JOURNAL OF BACTERIOLOGY, June 1988, p. 2886–2889
0021-9193/88/062886-04$02.00/0

Vol. 170, No. 6

Complete Nucleotide Sequence of the 16S rRNA Gene of *Mycobacterium bovis* BCG

YASUHIKO SUZUKI,* AKIHISA NAGATA, YASUKO ONO, AND TAKESHI YAMADA

Research Institute for Microbial Diseases, Osaka University, Yamada-Oka 3-1, Suita City, Osaka 565, Japan

Received 12 January 1988/Accepted 18 March 1988

The complete nucleotide sequence of the 16S rRNA gene of *Mycobacterium bovis* BCG was determined. Its coding region was estimated to be 1,536 base pairs long. The nucleotide sequence of the gene in *M. bovis* BCG has homologies of 75 and 89% with those of *Escherichia coli* and *Streptomyces lividans*, respectively.

The special interest in the structure of the 16S rRNA from *Mycobacterium bovis* BCG arose from the finding that mycobacterial ribosome does not translate f2 bacteriophage RNA in a cell-free system (12). Since the 3′ terminus of bacterial 16S rRNA has been implicated in the recognition of polypeptide chain initiation (6, 7), we analyzed the structure of 16S rRNA from *M. bovis* BCG. Southern hybridization analysis of rRNAs from *M. bovis* BCG was performed to determine the number of rRNA genes and the strategy for cloning. The results suggest that *M. bovis* BCG possesses only a minimum set of rRNA genes and that the spacer region between the genes for 16S and 23S rRNAs contains no tRNA genes (10). This paper describes the complete nucleotide sequence of the 16S rRNA gene and the spacer region between the 16S and 23S rRNA genes.

All the restriction enzymes, T4 ligase, and T4 polynucleotide kinase were purchased from Toyobo Co. Ltd. (Tokyo, Japan); T4 RNA ligase was from Pharmacia Japan Co. Ltd. (Tokyo); phage M13 DNA, the 17-mer primer, and the deaza-DNA sequencing kit were from Takara Shuzo Co. Ltd. (Kyoto, Japan); [α-^{32}P]dCTP and [5′-^{32}P]cytidine-3′, 5′-bisphosphate were purchased from Amersham Japan Co. Ltd. (Tokyo, Japan). The oligonucleotide primer for primer extension was a gift from Yamazaki and Matsuo, Central Institute for Fundamental Science, Ajinomoto Co. Ltd. (Kanagawa, Japan).

The recombinant plasmid pBCG4 containing the entire rRNA operon of *M. bovis* BCG was propagated in *Escherichia coli* HB101 cells and isolated as described by Hughes and Meynell (1). The DNA was digested by *Bal*I and separated by 1.0% agarose gel electrophoresis. The 1.7-kilobase-pair (kbp) DNA fragment was purified by using DE81 paper (10), cloned into the *Sma*I site of pUC18, and named pBCG101. The 1.0-kbp *Bam*HI fragment was also cloned into the *Bam*HI site of pUC18 and named pBCG102.

Various restriction fragments were cut out from pBCG101 and pBCG102, cloned into M13mp18, and transfected into *E. coli* JM109 to prepare single-stranded sequencing templates. Every fragment was sequenced on both plus and minus strands at least twice by the dideoxynucleotide chain terminator method of Sanger et al. (5), with modifications as described elsewhere (2).

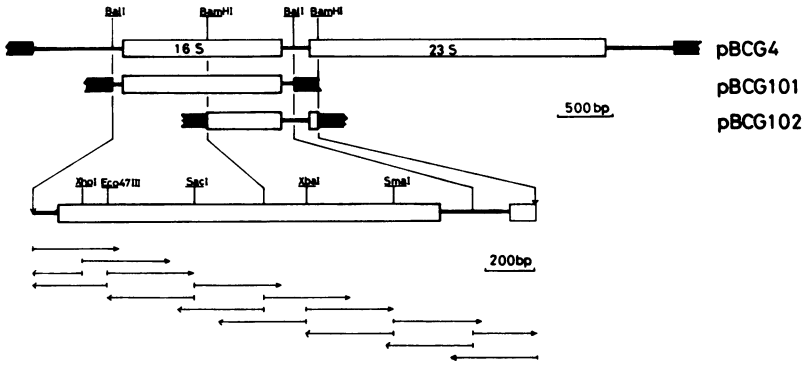

FIG. 1. Physical maps of the cloned rRNA gene cluster in plasmids pBCG4, pBCG101, and pBCG102 and sequencing strategy for the 16S rRNA gene. White boxes represent the rRNA coding regions. The vectors pBR322 and pUC18 are shown by black boxes. The lower part of the figure shows the sequencing strategy.

The 3′ terminal of 16S rRNA was labeled with ^{32}P by using T4 RNA ligase and [5′-^{32}P]cytidine-3′,5′-bisphosphate, partially digested by a chemical method (3) and subjected to 8 M urea-12% polyacrylamide gel electrophoresis. Results were obtained after exposure for 16 h at −70°C. RNA extraction was performed as described previously (11).

About 2 μg of 16S rRNA and 1 pmol of oligonucleotide primer (5′-TGCATGTGTTAAGCACG-3′) labeled with ^{32}P at the 5′ end were dissolved in 10 μl of 0.4 M NaCl-40 mM

* Corresponding author.

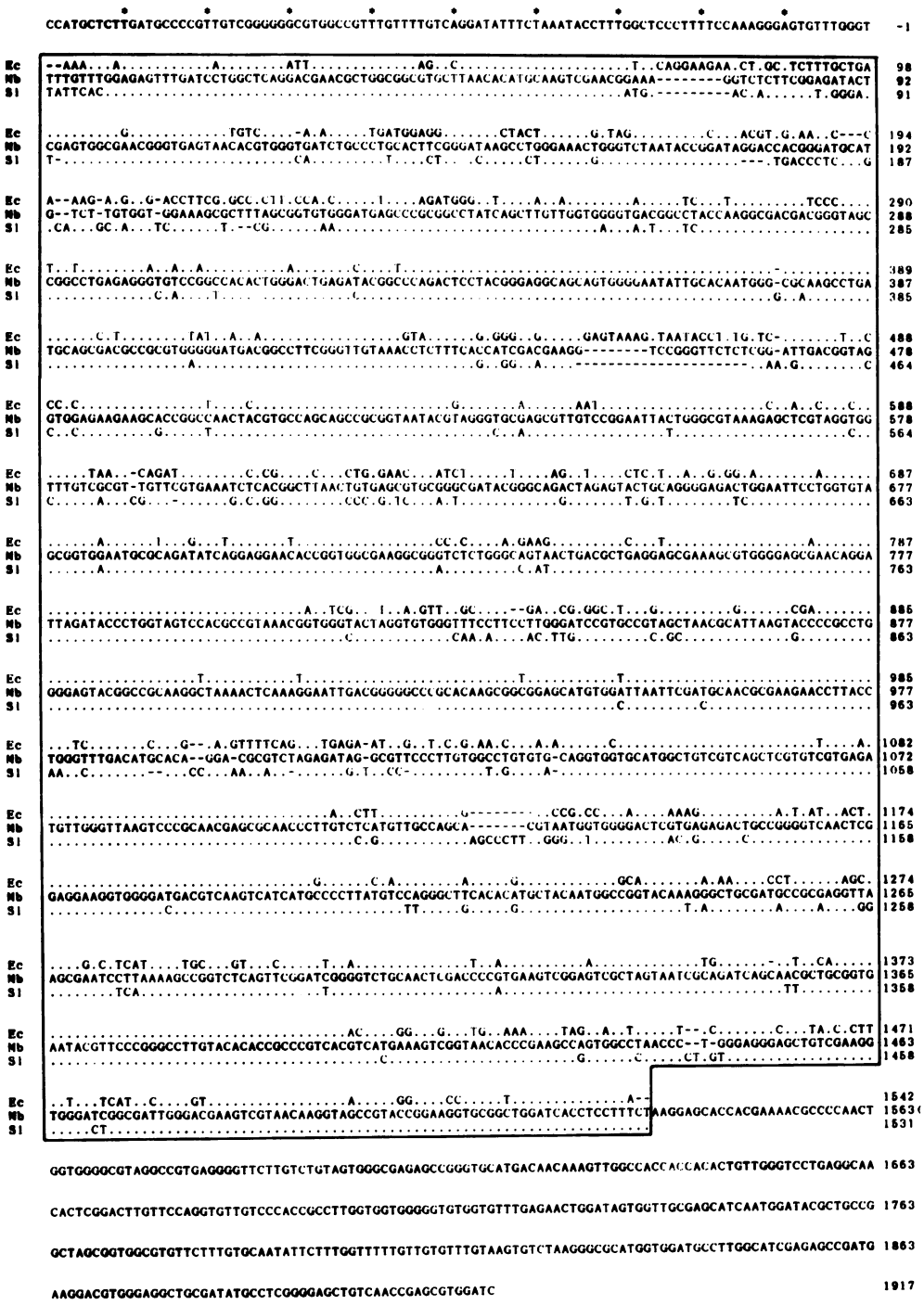

FIG. 2. Total nucleotide sequence of the 1.7-kbp *Bal*I and 1.0-kbp *Bam*HI fragments of pBCG4. The noncoding (RNA-like) strand of *M. bovis* BCG (Mb) 16S rRNA gene is shown, together with those of *E. coli* (Ec) and *S. lividans* TK21(S1) 16S rRNA genes for comparison. The 16S rRNA coding regions are boxed. Dots indicate nucleotides identical with those of *M. bovis* BCG. Bars indicate deletions.

PIPES (piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)-1 mM EDTA. The solution was incubated at 95°C for 5 min and cooled to 42°C. The nucleic acids were then precipitated with ethanol, washed twice with 70% ethanol, dried in vacuo, and redissolved in 20 μl of 50 mM Tris hydrochloride (pH 8.3)-120 mM KCl-8 mM $MgCl_2$-2 mM dithiothreitol-0.5 mM deoxynucleoside triphosphates with 20 U of reverse transcriptase. After incubation at 42°C for 1 h, the reaction was stopped by ethanol precipitation. The precipitate was washed twice with 70% ethanol, dried in vacuo, redissolved in 10 μl of 10 mM Tris hydrochloride-1 mM EDTA (pH 8), and then subjected to 8 M urea-6% polyacrylamide gel electrophoresis.

The mapping of the cloned 16S rRNA gene, including its 3′- and 5′-flanking regions, and the sequencing strategy are shown in Fig. 1. As has been pointed out previously (8), the

high G+C content increases the chances of sequencing artifacts due to compression. Every sequence was confirmed by sequencing both strands at least twice. Figure 2 shows the total nucleotide sequence of the noncoding (RNA-like) strand of the 1.7-kbp *Bal*I fragment and the 1.0-kbp *Bam*HI fragment. The 3′ end was determined by RNA sequencing. The 5′ end was determined by the primer extension experiment using oligonucleotide primer (Fig. 3). The 5′ end of the 23S rRNA gene was not determined experimentally but was assigned on the basis of sequence homologies with that of *E. coli* (11). The coding region of the *M. bovis* BCG 16S rRNA gene was 1,536 bp long. The 1.7-kbp *Bal*I fragment contained a 100-bp 5′-flanking sequence, while the 1.0-kbp *Bam*HI fragment contained the spacer between the 16S and 23S rRNA genes (276 bp) and the 5′-terminal sequence of the 23S rRNA gene (96 bp). The 16S rRNA gene of *M. bovis* BCG was 6 bp shorter than that of *E. coli* and 5 bp longer than that of *Streptomyces lividans* (9). The nucleotide sequence (Fig. 3) has homologies of 75 and 89% with those of *E. coli* and *S. lividans*, respectively. There are also homologies of secondary structures among the three organisms.

Relatively larger insertions and deletions compared with those of *E. coli* were seen within the 5′-terminal half of the 16S rRNA gene. The large deletions at positions 74 and 75 and 453 and 454 existed in both the *M. bovis* BCG and the *S. lividans* 16S rRNA genes. Insertions were found at positions 189 to 191, 834 and 835, and 1124. The larger insertion at position 1124 was also found in *S. lividans*.

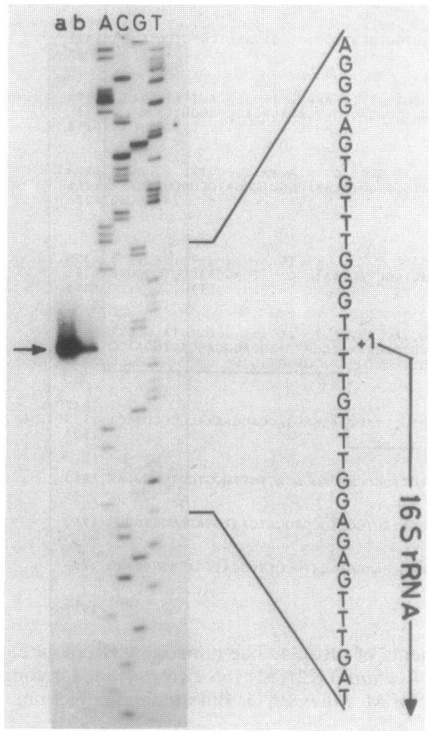

FIG. 3. Primer extension and dideoxy-sequence analysis with the 5′-labeled oligonucleotide primer (complementary to positions 45 to 61 in Fig. 2). For primer extension, the primer was hybridized to 16S rRNA of *M. bovis* BCG and extended by reverse transcriptase. Lanes a and b, Elongation products with different amounts of primer (0.1 and 1 pmol, respectively) hybridized to 2 μg of *M. bovis* BCG 16S rRNA. For dideoxy sequencing, the same primer was hybridized to the single-stranded template containing the 1.0-kbp *Sal*I-*Xho*I fragment of pBCG4 in the *Sal*I site of M13mp18 and extended with the Klenow fragment of *E. coli* DNA polymerase.

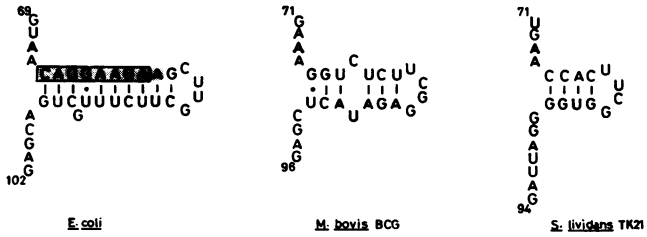

FIG. 4. Secondary structures at positions 71 to 96 of *E. coli*, *M. bovis* BCG and *S. lividans* TK21 16S rRNAs.

Secondary-structure models for *E. coli* 16S rRNA have been proposed by Woese et al. (11). Almost all stem and loop structures seen in *E. coli* 16S rRNA were observed in *M. bovis* BCG 16S rRNA, but there were two positions in which the stem structures were shortened by the large deletion. These positions were at 71 to 96 and 434 to 488. The stems of these regions in *M. bovis* BCG were intermediate between those of *E. coli* and *S. lividans* (Fig. 4 and 5). The insertion at position 1124 did not drastically change the secondary structure in *M. bovis* BCG, whereas it strongly affected that of *S. lividans*. The region from positions 132 to 223 was one of the examples of low conservation. More than 60% of residues were substituted, and the secondary structure could not be constructed in both *M. bovis* BCG and *S. lividans*. The most conserved stem and loop structure was the region from positions 490 to 535. There were only three substitutions.

The large stem and loop structure region at positions 566 to 754 where the S8 and S15 ribosomal proteins bind initially in *E. coli* was also of interest. We found many substitutions in the stem structure, but the structure of the stems was well conserved. The secondary structure here may be very important for the binding of these proteins. We could find similar situations at the three other positions 811 to 871, 989 to 1044, and 1401 to 1483.

The region near the 3′ end was conserved, except for one substitution and two additions. The Shine-Dalgarno sequences (6) were the same as those in *E. coli* and *Bacillus subtilis*. This similarity implies that the failure of ribosomes in mycobacteria to translate the f2 phage RNA is not attributable to the structural difference of the Shine-Dalgarno sequence. Whether the other bases adjacent to the

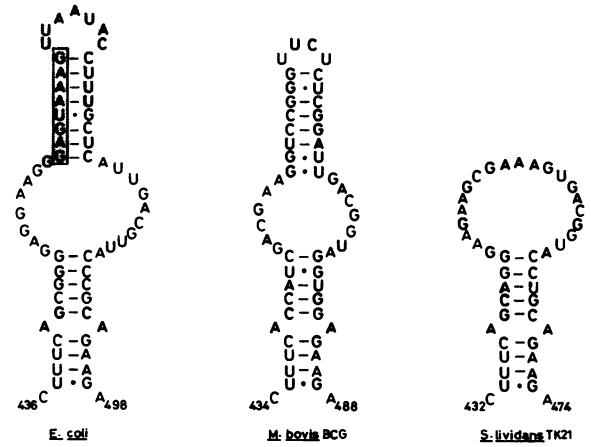

FIG. 5. Secondary structures at positions 434 to 488 of *E. coli*, *M. bovis* BCG and *S. lividans* TK21 16S rRNAs.

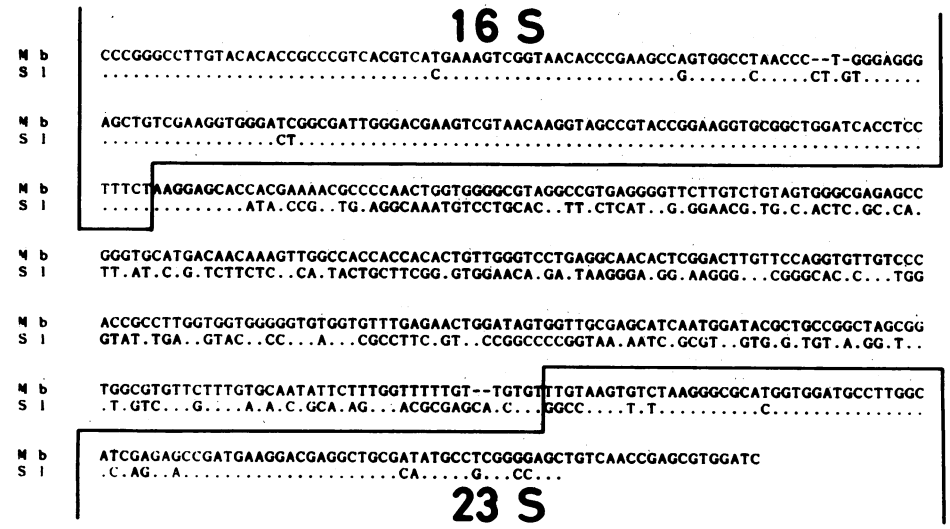

FIG. 6. Nucleotide sequence of the spacer region between 16S and 23S rRNA genes. The noncoding strand of the *M. bovis* BCG (M) spacer region is shown together with that of *S. lividans* (S) for comparison. Dots indicate nucleotides identical with those of *M. bovis* BCG. Bars indicate deletions.

CCTCC sequence are attributable to the translational barrier of mycobacterial ribosomes for f2 phage RNA remains to be elucidated.

Looking over all the sequences, we found that among these three strains the highly conserved regions were found more in loop regions than in stem regions.

With respect to 16S rRNA gene structure, these results show that *M. bovis* BCG is phylogenetically more closely related to *S. lividans* than to *E. coli*. However, we must further examine the structure and function of the promoter region of *M. bovis* BCG to draw a final conclusion on the question of whether the rRNA operon of this organism is more closely related to that of *S. lividans* than to that of *E. coli*. This work will be reported elsewhere.

With respect to the spacer region between the 16S and the 23S rRNA genes, no tRNA gene was found from the sequence analysis of the T-G-psi-C stem and loop structure of common tRNAs (Fig. 6). This conclusion was consistent with the result of Southern hybridization analysis (10).

Another interesting point we have to stress regarding the spacer region is that homology between the spacer regions of *M. bovis* BCG and *S. lividans* was strikingly lower than that between the 16S rRNA genes of the two organisms. The homologies between the 16S rRNA gene and a part of the 23S rRNA gene of *M. bovis* BCG and *S. lividans* were 89 and 80%, respectively, while that between the spacer regions was lower than 40%. As the structure of 16S rRNA is closely related to the function of ribosomes, the change of bases in the gene might be strongly restricted. Thus, the speed of substitution of bases in the structural gene might be slower than that in the spacer region.

We thank Takahisa Yamanouchi for his interest and encouragement.

This work was supported in part by grants from the Arima Memorial Foundation for Medical Research and grant-in-aid 61570210 from the Japanese Ministry of Education.

LITERATURE CITED

1. **Hughes, C., and G. G. Meynell.** 1977. Rapid screening for plasmid DNA. Mol. Gen. Genet. **151**:175–179.
2. **Mizusawa S., S. Nishimura, and F. Seela.** 1986. Improvement of the dideoxy chain termination method of DNA sequencing by use of deoxy-7-deazaguanosine triphosphate in place of dGTP. Nucleic Acids Res. **14**:1319–1324.
3. **Peattie, D. A.** 1979. Direct chemical method for sequencing RNA. Proc. Natl. Acad. Sci. USA **76**:1760–1764.
4. **Rubin, G. M.** 1975. Preparation of RNA and ribosomes from yeast. Methods Cell Biol. **12**:45–64.
5. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
6. **Shine, J., and L. Dalgarno.** 1974. The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding site. Proc. Natl. Acad. Sci. USA **71**:1342–1346.
7. **Shine, J., and L. Dalgarno.** 1975. Determinant of cistron specificity in bacterial ribosomes. Nature (London) **254**:34–38.
8. **Shinnick, T. M.** 1987. The 65-kilodalton antigen of *Mycobacterium tuberculosis* J. Bacteriol. **169**:1080–1088.
9. **Suzuki, Y., and T. Yamada.** 1988. The nucleotide sequence of a 16S ribosomal RNA gene in *Streptomyces lividans* TK21. Nucleic Acids Res. **16**:370.
10. **Suzuki, Y., K. Yoshinaga, Y. Ono, A. Nagata, and T. Yamada.** 1987. Organization of rRNA genes in *Mycobacterium bovis* BCG. J. Bacteriol. **169**:839–843.
11. **Woese, C. R., R. Gutell, R. Gupta, and H. F. Noller.** 1983. Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. Microbiol. Rev. **47**:621–669.
12. **Yamada, T.** 1982. Characterization of an S1-like protein in *Mycobacterium smegmatis* ribosomes. FEBS Lett. **142**:267–270.