The 16S rRNA nucleotide sequence of *Mycobacterium leprae*: phylogenetic position and development of DNA probes

Andreas Teske 1, Jörn Wolters 2 and Erik C. Böttger 1

1 Institut für Medizinische Mikrobiologie, Medizinische Hochschule Hannover, Hannover, and 2 Institut für Allgemeine Mikrobiologie, Christian-Albrechts-Universität, Kiel, F.R.G.

Received 25 January 1991
Accepted 5 February 1991

Key words: *Mycobacterium leprae*; 16S rRNA; Oligonucleotide probe; Phylogeny

1. SUMMARY

The almost complete 16S rRNA sequence from *Mycobacterium leprae* was determined by direct sequencing of the chromosomal gene amplified by the polymerase chain reaction. The primary sequence revealed an insertion of 12 nucleotides at the 5' end of the 16S rRNA gene, which consists of an A-T stretch and appears to be unique for *M. leprae*. Within the mycobacteria *M. leprae* branches off with a group of slow-growing species comprising *M. scrofulaceum*, *M. kansasii*, *M. szulgai*, *M. malmoense*, *M. intracellulare* and *M. avium*. A systematic comparison of the nucleotide sequence resulted in the characterization of oligonucleotide probes which are highly specific for *M. leprae*. The probes hybridized exclusively to 16S rRNA nucleic acids from *M. leprae*, but not to nucleic acids from 20 cultivable fast- and slow-growing mycobacteria.

2. INTRODUCTION

*Mycobacterium leprae*, the causative agent of leprosy, is a human pathogen which is poorly understood, largely due to the inability of the organism to grow in conventional media or tissue culture cell lines. The inability to cultivate *M. leprae* makes the phenotypic identification of this organism difficult. However, specific DNA probes have been developed and successfully used for the identification of a number of bacterial pathogens [1,2]. We have focused on small subunit ribosomal RNA, because it is present in several thousand copies in each bacterium [3] and thus facilitates the development of nucleic acid-based detection assays [4]. In addition, the high copy number of rRNA offers an opportunity which is unique among DNA probes targeted at bacterial nucleic acids, namely in situ hybridization of whole cells [5-7]. More recently, a polymerase chain reaction [8] mediated amplification of rDNA/rRNA sequences has been demonstrated to be an extremely sensitive tool for the detection of mycobacteria [9].

So far, only partial fragments of the 16S rRNA sequence from *M. leprae* have been published
[10,11], which were obtained by the reverse transcriptase technique. These partial sequences show several nucleotide differences as a result of the inaccuracy of the sequencing method employed [12,13]. Using this limited sequence information, however, the taxonomic placement of M. leprae within the genus Mycobacterium was demonstrated and this species was shown to belong phylogenetically to slow-growing mycobacteria [11].

We have recently developed a general and highly accurate sequencing procedure, in which the polymerase chain reaction is used [14], for the isolation and direct nucleotide determination of entire genes coding for 16S rRNA [15,16]. Phylogenetically, as determined by 16S rRNA comparative sequencing, mycobacteria have been reported to represent a very homogenous genus, characterized by a 16S rRNA sequence similarity greater than 94% [17-19]. Furthermore, some mycobacterial species differ only by a single nucleotide in their 16S rRNA sequence [19]. Given the limitations of the reverse transcriptase sequencing technique, the numerous differences between the published partial 16S rRNA sequences from M. leprae [10,11] and the phylogenetic homogeneity within the genus Mycobacterium, which is reflected at the high 16S rRNA sequence similarity within this taxon, it was thus necessary to determine the 16S rRNA primary structure from M. leprae more accurately for determination of its intrageneric taxonomic position, as well as for the development of specific DNA probes.

3. MATERIALS AND METHODS

M. leprae cells were kindly provided by J. Gerdes, Forschungsinstitut Borstel, F.R.G.

Bacterial cells were lysed using a lysozyme, proteinase K, sodium dodecylsulfate procedure and nucleic acids extracted with phenol [9]. The 16S rRNA gene was amplified by the polymerase chain reaction and the sequence determined as described previously [16].

For the hybridization analysis, the 16S rRNA gene from the different mycobacterial species was amplified using primers 246 AGA GTT TGA TCC TGG CTC AG (corresponding to E. coli positions 8-28) and reverse 264 TGC ACA CAG GCC ACA AGG GA (corresponding to E. coli positions 1046-1027). These primers direct the synthesis of a 1030-bp gene fragment. Agarose gel electrophoresis was used to confirm the success of the amplification. Approximately 2 µg of the amplified fragment was obtained (data not shown). One-tenth of the amplified sample was denatured by the addition of 100 µl of 0.5 M NaOH-25mM EDTA, incubated for 30 min at room temperature, and neutralized with 120 µl of 2 M ammonium acetate. The samples were loaded into wells of a manifold (Minifold II; Schleicher & Schuell, Dassel, F.R.G.) fitted with a nitrocellulose membrane (BA 85; Schleicher & Schuell) previously wetted in 10 × SSC (1 × SSC is 0.15 M NaCl plus 0.015 M sodium citrate, pH 7.0). Each well was rinsed with 0.4 ml of 10 × SSC, and the membranes were heated for 2 h at 80°C. Hybridization was performed by standard techniques. The membranes were rinsed in 6 × SSC and prehybridized in 6 × SSC, 5 × Denhardt solution (1 × Denhardt solution is 0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin), 0.1% sodium dodecyl sulfate, and 100 µg of tRNA (Boehringer) ml⁻¹ for 1 to 2 h at 50°C. Hybridization was performed in the same solution containing 2 × 10⁶ to 1 × 10⁷ cpm of 5'-end-labelled oligonucleotide ml⁻¹ overnight. The blots were washed three times for 20 min each at room temperature in 6 × SSC and three times for 20 min each in 6 × SSC-0.1% sodium dodecyl sulfate at the appropriate temperature given by the Tm (°C) of the oligonucleotide. The blots were exposed to XAR-5 films (Eastman Kodak Co., Rochester, NY) for 2 h at room temperature.

Oligonucleotides were labelled in 15 to 20 µl of 50 mM Tris (pH 7.6), 10 mM MgCl₂, 5 mM dithiothreitol, 0.1 mM spermidine, and 0.1 mM EDTA with 8 pmol of synthetic oligonucleotide primer, 100 µCi of [γ-³²P]ATP (specific activity > 5000 Ci/mmol; Amersham, Braunschweig, F.R.G.), and 10 U of T4 polynucleotide kinase (New England BioLabs, Schwalbach, F.R.G.) by incubation for 1 h at 37°C. The kinase was inactivated by incubation for 10 min at 75°C; the nucleotides were extracted with phenol, and the
labelled oligonucleotide was purified and separated from residual \([\gamma^32P]ATP\) using push columns (Stratagene, Heidelberg, F.R.G.). The resulting specific activity of the labelled oligonucleotide was \(2 \times 10^8\) to \(5 \times 10^8\) cpm/\(\mu\)g.

16S rRNA sequences were aligned using the multisequence alignment of Krüger and Osterburg [20]. For the phylogenetic analysis, regions of alignment uncertainty were omitted, reducing the number of positions from 1448 to 1400. Pairwise distances were calculated by weighting nucleotide differences and insertions–deletions equally (Hamming distance, data not shown). The phylogenetic tree was constructed by using the neighbourliness method [21].

4. RESULTS AND DISCUSSION

The sequencing strategy used resulted in the determination of a contiguous stretch of 1448 nucleotide positions, covering 95% of the 16S rRNA gene (see Fig. 1; sequence accession number EMBL X55022). Comparison with published fragmentary rRNA sequences [10,11] revealed several sequencing errors therein which have to be corrected and are most likely due to the known problems of the reverse transcriptase technique used.

The obtained sequence was aligned with published 16S rRNA sequences from 12 different mycobacterial species [19] covering most slow-growing mycobacteria [22] and the homologous regions were used for the phylogenetic analysis. Even though the primary sequence of the 16S rRNA is known for its conserved structure [3], numerous differences in its nucleotide sequence allowed us to characterize the taxonomic position of \(M. \text{leprae}\) in detail and to elucidate its relationships. The tree in Fig. 2 reveals that \(M. \text{leprae}\) branches off within the slow-growing mycobacteria by showing a common ancestry with \(M. \text{scrofulaceum}, M. \text{kansasii}, M. \text{szulgai}, M. \text{malmoense}, M. \text{intracellulare}\) and \(M. \text{avium}\) (interspecies sequence homology > 97.5%), while \(M. \text{tuberculosis}\), the other common human pathogen, forms a slightly more ancient line of descent.

In a first attempt to determine the phylogenetic position of \(M. \text{leprae}\) by 16S rRNA comparative sequencing, this species has been characterized as...
phylogenetically closely related to slow-growing mycobacteria [11]. However, the relatively high level of sequencing errors of the reverse transcriptase sequencing method, which has been estimated to be some 2–5% [12,13], may cause organisms to branch off at positions which do not necessarily reflect their actual phylogenetic position. This precludes the determination of relationships characterized by a high level of sequence similarity, where the actual sequence difference may equal the sequencing error of the method used. In addition, phylogenetic trees are dynamic, their topography changing when a more balanced selection of investigated organisms is used. The majority of reference strains used in this study have not been analysed by the reverse transcriptase sequencing method. The reinvestigation of the 16S rRNA sequence by a more advanced and accurate procedure has allowed us to define the phylogenetic position of \textit{M. leprae} more precisely and to place this organism within the confines of \textit{M. tuberculosis}, a pathogen, and a group of mycobacteria consisting of \textit{M. scrofulaceum}, \textit{M. kansasii}, \textit{M. szulgai}, \textit{M. malmoense}, \textit{M. intracellulare} and \textit{M. avium} which are characterized as occasionally pathogenic species [22].

![Fig. 2. Phylogenetic tree showing the relationship of \textit{M. leprae} to other slow-growing mycobacteria. The tree was constructed using the neighbourliness method [21] and rooted by using \textit{M. smegmatis} as an outgroup. The bar indicates 10 nucleotide differences.](image)

![Fig. 3. Secondary structure of the 16S rRNA of \textit{M. leprae} and \textit{M. bovis}. \textit{M. leprae} is characterized by an insertion of 12 nucleotides resulting in an extended helix 6 in variable region VI.](image)

determination of the 16S rRNA primary structure has resolved the problems concerning the relatedness of this organism as DNA-DNA hybridization data gave conflicting results [23,24].

When compared with the 16S rRNA sequences of other mycobacterial species [19], the 16S rRNA primary structure of \textit{M. leprae} reveals an insertion of several nucleotides at positions 81 and 89. This insertion is unique for \textit{M. leprae} among the mycobacteria investigated and results in an extended helix 6 in variable region VI [25] as shown in Fig. 3.

For the development of nucleic acid probes specific for \textit{M. leprae} we have focused on two segments (indicated in Fig. 1), which are contained in a region, the primary structure of which is known to be variable in eubacteria [3], i.e. position 180–240 of the \textit{E. coli} numbering system [26]. Comparison with published mycobacterial 16S rRNA sequences [19,27] reveals that this region contains sequence stretches which are unique for \textit{M. leprae}. A 20- and 22-mer probe were chosen which differ by at least two or more bases from other mycobacteria, as shown by an alignment of the homologous regions in the 16S rRNA molecule in Fig. 4. The probe sites were designed to place the mismatches between mycobacterial sequences as near the center of the segment as possible. The specificity of the oligonucleotides was tested against amplified nucleic acids from 20
slow- and fast-growing mycobacterial species, including *M. avium*, *M. intracellulare*, *M. scrofulaceum*, *M. kansasii*, *M. terrae*, *M. gastri*, *M. tuberculosis*, *M. fortuitum*, *M. flavescens*, *M. marinum*, *M. nonchromogenicum*, *M. xenopi*, *M. malmoense*, *M. chelonae*, *M. cookii*, *M. sphagni*, *M. komossense* and *M. smegmatis*. As demonstrated in Fig. 5 both probes, i.e. oligonucleotides 418 and 419, show exclusive hybridization to *M. leprae*. The region targeted by oligonucleotide 419 has been described previously in part as useful for development of a DNA probe specific for *M. leprae*, although the published probe sequence [11] does not correspond to the actual 16S rRNA sequence of *M. leprae* by showing two nucleotide mismatches.

Specific nucleic acid probes have been suggested for the detection of *M. leprae* [28]. More recently, polymerase chain reaction based detection methods for this pathogen have been described [29-33]. However, oligonucleotide probes directed at specific sequences in bacterial ribosomal RNA promise to be especially useful for the rapid identification of bacteria and the 16S rRNA probes described here offer a dual advantage. Following general amplification of mycobacterial 16S rDNA/rRNA sequences by the polymerase chain reaction technique, such species-specific probes can be used for a highly sensitive and species-specific detection of mycobacteria as has been demonstrated recently [9]. Furthermore, due to the high abundancy of rRNA molecules, such probes offer the unique advantage of a direct in situ staining [5-7].

ACKNOWLEDGEMENTS

The authors thank J. Gerdes for providing *M. leprae* cells, A. Ohrdorf for expert technical assistance, S. Maibom for typing the manuscript, A. Rosenmeier for the artwork and D. Bitter-Suer-
Fig. 5. Hybridization of amplified nucleic acids from the indicated bacterial species to *M. leprae*-specific oligonucleotide probes. (A) Sequence ATA GGA CTG CAA GGC GCA TG, hybridization was performed at 56°C, washings at 59°C; (B) Sequence AAG CTG TGT GCG GTG CAG GAT G, hybridization was performed at 56°C, washings at 64°C.

REFERENCES


mann for continuous encouragement. This work was supported by a grant from the Bundesministerium für Forschung und Technologie (OI KI 89117) to E.C. Böttger. J. Wolters was supported by a grant from the Deutsche Forschungsgemeinschaft.