JOURNAL OF BACTERIOLOGY, Sept. 1995, p. 4836–4840 Vol. 177, No. 17
0021-9193/95/$04.00+0

Isolation and Sequencing of the Replication Region of *Mycobacterium avium* Plasmid pLR7

MARJORIE L. BEGGS,[1]* JACK T. CRAWFORD,[2] AND KATHLEEN D. EISENACH[1]

Department of Pathology, University of Arkansas for Medical Sciences, and J. L. McClellan Memorial Veterans Hospital, Little Rock, Arkansas 72205,[1] and Division of Bacterial and Mycotic Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia 30333[2]

Received 18 January 1995/Accepted 24 June 1995

The *Mycobacterium avium* plasmid pLR7 is representative of a group of small plasmids that are common in isolates from AIDS patients with disseminated *M. avium* infections. Determination of the functions of these and other plasmids has been hampered by the lack of methods for genetic manipulation of *M. avium*. In this study, the region of pLR7 capable of replication was identified and sequenced. Fragments of pLR7 were cloned into a pUC18 derivative carrying a kanamycin resistance marker and introduced into a plasmid-free *M. avium* strain by electroporation. The origin of replication was located on a 1.8-kb *Pvu*II-to-*Sma*I fragment. An open reading frame encoding a putative Rep protein was identified. Two other open reading frames were identified in this region. A shuttle vector, pMB351, was constructed with the pLR7 origin of replication, pUC18, and the kanamycin resistance gene from Tn*5*. This vector was successfully transformed into *M. avium*, *Mycobacterium tuberculosis*, and *Mycobacterium bovis*.

Mycobacterium avium is predominately a pathogen of birds and other animals but is also capable of causing infection in humans. Since the emergence of human immunodeficiency virus, *M. avium* infections in humans have increased dramatically. *M. avium* is also commonly found in the environment.

M. avium isolates frequently carry plasmids—in some cases, multiple plasmids (5). pLR7 is a 15.3-kb plasmid that was described previously (3). Plasmids showing close homology with pLR7 are often present in *M. avium* isolates from AIDS patients (4). At the present time, only two other mycobacterial plasmid origins of replication have been described: pAL5000 from *Mycobacterium fortuitum* and pMSC262 from *Mycobacterium scrofulaceum* (13, 14). Both of these plasmids have been used to construct mycobacterium-*Escherichia coli* vectors. It is of interest to determine how these plasmids compare with pLR7 in their origins of replication, since this will expand our understanding of how plasmids replicate and are maintained in mycobacteria. Characterization of pLR7 will also provide insight into the functions of native plasmids in *M. avium*.

A number of investigators have been unsuccessful in attempts to transform various *M. avium* strains. One possible explanation for this difficulty is interference by preexisting plasmids in the strains used. Hellyer et al. described several plasmid-free strains of *M. avium* that were isolated from AIDS patients in the United Kingdom (9). With these strains, we were able to obtain transformants after electroporation with plasmids carrying the origin of the *M. fortuitum* plasmid pAL5000. With this system, we were able to identify the region of pLR7 required for replication in *M. avium*. This region was sequenced and is described. This origin was used to create a new mycobacterial shuttle vector.

* Corresponding author. Mailing address: John McClellan Memorial Veterans Hospital, Medical Research Department, Slot 151, Room GB126, 4300 W. 7th St., Little Rock, AR 72205. Phone: (501) 661-1202, ext. 2965. Fax: (501) 671-2510. Electronic mail address: MLBEGGS@life.uams.edu.

MATERIALS AND METHODS

Cells and growth media. Plasmid-free *M. avium* strains A4 and A5 (both serotype 4) were obtained from Tobin Hellyer (St. Mary's Hospital Medical School, London, United Kingdom). *M. avium* LR113, a serotype 4 strain that carries plasmid pLR7, was obtained from the culture collection at National Jewish Hospital. *Mycobacterium tuberculosis* H37Rv and H37Ra and *Mycobacterium bovis* 410 were obtained from the American Type Culture Collection (Rockville, Md.). *Mycobacterium smegmatis* LR222 is a highly transformable strain isolated in our laboratory from *M. smegmatis* mc^26 (17). *Escherichia coli* DH5α was obtained from Bethesda Research Laboratories (Gibco BRL, Grand Island, N.Y.). All mycobacterial cells were cultivated in Middlebrook 7H9 broth with 0.1% Tween and Dubos oleic albumin complex enrichment. All transformants were isolated on 7H10 agar containing 50 μg of kanamycin per ml. For plasmid isolations, transformants were grown in 7H9 broth with 50 μg of kanamycin per ml.

DNA. pUC18 and pUC19 were obtained from Gibco BRL. pJC74 was constructed in our laboratory and consists of pUC19 with the Tn*5* kanamycin resistance gene replacing the ampicillin gene. pJC85 was also created in our laboratory and is an *E. coli*-mycobacterium shuttle vector that contains the pAL5000 origin of replication linked to the kanamycin resistance gene inserted into pUC19. Table 1 lists clones used in this study.

Electroporation. Mycobacterial cells were grown in sidearm flasks or tissue culture plates with gentle agitation until a Klett unit reading of 100 to 200 was obtained. The cells were collected, washed three times with cold sterile 10% glycerol, resuspended in 1/100 volume of 10% glycerol, and dispensed into 60-μl aliquots. DNA (1 μg or less) was mixed with the cells, and the samples were set on ice for 10 min. Electroporation was performed with a BTX cell manipulator 600 (Biotechnologies and Experimental Research Inc., San Diego, Calif.) in 1 mm gap disposable cuvettes at 1.2 kV for 5 ms. These settings correlate to a pulse of 12 kV/cm. After electroporation, the cells were washed out of the cuvette with 1 ml of medium and incubated at 37°C for 2 h before being plated.

Cloning and sequencing of pLR7. A restriction endonuclease map of pLR7 is presented in Fig. 1. The *Hin*dIII to *Eco*RI_1 fragment of pLR7 was digested with various enzymes to create subclones for sequencing (Table 1). Standard procedures were used for digestion, ligation, and transformation (15). Sequencing was performed on an ABI 373 DNA Sequencer with the *Taq* DyeDeoxy Terminator Cycle sequencing kit (Perkin-Elmer, Applied Biosystems Inc., Foster City, Calif.). DNA for sequencing was prepared with the Qiagen mini (tip 20) or midi (tip 100) kit according to the manufacturer's protocol (Qiagen, Chatsworth, Calif.). Sequencing primers were designed with the Oligo 4.0 Primer Analysis software (National Biosciences, Inc., Plymouth, Minn.). Primers were synthesized by the University of Arkansas Medical Sciences core facility. Sequence analysis was performed with Seqed (Applied Biosystems Inc.), DNAstar (Madison, Wis.), and the GCG package (Genetics Computer Group, Madison, Wis.).

Nucleotide sequence accession number. The nucleotide sequence reported here will appear in the GenBank nucleotide sequence database under accession number U18777.

TABLE 1. Plasmids used in this study

Plasmid[a]	Relevant characteristics
pJC74	*E. coli* plasmid vector, Km^r
pJC85	*E. coli*-mycobacterium (pAL5000) shuttle vector, Km^r
pMB4	*Sal*I_6-*Sal*I_7 pLR7 fragment in pUC18
pMB5	*Sal*I_5-*Sal*I_6 pLR7 fragment in pUC18
pMB45	*Hin*dIII-*Xba*I pLR7 fragment in pJC74
pMB46	*Xba*I-*Hin*dIII pLR7 fragment in pJC74
pMB47	Complete pLR7 at *Hin*dIII site in pJC74
pMB50	*Bam*HI-*Eco*RI_1 pLR7 fragment in pJC74
pMB51	*Eco*RI_3-*Bam*HI pLR7 fragment in pJC74
pMB67	pMB45 with the *Bam*HI site destroyed
pMB80	*Sma*I_5-*Bam*HI pLR7 fragment in pUC18
pMB126	*Bam*HI-*Sal*I_1 pLR7 fragment in pUC18
pMB131	*Sal*I_7-*Sma*I_5 pLR7 fragment in pUC18
pMB144	*Sal*I_1-*Sal*I_2 pLR7 fragment in pUC18
pMB150	*Sma*I_5-*Xba*I pLR7 fragment in pJC74
pMB155	*Sca*I-*Xba*I pLR7 fragment in pJC74
pMB158	*Sca*I-*Sal*I_1 pLR7 fragment in pJC74
pMB168	*Sca*I-*Pvu*II pLR7 fragment in pJC74

[a] All plasmids were created in our laboratory.

RESULTS

Transformation of *M. avium*. Initial transformations of *M. avium* A4 and A5 were performed with pJC85, which carries a known mycobacterial origin of replication from pAL5000 (14). Electroporation controls were *M. smegmatis* LR222 and *M. avium* LR113. Transformation was successful with strain LR222 and strains A4 and A5, with efficiencies of transformation of 10^5 and 10^4/μg of DNA, respectively. There were no transformants with LR113, which carries the pLR7 plasmid. Several colonies from each of the transformations were selected, and plasmid isolation was performed. All contained the pJC85 plasmid. Next, plasmids pMB45, pMB46, pMB47, pMB50, and pMB51 (Table 1) were used to transform LR222 and A4. These clones contain large, overlapping fragments of pLR7 and represent the entire plasmid. The pMB47 clone contains the entire plasmid cloned at the *Hin*dIII site. Plasmids pMB45 and pMB47 yielded transformants in A4, whereas the other plasmids did not. Plasmids isolated from the transformants appeared to be identical to the input plasmids. More transformants were obtained with pMB45 than with pMB47, which may reflect the larger size of pMB47. Repeated attempts to transform pMB45 into *M. smegmatis* LR222 were unsuccessful.

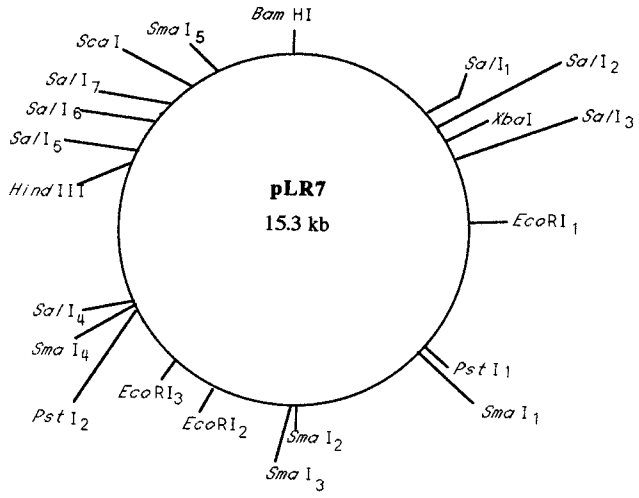

FIG. 1. Restriction endonuclease map of pLR7. The numbering of sites proceeds to the right of the *Bam*HI site, which is designated as 0.

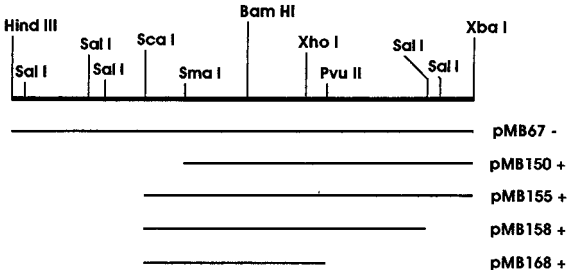

FIG. 2. Subclones and deletions created in the region of pLR7 known to contain the origin of replication. The clones are identified, and the + and − signs indicate whether (+) or not (−) this clone was able to replicate in A4. pMB67 contains the entire region known to contain the ability to replicate, with the *Bam*HI site eliminated.

Determination of the minimum region required for replication. The results presented above indicated that the replication region was contained in the *Hin*dIII-to-*Xba*I fragment (Fig. 1). To more closely delineate the essential region, *Sca*I-to-*Xba*I (pMB155), *Sma*I-to-*Xba*I (pMB150), *Sca*I-to-*Sal*I (pMB158), and *Sca*I-to-*Pvu*II (pMB168) fragments were cloned into pJC74 (Fig. 2). In addition, pMB45 was digested with *Bam*HI, end filled, and religated (pMB67), which eliminates the *Bam*HI site and adds four nucleotides to the reading frame. All of the clones described above were electroporated into A4. Figure 2 indicates which clones were able to replicate in A4. Thus, the region necessary for replication was localized in a fragment extending from *Sma*I to *Pvu*II. Lack of replication by pMB67 indicates that the *Bam*HI site is located in an essential reading frame.

Sequencing of the region containing the pLR7 origin of replication. The clones used to sequence the pLR7 origin of replication are indicated in Table 1. Linkage of the *Sal*I fragments was accomplished with the terminal primers in a region with pMB45 to ascertain the sequence of adjacent DNA. Sequence data from the *Pvu*II site to the *Sma*I site are presented in Fig. 3. This fragment is 1,828 bp in length and has a GC content of 70%. This region demonstrates 69% identity with the *M. scrofulaceum* plasmid pMSC262 (nucleotides 570 to 1735) (15) and 54.9% identity in a 204-bp overlap with pAL5000 (nucleotides 2658 to 2862) (17). There are unique restriction endonuclease sites for *Bam*HI, *Xho*I, *Pvu*II, *Xma*I, and *Sma*I. There are no restriction endonuclease sites for *Hin*dIII, *Eco*RI, *Eco*RV, *Kpn*I, *Pst*I, *Xba*I, or *Ssp*I. The sequence for the entire plasmid will be published elsewhere.

ORF analysis. There are a number of open reading frames (ORFs) in this region of pLR7. The locations of ORF1, ORF4, and ORF9 are shown in Fig. 4. Both ORF1 and ORF4 span the *Bam*HI site and would have been disrupted in pMB67. ORF1 is 1,083 nucleotides in size, with an initiation codon of GTG at position 284 and a stop codon at position 1364 (Fig. 3). ORF1 demonstrates typical mycobacterial codon usage (6, 7). The ribosome binding site and putative −10 and −35 promoter regions are indicated. ORF1 is preceded by an AT-rich (58%) region from 127 to 150 nucleotides and a GC-rich region (86%) from 183 to 226 nucleotides. There are three inverted repeats (IRs) in the sequence upstream, which are indicated. ORF1 codes for a putative protein with a length of 356 amino acids (39,454 Da). The protein contains a helix-turn-helix motif (Fig. 3) present in proteins able to bind to DNA (1, 12). The characteristics of the ORF1 protein indicate that this may be a

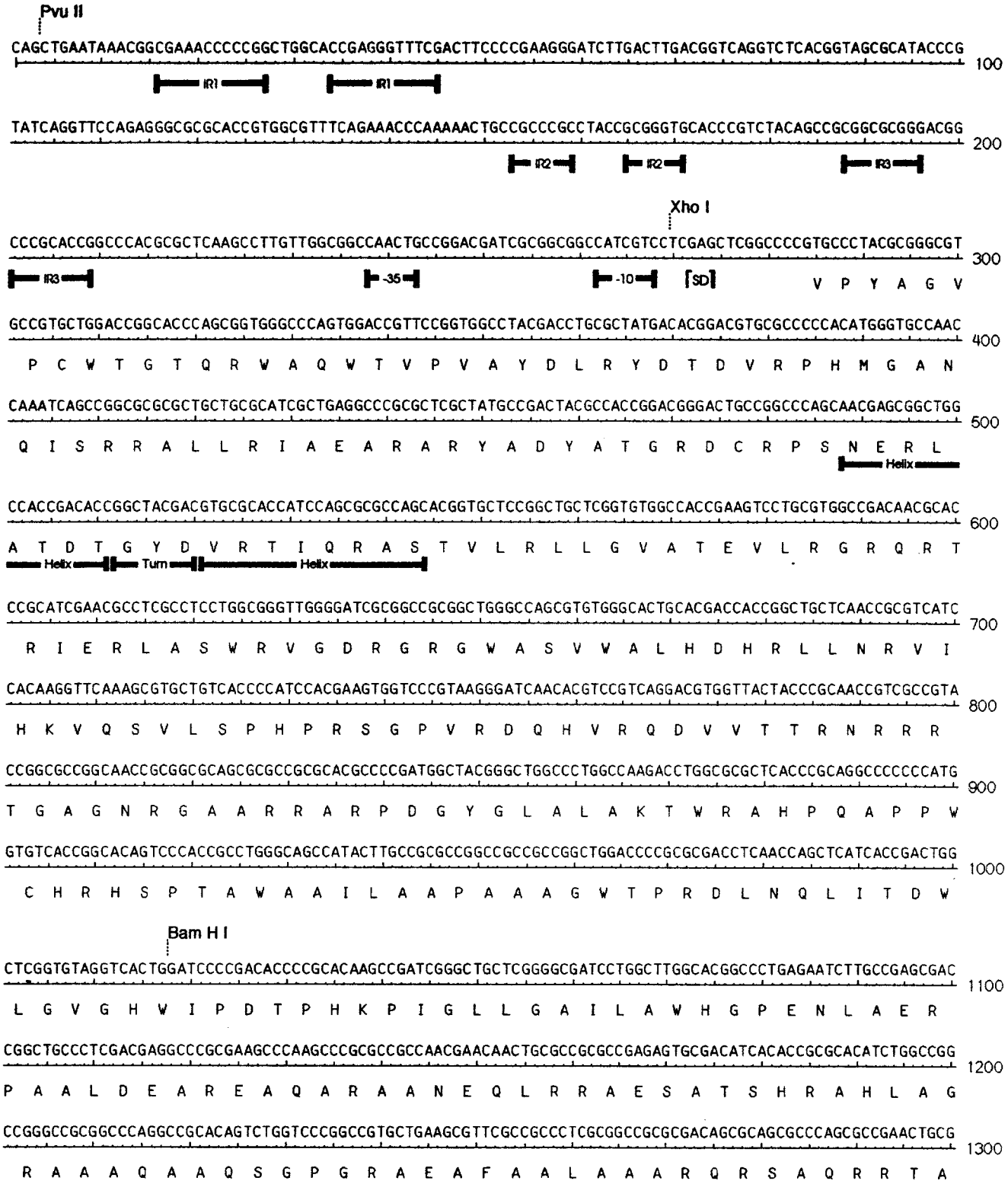

FIG. 3. Complete nucleotide sequence of the pLR7 origin region. IRs are indicated below the sequence. The putative *rep* gene (ORF1) is indicated by the deduced amino acid sequence below the DNA sequence. SD, putative Shine-Dalgarno sequence.

Rep protein (16). Alignment with the Bestfit program (Genetics Computer Group) of the pLR7 Rep protein and the *M. scrofulaceum* pMSC262 Rep protein presented 43% sequence similarity. Since the homology at the DNA level was much more significant, it is probable that the ORF chosen for either pLR7 or pMSC262 was in error. We translated every possible ORF in this region of pLR7 and were unable to find one that correlated with the pMSC262 putative Rep protein. When we

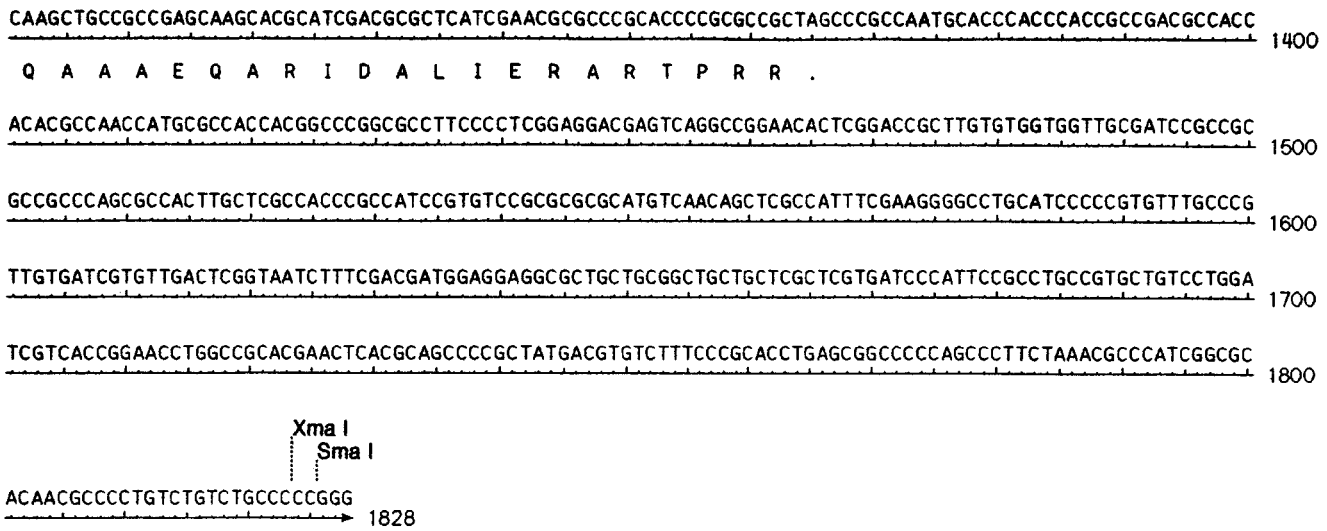

FIG. 3—*Continued.*

translated all possible ORFs in pMSC262, we found one that produced a protein that is 60% identical and 74% similar to the pLR7 ORF1 protein. This ORF is on the opposite strand from the ORF described for pMSC262. Both sequences contain an *Xho*I site preceding the putative *rep* ORF, as well as IR1.

The region of pAL5000 demonstrating homology with pLR7 resides in ORF4, which has been postulated to be an exported protein (14). Alignment of the pAL5000 ORF4 protein product and the protein products of the three putative ORFs (ORF1, ORF6, and ORF9) in the homologous area of pLR7 was performed. The protein product from ORF9 was 24% identical and 56% similar to the pAL5000 protein from ORF4. ORF9 is located on the opposite strand (complement 179 to 703 nucleotides) from the Rep protein, with a termination site upstream of ORF1. Overall, the codon usage of ORF9 is acceptable for a mycobacterial gene, but in a few instances, it does not conform.

The pLR7 ORF, designated ORF4, is located within the region encoding the pLR7 Rep protein but on the opposite strand (complement 721 to 1075 nucleotides). The putative ORF4 protein contains the prokaryotic membrane lipoprotein lipid attachment site, which is found on numerous proteins associated with membranes, including the Col plasmid lysis proteins, *E. coli* plasmid TraT proteins, *Yersinia* virulence plasmid protein YscJ, and the *Shigella* invasion plasmid proteins (8). The codon usage of this gene does not conform well to the published mycobacterial codon usage.

The other ORFs found in this region did not demonstrate mycobacterial codon usage, have distinguishable ribosome binding sites and promoters, or show similarity to any other known sequences.

Construction of a pLR7-based vector. The *Sca*I-to-*Pvu*II fragment of pLR7 was isolated and ligated with pUC18 digested with *Sma*I. A clone was identified in which the *Sma*I site of the pLR7 fragment is nearest to the pUC18 *Hin*dIII site. This clone was digested with *Hin*dIII and *Sma*I and ligated with a *Hin*dIII-*Sma*I fragment containing the kanamycin resistance gene. A transformant carrying both the kanamycin resistance gene and the pLR7 origin of replication was selected. From this plasmid, the *Kpn*I-*Hin*dIII fragment was isolated, end filled to produce two blunt ends, and cloned into pUC18 digested with *Dra*I and *Ssp*I. This removes the ampicillin resistance gene and introduces the pLR7::kanamycin resistance gene. This vector is designated pMB351 (Fig. 5) and contains unique restriction sites for *Eco*RI, *Hin*dIII, *Sal*I, *Sma*I, *Xba*I, and *Kpn*I. The pMB351 vector was successfully transformed into *M. avium* A4, a clinical isolate of *M. tuberculosis*, *M. tuberculosis* H37Rv and H37Ra, *M. bovis* 410, and *M. bovis* BCG Denmark with a transformation efficiency of $10^5/\mu g$ of DNA.

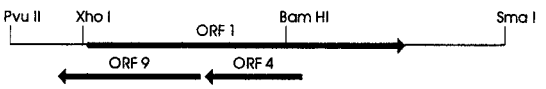

FIG. 4. ORF map of the pLR7 origin region. Arrows indicate the direction of transcription.

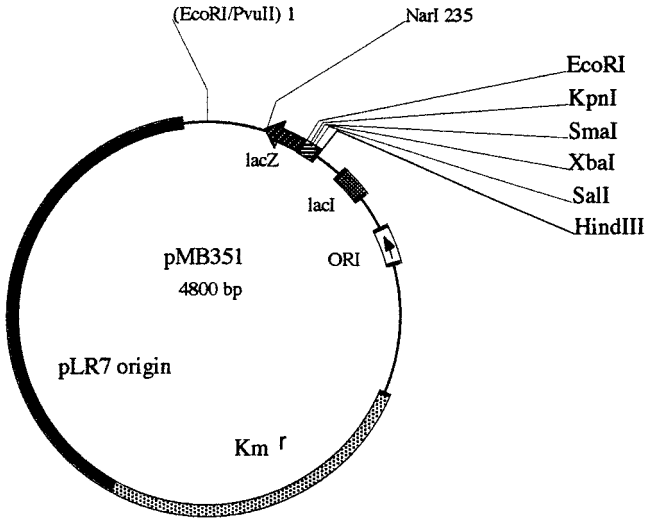

FIG. 5. pMB351 mycobacterium-*E. coli* shuttle vector. The available cloning sites are indicated in the *lacZ* region.

DISCUSSION

Sequence analysis of pLR7 indicates that it has a high degree of homology with the *M. scrofulaceum* plasmid pMSC262, and therefore these plasmids may be members of an incompatibility group (16). Compatibility between pMSC262 and pAL5000 has been demonstrated (13). It is interesting that we were unable to transform *M. smegmatis* LR222 with the plasmid carrying the origin of pLR7. A vector carrying the origin of pMSC262 was found to be unable to replicate in mc^2155, another highly transformable *M. smegmatis* strain, although it did replicate in *M. smegmatis* J15cs (13, 17). The transformable *M. smegmatis* strains LR222 and mc^2155 were derived from the same parent strain (mc^26) and have been found to carry a small cryptic plasmid (11a). This plasmid may be compatible with the pAL5000 origin but incompatible with the pLR7 and pMSC262 origins, resulting in their inability to replicate in these cells. The sequence data of this new *M. smegmatis* plasmid would help elucidate the relationship among the known mycobacterial plasmids. Other possibilities exist for the inability of pLR7 to replicate in LR222, such as restriction modification or the presence of undetected plasmids either free or incorporated into the chromosome. The observation that the LR113 strain carrying only the pLR7 plasmid was nontransformable with pJC85 reiterates that the inability to transform some strains of *M. avium* is complex.

The region of pLR7 containing the origin of replication has many features in common with other plasmids. Some well-characterized *E. coli* plasmids have *rep* genes which are preceded by a region containing many direct repeats and IRs. Repeats have been demonstrated to be important features of origin regions in various *E. coli* and *Salmonella* plasmids (2, 10, 11, 16). The repeats function as binding sites for the plasmid Rep protein for the initiation of plasmid replication. The role of the various repeats upstream of the pLR7 Rep region still needs to be determined, but it is likely that some will be found to be essential for replication or, perhaps, plasmid segregation.

It is interesting that the ORFs chosen for pMSC262 and pLR7 are so different. Since the putative *rep* gene of pLR7 also appears to be present in pMSC262, whereas the pMSC262 *rep* gene is not present in pLR7, we conclude that the pLR7 *rep* gene is more likely the correct one. We are in the process of creating additional subclones and deletions to verify that ORF1 codes for the putative Rep.

The pLR7 ORF9 protein had 56% homology with a putative surface-associated protein produced by the *M. fortuitum* pAL5000 plasmid. A BLAST search with the putative protein sequence from ORF9 resulted in 47% of the 220 sequences identified being related to collagen, most frequently the collagen alpha precursor from a variety of species. The BLAST search also found homology with various *Plasmodium falciparum* surface antigens. If ORF9 does in fact code for a protein, this protein may be associated with the cell surface, as has been postulated for the pAL5000 protein.

ORF4 is also a possible choice for production of a Rep protein, since this ORF spans the *Bam*HI site and production of its protein product would have been disrupted in pMB67. It was not chosen because it lacks the characteristics of Rep proteins and does not have typical mycobacterial codon usage. If a protein is produced by ORF4, it may be associated with the cell membrane, since it contains a motif for the addition of lipid to the protein. We plan to perform experiments to ascertain if any product is produced by ORF4 and the other ORFs in this region of pLR7.

At this time, it is unclear if pLR7 or any mycobacterial plasmid serves to enhance the virulence of *M. avium* species. Since we now have the ability to transform a plasmid-free *M. avium* strain, we can test for differences between the parent and transformed strains. In addition, the availability of vectors will allow genetic manipulation of *M. avium* and will enhance our ability to understand this important opportunistic pathogen.

REFERENCES

1. **Bramhill, D., and A. Kornberg.** 1988. A model for initiation at origins of DNA replication. Cell **54:**915–918.
2. **Brennan, R. G., and B. W. Mathews.** 1989. The helix-turn-helix DNA binding motif. J. Biol. Chem. **264:**1903–1906.
3. **Crawford, J. T., and J. H. Bates.** 1984. Restriction endonuclease mapping and cloning of *Mycobacterium intracellulare* plasmid pLR7. Gene **27:**331–333.
4. **Crawford, J. T., and J. H. Bates.** 1986. Analysis of plasmids in *Mycobacterium avium-intracellulare* isolates from persons with acquired immunodeficiency syndrome. Am. Rev. Respir. Dis. **134:**659–661.
5. **Crawford, J. T., M. D. Cave, and J. H. Bates.** 1981. Characterization of plasmids from strains of *Mycobacterium avium-intracellulare*. Rev. Infect. Dis. **3:**949–952.
6. **Dale, J. W., and A. Patki.** 1990. Mycobacterial gene expression and regulation, p. 173–198. *In* J. McFadden (ed.), Molecular biology of the mycobacteria. Academic Press, Ltd., London.
7. **Hatfull, G. F., and G. J. Sarkis.** 1993. DNA sequence, structure and gene expression of mycobacteriophage L5: a phage system for mycobacterial genetics. Mol. Microbiol. **7:**395–405.
8. **Hayashi, S., and H. C. Wu.** 1990. Lipoproteins in bacteria. J. Bioenerg. Biomembr. **22:**451–471.
9. **Hellyer, T. J., I. N. Brown, J. W. Dale, and C. S. F. Easmon.** 1991. Plasmid analysis of *Mycobacterium avium-intracellulare* (MAI) isolated in the United Kingdom from patients with and without AIDS. J. Med. Microbiol. **34:**225–231.
10. **Kües, U., and U. Stahl.** 1989. Replication of plasmids in gram-negative bacteria. Microbiol. Rev. **53:**491–516.
11. **Manen, D., and L. Caro.** 1991. The replication of plasmid pSC101. Mol. Microbiol. **5:**233–237.

11a. **McCarthy, C. (New Mexico State University).** Personal communication.

12. **Pabo, C. O., and R. T. Sauer.** 1984. Protein-DNA recognition. Annu. Rev. Biochem. **53:**293–321.
13. **Qin, M., H. Taniguchi, and Y. Mizuguchi.** 1994. Analysis of the replication region of a mycobacterial plasmid, pMSC262. J. Bacteriol. **176:**419–425.
14. **Ranes, M. G., J. Rauzier, M. Lagranderie, M. Gheorghiu, and B. Gicquel.** 1990. Functional analysis of pAL5000, a plasmid from *Mycobacterium fortuitum*: construction of a "mini" mycobacterium-*Escherichia coli* shuttle vector. J. Bacteriol. **172:**2793–2797.
15. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Scott, J. R.** 1984. Regulation of plasmid replication. Microbiol. Rev. **48:**1–23.
17. **Snapper, S. B., R. E. Melton, S. Mustafa, T. Kieser, and W. R. Jacobs, Jr.** 1990. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. Mol. Microbiol. **4:**1911–1919.