, amounted to 0.77 for So93, 1.03 for So93R, and 0.27 for H37Rv, the lower and upper limits being 0 and 1.34, respectively, for 8 weeks of survival. This suggests that the Somali strain, So93, and its rough derivative, So93R, are more virulent than H37Rv in the guinea pig and that So93R is slightly more virulent than So93. Microscopically, the tuberculous epithelioid granulomas in all six animals were essentially the same (data not shown). Caseous necrosis with calcification was restricted to the primary complex and other severely tuberculous organs. Mycobacteria were easily observed after auramine-rhodamine staining (data not shown), whereas in serial tissue sections, a few bacilli were found only by the Ziehl-Neelsen staining method. The numbers of CFU cultured from tissue specimens of the guinea pigs did not differ significantly among animals infected with the different strains. All colonies grown from different tissues from guinea pigs infected with So93R were rough. In contrast, 10 to 30\% of the colonies grown from the spleens and lymph nodes from the So93-infected guinea pigs were rough (Fig. 2B). No such variation was observed among colonies grown from liver tissue of So93-infected guinea pigs.

**DISCUSSION**

In this study, we used novel genetic markers which have not been used earlier to classify different species of the \textit{M. tuberculosis} complex. Although this study was mainly focused on the characterization and classification of the two mycobacterial isolates with an unusual smooth colony morphology, the data for both the \textit{M. tuberculosis} complex group of bacteria and various non-\textit{M. tuberculosis} complex species allow us to draw more general conclusions about the taxonomic classification of the \textit{M. tuberculosis} complex bacteria. DNA sequence analyses of the 16S rRNA, \textit{popB}, \textit{katG}, \textit{psl}, and \textit{gyrA} genes reveal a very strong identity among bacteria of the \textit{M. tuberculosis} complex, including the strains So93 and Canetti. Furthermore, the results obtained by RFEL and two-dimensional whole-cell protein electrophoresis also showed a high degree of similarity among isolates of the complex and a distinction from non-\textit{M. tuberculosis} complex bacteria, such as \textit{M. avium} and Mycobac-

---

**TABLE 2.** Comparison of the whole-cell protein profiles of strains So93, So93R, H37Rv, and myc 94-2165 by the similarity coefficient of Jaccard according to spot positions

<table>
<thead>
<tr>
<th>Strain</th>
<th>So93</th>
<th>So93R</th>
<th>H37Rv</th>
<th>myc 94-2165</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Similarity to strain</td>
<td>97.0</td>
<td>93.3</td>
<td>85.7</td>
<td>84.0</td>
</tr>
<tr>
<td>So93</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>So93R</td>
<td>100</td>
<td>85.0</td>
<td>83.3</td>
<td></td>
</tr>
<tr>
<td>H37Rv</td>
<td>100</td>
<td></td>
<td>89.5</td>
<td></td>
</tr>
<tr>
<td>myc 94-2165</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 3.** Sugar composition of the apolar and polar lipids of So93 and So93R as determined by GC

<table>
<thead>
<tr>
<th>Strain</th>
<th>Lipid fraction</th>
<th>Sugar composition</th>
<th>Nature of lipid</th>
</tr>
</thead>
<tbody>
<tr>
<td>So93</td>
<td>Apolar</td>
<td>2,3,4-Tri-O-Me-Fuc</td>
<td>Phenolic glycolipid-Tbl</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2-O-Me-Rha</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polar</td>
<td>2-O-Me-Fuc</td>
<td>2-O-Me-Rha</td>
<td>LOS</td>
</tr>
<tr>
<td></td>
<td>4-O-Me-Rha</td>
<td></td>
<td></td>
</tr>
<tr>
<td>So93R</td>
<td>Apolar</td>
<td>2,3,4-Tri-O-Me-Fuc</td>
<td>Phenolic glycolipid-Tbl</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2-O-Me-Rha</td>
<td></td>
</tr>
<tr>
<td>Polar</td>
<td>No characteristic</td>
<td>6-deoxyhexoses</td>
<td></td>
</tr>
</tbody>
</table>
terium kansasii. Finally, we showed, that in contrast to other mycobacterial species, all groupings belonging to the *M. tuberculosis* complex (including So93 and the Canetti strain) share the presence of the repetitive DNA elements such as IS6110 and DR. Based on these data, *M. tuberculosis*, *M. bovis*, *M. africanum*, So93, and the Canetti strain seem to constitute a very closely related grouping of mycobacteria, clearly distinct from other mycobacterial species. This idea is consistent with the view that all groupings belonging to the *M. tuberculosis* complex should be regarded as a single bacterial species (39, 47) and that the previously established species *M. tuberculosis*, *M. bovis*, *M. africanum*, and *M. microti* are subspecies of this single species, which likely will be designated "*M. tuberculosis."*

In due time, we will propose a revision of the nomenclature of the taxa within the *M. tuberculosis* complex group of bacteria.

Smooth strains of *M. tuberculosis* show up rarely in diagnostic laboratories, and such strains usually are not studied further. Strain So93 attracted our attention because of its extremely smooth and glossy appearance, while its other phenotypic properties corresponded with those of *M. tuberculosis* complex bacteria. At the mycobacterial reference laboratory in Bilthoven, this was the first *M. tuberculosis* strain encountered among over 10,000 isolates investigated in the past 10 years with this distinct morphotype. To our knowledge, the only other similarly smooth *M. tuberculosis* isolate described is the Canetti strain, isolated in France from a native Frenchman in 1969 and used in extensive studies in search of antigens for immunodiagnostic tests (12).

Although So93 clearly belongs to the *M. tuberculosis* complex, the strain is distinct from other (sub)species in this group of taxonomically closely related bacteria. First, So93 is highly exceptional in harboring only a single copy of the insertion sequence IS1081. All of the nearly 2,000 *M. tuberculosis* complex isolates examined in Bilthoven (RIVM), including about 100 *M. bovis* and 10 *M. microti* strains, carried at least four IS1081 copies, usually located at the same PvuII restriction fragments (43). Genomic mapping of strain H37Rv has shown that each of the IS1081 copies is located at widely different loci of the *M. tuberculosis* chromosome (11). This means that the absence of all but one of the IS1081 copies in So93 cannot be attributed to a single genomic deletion. This is consistent with preliminary observations suggesting that the loci in *M. tuberculosis* strain H37Rv where IS1081 elements are usually integrated lack any sequence similarity over stretches of more than 10 kb (9).

Second, So93 was found to have a DR locus which is unique within the *M. tuberculosis* complex. So far, the DR locus has been found only in *M. tuberculosis* complex bacteria, and it is highly polymorphic because of variation in the number of short DRs and nonrepetitive spacer sequences. So93 was found to carry only two of the currently characterized spacers which are used in spoligotyping. Virtually all of the approximately 2,000 other *M. tuberculosis* complex strains examined so far share at least 8 of their spacer sequences with any of the set of 43 spacer sequences which have been derived from *M. tuberculosis* H37Rv and *M. bovis BCG* (20). Most *M. microti* strains share as few as two spacers with those of H37Rv and *M. bovis BCG*; however, the *M. microti* spacers differ from those found in So93 (9). It is unlikely that the polymorphism in the DR locus of strain So93 resulted from a few mutations within the DR region of any of the known *M. tuberculosis* strains, because the DR locus of strain So93 contains over 40 spacers (9), most of which are different from those of H37Rv and *M. bovis BCG* (9). This reinforces the idea that So93 is a distinctly related member within the *M. tuberculosis* complex. It should be noted that large DNA fingerprint databases, such as those of IS6110, IS1081, and the spoligotypes, appear to be extremely useful in mycobacterial systematics. Strain So93 was displayed as an unusual type in these three databases, and this greatly stimulated interest in this strain.

Third, the recA gene of So93 was found to differ significantly from those of other *M. tuberculosis* complex strains. So93 and the Canetti strain differed in 18 and 20 residues in the 2.7-kb recA sequence, respectively; most of these were synonymous mutations in a restricted area of the recA intein. No DNA sequence variation was found among strains of *M. tuberculosis*, *M. bovis*, and *M. microti*. This high degree of variation in recA of strain So93 and Canetti is highly exceptional. Musser et al. (21) found only four synonymous nucleotide substitutions among 240,000 bases in eight housekeeping genes among 30 *M. tuberculosis* isolates obtained from widely different geographic origins. Because the Canetti strain was found to be nearly

---

**TABLE 4. Virulence of So93 and So93R in guinea pigs: results of pathology and bacteriology**

<table>
<thead>
<tr>
<th>Site</th>
<th>So93</th>
<th>So93R</th>
<th>H37Rv</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10^3</td>
<td>10^3</td>
<td>10^3</td>
</tr>
<tr>
<td>Inoculation</td>
<td>+ + 5.00</td>
<td>+ + 4.88</td>
<td>+ + 5.85</td>
</tr>
<tr>
<td>Lymph station</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>++ + 4.60</td>
<td>++ + 4.60</td>
<td>+ + 5.85</td>
</tr>
<tr>
<td>2</td>
<td>++ + 4.88</td>
<td>++ + 4.70</td>
<td>++ + 5.54</td>
</tr>
<tr>
<td>3</td>
<td>++ + ND</td>
<td>++ + ND</td>
<td>++ + ND</td>
</tr>
<tr>
<td>Lungs</td>
<td>− + ND</td>
<td>+ + ND</td>
<td>+ + 4.54</td>
</tr>
<tr>
<td>Liver</td>
<td>+ + 2.48</td>
<td>+ + 2.70</td>
<td>+ + 2.78</td>
</tr>
<tr>
<td>Spleen</td>
<td>++ + 4.48</td>
<td>++ + 4.00</td>
<td>++ + 4.30</td>
</tr>
</tbody>
</table>

*°* Gross, gross extent of tuberculosis: -, absent; +, mild to moderate; ++, marked; ++++, severe; NF, organ not found.

Micro, microscopical confirmation of tuberculosis: -, no granuloma; ±, minimal epithelioid granuloma; +, clear granuloma formation.

CFU, colony-forming unit on 7H10 per g of selected tissue specimens (4 weeks of growth); ND, not done; NG, not grown.

* Lymph station 1, left deep and superficial inguinal lymph nodes; 2, left iliac lymph nodes; 3, lumbar aortie and celiac lymph nodes.

* Spleen of this animal was grossly normal, except for one 2-mm white nodule that was included in toto in the tissue for bacterial enumeration.
identical to So93 with respect to its distinct genetic properties, we propose to designate the two strains as a new (sub)species, "Mycobacterium tuberculosis subsp. Canetti," or "M. canetti."

As previously observed for the Canetti strain, strain So93 grew with smooth colonies. So93 switched to variants with a rough colony morphology. Reversion was not observed, even after passage in the guinea pig. The preliminary data show that transition frequency from smooth to rough was higher in vivo than in vitro, suggesting that the rough variant is better adapted to in vivo conditions. Both colony variants were about equally virulent in the guinea pig model.

The transition from smooth to rough was associated with disappearance of the LOS component. The morphotypes also differed in the expression of approximately 20 polypeptides as measured by two-dimensional SDS-PAGE. This indicates that the colony variation probably results from an alteration in the expression of a large number of genes.

Analysis of the glycolipids of So93 and So93R showed the presence of triglycosyl phenol phthiocerol glycolipid (phenolic glycolipid) with the triglycosyl unit [2,3,4-tri-O-methyl-1-fuco-pyranosyl-(1→1)-rhamnopyranosyl-(1→3)-2-O-methyl-L-rhamnopyranosyl-(1→3)] characteristic of the Canetti strain. In addition, the smooth So93 strain, but not the So93R equivalent, contained the characteristic LOS of the Canetti strain, distinguished by the presence, among other sugars, of 2-O-Me-Fuc, 2-O-Me-Rha, and 4-O-Me-Rha.

No genetic rearrangement associated with colony variation was observed by Southern blot analyses with various probes. Therefore, the genetic mechanism underlying the phase variation remains unknown. Variation in colony morphology is more common among non-M. tuberculosis complex mycobacteria. In M. avium, the transition from smooth to rough colonies was found to be associated with the loss of a group of glycolipids known as the glycopeptidolipids due to chromosomal deletions (4). Perhaps of more direct relevance to the present So93 versus So93R switch is the example of M. kansasii. All known isolates of M. kansasii, whether of rough or smooth phenotype, contain an M. kansasii-specific phenolic glycolipid (3). However, only the smooth varieties contain the M. kansasii version of the LOSs. The genetic basis of this switch has not been examined.

Presently, So93 and the Canetti strain are the only known isolates of the Canetti type, and therefore the incidence of the Canetti type in human tuberculosis patients is probably low, although the incidence may have been underestimated because of misdiagnosis due to the unusual colony morphology. Strain So93 was isolated in 1993 from a Somali child, who probably contracted the infection in Somalia or in a refugee camp in Kenya, where the child resided before coming to The Netherlands. The reservoir of M. tuberculosis subsp. Canetti is presently unknown. Perhaps animals constitute the natural host, which may explain the sporadic occurrence among humans. The detailed genetic characterization of the Canetti type of M. tuberculosis in this study enables us to easily recognize this subspecies by a variety of techniques, including PCR-based methods. This will facilitate future studies of the occurrence and host range of this novel (sub)group of the M. tuberculosis complex.

ACKNOWLEDGMENTS

We thank Marcel Sluijter, Joke Pijnenburg, Mirjam Dessens, Kristin Kremer, Annelies Bunschoten, Marcel Menkhorst, and Gerard van Leeuwen for technical assistance. Nicole Knetemann is acknowledged for secretarial assistance. S. Cole is acknowledged for performing single-strand conformation polymorphism analysis of housekeeping genes. We furthermore thank Herre Heersma and Nico Nagelkerke for statistical analysis.

This work was supported by the World Health Organization Programme for Vaccine Development, the European Community Programme for Science Technology and Development and Biomed 1, the Dutch Foundation for Technical Sciences, and the National Institute for Allergy and Infectious Diseases (NIH; grants AI 18357 and 35220). This work was also partly supported by the Diemien Foundation (Brussels).

REFERENCES


23. Leif-Frebault, V., and F. Portaels. 1902. Proposed minimal standards for the genus Mycobacterium and for description of slow growing Mycobacterium...


48. Wells, A. Q. 1945. The murine type of tubercle bacillus. Medical Research Council, Sir William Dunn School of Pathology, University of Oxford, Oxford, United Kingdom. (Special report series no. 259.)