Volume 281, number 1,2, 114–118 FEBS 09567 April 1991

© 1991 Federation of European Biochemical Societies 00145793/91/$3.50

ADONIS 001457939100277S

Complete nucleotide sequence of the *Mycobacterium leprae* 23 S and 5 S rRNA genes plus flanking regions and their potential in designing diagnostic oligonucleotide probes

W. Liesack[1], S. Sela[2], H. Bercovier[2], C. Pitulle[3] and E. Stackebrandt[1]

[1]*Department of Microbiology, University of Queensland, St. Lucia, Australia 4072,* [2]*Department of Clinical Microbiology, The Hebrew University, Medical School, Jerusalem 91010, Israel* and [3]*Institut für Allgemeine Mikrobiologie, Christian-Albrechts-Universität, 2300 Kiel, Germany*

Received 30 January 1991

The complete nucleotide sequences of the *Mycobacterium leprae* 23 S and 5 S rRNA genes and their flanking regions are presented. As compared to other eubacterial homologous molecules the 23 S rDNA exhibits two insertions. A 16 nucleotide long insertion is almost unique to members of the genus *Mycobacterium*, while the second represents an extended version of helix 54. The potential of both insertions to serve as target for diagnostic oligonucleotide probes was proven by comparative sequence analysis of 23 S rRNA of several *Mycobacterium* species and by dot blot hybridization. In addition, a 19-mer oligonucleotide probe is described, which can be considered genus *Mycobacterium*-specific.

Mycobacterium; *Mycobacterium leprae*; rDNA, 23 S and 5 S; rrn Operon; Diagnostic probes

1. INTRODUCTION

The genus *Mycobacterium* represents a group of Gram-positive, acid-fast bacteria which comprises a number of significant human and animal pathogens. Besides the well-known pathogens *M. leprae* and *M. tuberculosis*, several opportunistic pathogens exist (e.g. *M. kansasii, M. avium, M. intracellulare, M. simiae* and *M. flavescens*) which can cause severe infections whenever the normal cellular defence is depressed, e.g. AIDS [1], and other immunodeficiency diseases [2]. Despite progress in biochemical and immunological techniques, the needs for fast and reliable identification of mycobacterial species are obvious. This is particularly true for *M. leprae* because of its inability to grow outside its host.

The potential of large ribosomal (r) RNAs to serve as a most valuable source for both delineating phylogenetic relationships and taxon identification has been exploited over the last 12 years. Recent comparitive sequence analysis of 16 S rRNA/rDNA confirmed that the phenotypic division of mycobacteria into two separated clusters (the fast-growers, represented by basically harmless inhabitants of soil and water, and the slow-growers containing most of the overt pathogenic mycobacteria [3] is supported by their phylogenetic relationships [4–6]. The division between fast- and slow-growing mycobacteria is also reflected by the rRNA gene copy number: fast-growing mycobacteria contain two sets of rRNA genes, whereas slow-growers contain only one set [7,8].

16 S rRNA/rDNA sequence data permitted the first step in developing fast diagnostic assays for mycobacteria by specific rDNA oligonucleotide probes used either in solution (Gene-Trak, Framingham, MS, USA), in dot blot hybridization against bulk RNA [9] or in diagnostic PCR-mediated analysis [10]. High 16 S rRNA sequence similarities, on the other hand, detected especially between the overt pathogenic mycobacterial species, do only in a few cases permit the design of species-specific probes. This is in particular important with regard to PCR diagnosis, where a pair of specific sequences is advantageous for generating amplification products which in turn would indicate the presence of the respective organism in an infected tissue or biopsy. It is therefore obvious to investigate the 23 S rRNA/rDNA and the intracistronic spacer regions in order to evaluate the degree to which the variable regions can be used for the designation of oligonucleotide probes and primers.

Here we present the primary structure of the *Mycobacterium leprae* 23 S and 5 S rRNA genes and their flanking regions. Together with sequence information on the promoter region [11] and the primary structure of the 16 S rRNA of *M. leprae* [12] these data complete the nucleotide sequence of the one and only rDNA operon of *M. leprae*.

Correspondence address: W. Liesack, Department of Microbiology, University of Queensland, St. Lucia, Australia 4072. Fax: (61)(7) 365 1566

Published by Elsevier Science Publishers B.V.

2. MATERIALS AND METHODS

2.1. *Cloning*

The primary structure of the 23 S and 5 S rRNA genes from *M. leprae* was determined by sequencing the 5.3 kb *Pst*I-insert of plasmid pMLS2 [13]. For subcloning, *Pst*I/*Bam*HI and *Bam*HI restriction fragments of plasmid pMLS2 were separated on an 0.7% agarose gel. Gel slices containing restricted DNA were frozen at -70°C for 2 h and the DNA subsequently recovered by centrifuging through 0.22 μm filter units (Ultrafree-MC, Millipore, Bedford, MA 01730). Ligation of subfragments with phage vectors M13 mp18 or mp19, transfection, screening of recombinants and preparation of single stranded DNA were done as described [14].

2.2. *Oligonucleotide synthesis and sequencing*

Sequencing primers and oligonucleotide probes were synthesized using an Applied Biosystem 381A DNA assembler. DNA sequencing was carried out with Sequenase according to the manual of the manufacturer (U.S. Biochemicals Corporation, Cleveland, USA) using either plasmid pMLS2, prepared by CsCl-gradient centrifugation [15], or ssDNA of subfragments cloned into vector M13 as templates. Direct analysis of parts of the 23 S rRNA were performed according to the 16 S rRNA sequencing method [16]. PCR products of helix 27/31 were generated using the primer pair 5'GGGAGTG-AAATAGTACCTG and 5'TGGCCATGGGTAGATCACTC (*M. leprae* 23 S rDNA positions 584-602 and 788-807, respectively). PCR products of helix 54 were generated using the primer pair 5'GGACCTAAGGCGAGGCCG and 5'CGACGGATTT(A,G)CC-TA (positions 1460-1477 and 1710-1724, respectively). Generation of single stranded DNA via asymmetric PCR followed published procedures [17]. PCR products were purified by ammonium-acetate precipitation [18] and sequenced with sequenase.

2.3. *Dot blot hybridization*

The procedure for testing the specificity of diagnostic probes on bulk RNA and the names of 39 *Mycobacterium*- and 30 non-*Mycobacterium* reference strains used in dot blot hybridization have been published [19].

2.4. *Data analysis*

Sequences were aligned and homologies determined by means of Microgenie program [20, Beckman Instruments, Palo Alto, California, U.S.A.].

3. RESULTS AND DISCUSSION

The primary structure of the 5.3 kb insert of plasmid pMLS2 was determined over a length of 3.9 kb covering the sequence of the 23 S and 5 S rRNA genes plus flanking regions (Fig. 1) up to the 3′ terminus of the 16 S rDNA gene. The sequence has been deposited at EMBL data library under the accession number X56657. The 16 S rDNA sequence has been published recently [12].

3.1. *Primary structure of the 23 S rDNA and application for diagnostic oligonucleotide probes*

The putative 5′ and 3′ termini of the 23 S RNA gene from *M. leprae* (Fig. 1) were assigned by comparison with the 23 S rDNAs of two other actinomycetes, i.e. *Micrococcus luteus* [21] and *Streptomyces ambofaciens* [22]. Sequence similarities between *M. leprae* and *Mc. luteus* and between *M. leprae* and *S. ambofaciens* are 79.4% and 80.5%, respectively. With 3122 bp the length of the *M. leprae* 23 S rDNA is about the same as those of the two actinomycetes (*Mc. luteus*: 3095 bp; *S. ambofaciens*: 3120 bp), while those of other eubacteria is smaller, e.g. 2904 bp in *E. coli* [23], 2928 bp in *B. subtilis* [24] and 2876 bp in *Anacystis nidulans* [25].

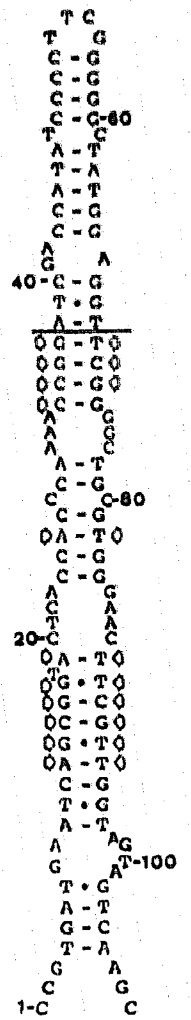

Fig. 1. Sequence alignment of the 16 S/23 S rDNA spacer and 5 S rRNA gene regions from *M. leprae* (Ml) and *M. bovis* BCG (Mb) [8,31]. Dots and bars indicate identical nucleotides and gaps, respectively. The ribosomal genes (16 S rDNA [EMBL accession number X53999], 23 S and 5 S rDNA [X56657]) are framed.

Comparison with eubacterial reference molecules [26] indicates the presence of a rather unique insertion of 16 nucleotides in the 23 S rDNA of the *M. leprae* sequence (position 747-762), present so far only in *Thermus thermophilus* [27]. Sequence analysis of 23 S rRNA from several mycobacteria showed the insertion not only to be common to slow and fast growing mycobacteria and maintained at the rRNA level, but also variable in length and sequence (Fig. 2). The insertion can be folded into a helix structure, which is located between helices 27 and 31 (domain II) of the universal 23 S rDNA secondary structure model [28]. The presence of this insertion could be also verified at the rRNA level

Fig. 2. Sequence alignment of a 23 S rDNA stretch corresponding to *M. leprae* 23 S rDNA position 747–762. Specific insertions for the genus *Mycobacterium* and for *Thermus thermophilus* are emphasized by brackets. The target region of the highly specific *M. leprae* probe is framed. Nucleotide positions involved in putative helix structures are underlined. Sequences were obtained as follows: [1]DNA sequence of a M13 clone; [2]direct sequencing of 23 S rRNA via reverse transcriptase and DNA sequencing of the opposite strand of cDNA synthesized by asymmetric PCR [17,18]; [3]direct sequencing of 23 S rRNA. All sequences are written in the DNA nomenclature to facilitate comparison.

for *M. bovis* and *M. tuberculosis* as well as for *M. avium* and *M. intracellulare*, with members of each pair exhibiting identical sequences (data not shown). This finding points towards the high potential of the primary structure of the insertion to serve as a target for diagnostic oligonucleotide probes with highly discriminating character for closely related species. In order to check the specificity, a 22-mer oligonucleotide probe directed against the stretch 5′GTATCACGTGTGAGCGTGTGTA (T_m of 66°C) of the *M. leprae* 23 S rRNA (Fig. 2) was tested by dot blot hybridization against bulk rRNA of *M. leprae* and 68 reference organisms (38 mycobacterial and 30 non-mycobacterial strains as indicated in Fig. 1 of ref. 19). Even under relaxed hybridization and washing conditions (both at 45°C, 21°C below T_m) the high specificity of this probe is obvious in that the only signals obtained were those with rRNA from *M. leprae* (homologous) and *M. kansasii* (heterologous). The *M. kansasii* signal, however, disappeared completely by increasing the washing temperature to 54°C (12°C below T_m). The signal of the homologous hybrid remained detectable even at the T_m of the DNA/rRNA duplex.

One of the most characteristic features of 23 S rRNAs from actinomycetes is the presence of an extended version of helix 54 (domain III, as defined by Höpfle et al. [28]. As compared to the length of the helix of *B. subtilis*, *M. leprae* possesses additional 103 nucleotides (Fig. 3). Comparative sequence analysis of helix 54 from several actinomycetes revealed a high degree of inter- and intra-generic sequence variations and it has therefore been postulated that this part of the 23 S rDNA constitutes an important feature for both intrageneric taxonomic analysis and designing targets for diagnostic oligonucleotide probes [29]. The sequence data obtained for helix 54 of different mycobacteria

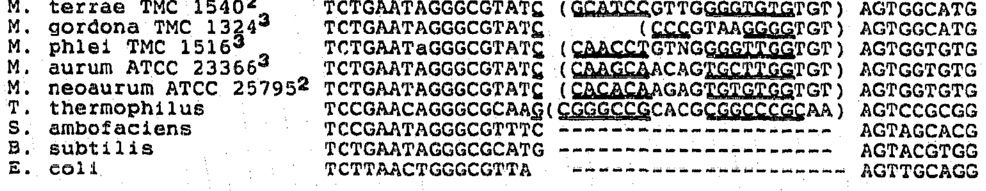

Fig. 3. Sequence alignment of helix 54 from 14 *Mycobacterium* species. The stretch corresponds to position 1522–1647 of the *M. leprae* 23 S rDNA. The target region of the published *M. leprae* probe [19] is underlined and the discriminating signature nucleotide marked by an asterisk. Identical sequences were determined for *M. bovis* and *M. tuberculosis* as well as for *M. avium* and *M. intracellulare* ATCC 23434. Sequences were obtained as indicated in the legend to Fig. 2.

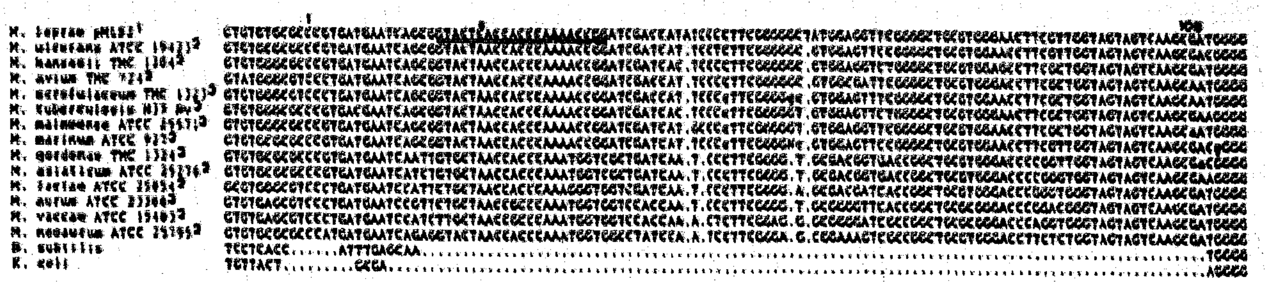

Fig. 4. Secondary structure model of the *M. leprae* insertion in helix 54 of the 23 S rDNA. Positions supported by co-ordinated base changes are marked by arrows. Positions 1-37 and 71-108 enclose the proposed universal secondary structure for the genus *Mycobacterium*, as derived from the information obtained for the 15 mycobacterial species investigated (Fig. 3). Since no general secondary structure can be obtained for the region between positions 38-70 this stretch is considered *M. leprae*-specific.

support this hypothesis in that the insertion exhibits both highly conserved and variable nucleotide positions (Fig. 3). Up to 40 differences are found between the 125 nucleotides of helix 54 of any pair of mycobacterial strains investigated. The secondary structure proposed for the insertion of the 23 S rRNA (Fig. 4) shows a stem-structure which for region 1-37 and 71-108 is supported by several co-ordinated base changes (see also Fig. 3). This part of the secondary structure can therefore be considered universal for the genus *Mycobacterium*. For region 38-70 we cannot propose a convincing universal secondary structure model. The potential of the insertion to be used for defining diagnostic probes has already been demonstrated with a highly specific 21-mer oligonucleotide probe for *M. leprae* [19]. The specificity of this probe is based upon a single *M. leprae*-specific signature nucleotide (C versus A, pos. 22) (Fig. 3).

Except for *M. phlei*, all mycobacterial strains investigated in this study have an identical stretch of 19 nucleotides between position 1661-1679 of the 23 S rRNA (according to the *M. leprae* numbering). A 19-mer oligonucleotide probe 5'ACCACTGAC(C/T)-GGTACGGCT, targeting this stretch at the rRNA level, was tested against the same dot blot matrix as cited above. The C/T degeneration within the probe was introduced to match a G-A exchange in the *M. phlei* sequence. The high specificity of this probe for the genus *Mycobacterium* was already obvious at low hybridization and washing temperature (40°C, corresponding to 20 (22)°C below T_m). Under these conditions rRNA from all mycobacteria and from *Tsukamurella aurantiaca* (a member of a neighboring genus [30]) displayed a strong signal. The *T. aurantiaca* signal disappeared at a washing temperature of 45°C while those of the mycobacteria were clearly visible. Increasing the washing temperature to 60°C resulted in the disappearance of signals from *M. chelonei* and *M. borstelense*. A mixture of this probe and species-specific probes for *M. chelonei* and *M. borstelense*, designed to match the same T_m values, could therefore be used under highly stringent conditions to detect all members of *Mycobacterium*.

3.2. *5 S rDNA*

The 5 S rDNA of *M. leprae* exhibits a length of 117 bp as determined by comparison with the 5 S rRNA from *M. bovis* BCG [31] (Fig. 1). The homology of 90.6% (11 mismatches) is significantly lower than that of the respective 16 S rRNA sequences (97.7%). Differences in the 5 S rRNA of *M. leprae* and other mycobacterial strains are especially conspicuous in the e-loop of the minimal model of the 5 S rRNA secondary structure [32].

3.3. *Spacer regions*

The 16 S/23 S rDNA spacers of *M. leprae* and *M. bovis* [33] are highly similar with respect to length (282 and 277 bp, respectively) and the absence of a tRNA gene. Since this organization is also found in the homologous spacer regions of *Streptomyces lividans* (278 bp [34]), *S. ambofaciens* (303 bp, [22]) and *S. coelicolor* A3(2) (277 bp, [35]) it can be assumed that the absence of a tRNA gene and an approximate length of 300 nucleotides are general features of the 16 S/23 S rDNA spacer of actinomycetes. Interestingly, the spacer regions of the two mycobacteria share a sequence homology of only 75.9% (Fig. 1). Thus, the degree of sequence variability of the 16 S/23 S rDNAs spacer is significantly higher than that between rRNA genes, even between closely related organisms. The possibility of rapid sequence analysis of 16 S/23 S rDNA spacers by PCR-mediated amplification makes therefore this part of the rrn operon of actinomycetes an interesting region for the designation of taxon-specific PCR- and sequencing primers as well as of diagnostic probes. Further studies are however necessary to determine the evolutionary stability of the intergeneric spacer sequence at the species- and strain level.

Acknowledgement: This work was supported by a grant from the NHMRC to E.S.

REFERENCES

[1] Collins, F.M. (1988) Springer Semin. Immunopathol. 10, 375-391.

[2] Wolinsky, E. (1984) Microbiol. Ser. 15, 1141–1207.
[3] Wayne, L.G. and Kubica, G.P. (1986) in: Bergey's Manual of Systematic Bacteriology (Sneath, P.H.A., Mair, N.S., Sharpe, M.E. and Holt, J.G. eds.) pp. 1436–1457, Williams and Wilkins, Baltimore.
[4] Stackebrandt, E., Smida, J. and Kazda, J. (1989) Acta Leprologica 7, 222–225.
[5] Stahl, D.A. and Urbance, J.W. (1990) J. Bacteriol. 172, 116–124.
[6] Rogall, T., Flohr, T. and Boettger, E.C. (1990) J. Gen. Microbiol. 136, 1915–1920.
[7] Bercovier, H., Kafri, O. and Sela, S. (1986) Biochem. Biophys. Res. Commun. 136, 1136–1141.
[8] Suzuki, Y., Yoshinaga, K., Ono, Y., Nagata, A. and Yamada, T. (1987) J. Bacteriol. 169, 839–843.
[9] Pitulle, C., Stackebrandt, E. and Kazda, J. (1990) Int. J. Leprosy 58, 130–133.
[10] Boeddinghaus, B., Rogall, T., Flohr, T., Bloecker, H. and Boettger, E.C. (1990) J. Clin. Microbiol. 28, 1751–1759.
[11] Sela, S. and Clark-Curtiss, J.E. (1991) Gene, in press.
[12] Liesack, W., Pitulle, C., Sela, S. and Stackebrandt, E. (1990) Nucleic Acids Res. 18, 5558.
[13] Sela, S., Clark-Curtiss, J.E. and Bercovier, H. (1989) J. Bacteriol. 171, 70–73.
[14] Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1987) Cur. Prot. Mol. Biol., Wiley, New York.
[15] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory.
[16] Stackebrandt, E. and Charfreitag, O. (1990) J. Gen. Microbiol. 136, 37–43.
[17] McCabe, P.C. (1990) in: PCR Protocols – A Guide to Methods and Applications (Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J. eds) pp. 76–83, Academic Press, San Diego.
[18] Brown, M.A.D. (1990) in: PCR Protocols – A Guide to Methods and Applications (Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J. eds) pp. 189–196, Academic Press, San Diego.
[19] Liesack, W., Pitulle, C. and Stackebrandt, E. (1990) Lett. Appl. Microbiol. 11, 96–99.
[20] Queen, C. and Korn, L.J. (1984) Nucleic Acids Res. 12, 581–599.
[21] Regenburger, A., Ludwig, W., Frank, R., Bloecker, H. and Schleifer, K.H. (1988) Nucleic Acids Res. 16, 2344.
[22] Pernodet, J.-L., Boccard, F., Alegre, M.-T., Gagnat, J. and Guerineau, M. (1989) Gene 79, 33–46.
[23] Brosius, J., Dull, T.J. and Noller, H.F. (1980) Proc. Natl. Acad. Sci. USA 77, 201–204.
[24] Green, C.J., Stewart, G.C., Hollis, A.H., Vold, B.S. and Bott, K.F. (1985) Gene 37, 261–266.
[25] Douglas, S.E. and Doolittle, W.F. (1984) Nucleic Acids Res. 12, 3373–3386.
[26] Lane, D.J. (1991) in: Nucleic Acid Techniques in Bacterial Systematics (Stackebrandt, E. and Goodfellow, M. eds) pp. 115–174, Wiley, Chichester.
[27] Höpfl, P., Ulbrich, N., Erdmann, V.A., Ludwig, W. and Schleifer, K.H. (1988) Nucleic Acids Res. 16, 9043.
[28] Höpfl, P., Ludwig, W., Schleifer, K.H. and Larsen, N. (1990) Eur. J. Biochem. 185, 355–364.
[29] Stackebrandt, E., Witt, D., Kemmerling, C., Kroppenstedt, R. and Liesack, W. (1991), J. Appl. Env. Microbiol., in press.
[30] Collins, M.D., Smida, J., Dorsch, M. and Stackebrandt, E. (1988) Int. J. Syst. Bacteriol. 38, 385–391.
[31] Ohkubo, S., Iwasaki, H., Hori, H. and Osawa, S. (1986) J. Biochem. 100, 1261–1267.
[32] Wolters, J. and Erdmann, V.A. (1988) Nucleic Acids Res. 16, r1–r70.
[33] Suzuki, Y., Nagata, A., Ono, Y. and Yamada, T. (1988) J. Bacteriol. 170, 2886–2889.
[34] Suzuki, Y., Ono, Y., Nagata, A. and Yamada, T. (1988) J. Bacteriol. 170, 1631–1636.
[35] Baylis, H.A. and Bibb, M.J. (1988) Mol. Microbiol. 2, 569–579.