12-1-1991

A Novel Transposon Trap for Mycobacteria: Isolation and Characterization of IS*1096*

Jeffrey D. Cirillo
Albert Einstein College of Medicine, Bronx, New York, jdcirillo@medicine.tamhsc.edu

Raul G. Barletta
University of Nebraska - Lincoln, rbarletta1@unl.edu

Barry R. Bloom
Albert Einstein College of Medicine, Bronx, New York

William R. Jacobs Jr.
Albert Einstein College of Medicine, Bronx, New York

JOURNAL OF BACTERIOLOGY, Dec. 1991, p. 7772–7780
0021-9193/91/247772-09$02.00/0

Vol. 173, No. 24

A Novel Transposon Trap for Mycobacteria: Isolation and Characterization of IS*1096*

JEFFREY D. CIRILLO, RAÚL G. BARLETTA,† BARRY R. BLOOM, AND WILLIAM R. JACOBS, JR.*

Howard Hughes Medical Institute, Department of Microbiology and Immunology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461

Received 12 July 1991/Accepted 7 October 1991

In the course of developing strategies to obtain a mutation in the aspartate semialdehyde dehydrogenase (*asd*) gene of *Mycobacterium smegmatis*, an efficient transposon trap was constructed which may be generally useful for the identification of transposable elements in mycobacteria. A DNA fragment containing the *asd* gene was replaced with an aminoglycoside phosphotransferase gene (*aph*) to generate a Δ*asd*::*aph* allele. Attempts to replace the wild-type *asd* gene with the Δ*asd*::*aph* allele were unsuccessful, suggesting that this deletion was lethal to the growth of *M. smegmatis*. The plasmid, pYUB215, which contains β-galactosidase expressed from a mycobacteriophage promoter and Δ*asd*::*aph*, was integrated into the chromosome of *M. smegmatis* by a homologous, single-crossover, recombination event. Visual screening for inactivation of the β-galactosidase gene in the resulting strain allowed the isolation of a novel mycobacterial insertion element from *M. smegmatis*. This insertion element, which is unique to *M. smegmatis*, was designated IS*1096* and transposes at a frequency of 7.2×10^{-5} per cell in an apparently random fashion. IS*1096* is 2,275 bp in length and contains two open reading frames which are predicted to encode proteins involved in transposition. This insertion element exhibits several characteristics that suggest it may be a useful tool for genetic analysis of mycobacteria, possibly including the study of mechanisms of pathogenesis.

Significant progress has been made in recent years toward understanding of the biology of mycobacteria with the development of genetic systems that permit the efficient introduction of recombinant DNA into mycobacterial cells (18, 35). There now exist a wide range of shuttle phasmid (18), shuttle plasmid (30, 35), and integration-proficient (21) vectors that can be used to transfer recombinant DNA between *Escherichia coli* and both fast- and slow-growing species of mycobacteria. Efficient transformation mutants of the fast-growing *Mycobacterium smegmatis* have been isolated (36) that provide a useful host for analyzing genes and gene libraries from slow-growing mycobacteria such as *M. leprae* (40), *M. tuberculosis* (19), and *M. avium* (3). Moreover, these systems have been used to develop vectors that permit the expression of genes encoding foreign antigens in *M. bovis* bacillus Calmette-Guérin (BCG), which can be used as a live recombinant vaccine vehicle for inducing multiple immune responses to foreign antigens (1, 38).

The generation of insertional mutations has provided a powerful approach for the study of bacterial genetics (44) and pathogenesis (5, 10, 16, 39). Gene replacement has been demonstrated recently for *M. smegmatis* (15). In addition, our laboratory has reported the successful isolation of auxotrophic mutants in *M. smegmatis* by using random shuttle mutagenesis (19). When this approach was extended to BCG or *M. tuberculosis*, a surprising result was obtained. The slow-growing mycobacteria randomly incorporated linear DNA into their chromosomes by illegitimate recombination to generate insertional mutations. Numerous repetitive DNA elements have been found for the mycobacteria, including *M. paratuberculosis* (25), *M. tuberculosis* (9, 24, 28, 41), and *M. leprae* (13). Sequence analysis has revealed that both the *M. paratuberculosis* (12) and *M. tuberculosis* (24, 42) repetitive elements share homologies with other known insertion elements, but transposition in mycobacteria has only recently been demonstrated with an *M. fortuitum* transposon (22).

In this work, we describe the construction of a strain of *M. smegmatis* containing a β-galactosidase reporter gene that facilitates the isolation of mutations generated by insertions of mobile genetic elements. This strain has been used to identify and characterize a novel insertion element of *M. smegmatis*. This element, designated IS*1096*, has been sequenced and shows promise as a useful tool for the genetic analysis of mycobacteria.

(The data presented are part of a thesis to be submitted by J.D.C. in partial fulfillment of the requirements for a Ph.D. from the Albert Einstein College of Medicine, Yeshiva University, New York.)

MATERIALS AND METHODS

Bacterial strains and media. Bacterial strains used in these experiments are described in Table 1. It was found that the ATCC 607 culture contained three unique morphotypes when plated on R5 regeneration media (14). The predominant morphotype, which had an orange-rough colonial morphology, was cloned as a single cell and designated mc^26. Two minor morphotypes were cloned in the same manner, the orange-smooth designated mc^223 and the white-rough designated mc^222. *E. coli* strains were grown in Luria broth (LB) or on LB agar. Mycobacterial strains were grown in M-ADC-TW broth (17) and plated on Middlebrook 7H9 agar containing 0.01% cycloheximide (M-ADC). Kanamycin and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) were added at 10 and 80 μg/ml, respectively, when needed in either broth or plates containing M-ADC-KX agar. Higher concentrations of X-Gal than the standard, 40 μg/ml, were used to allow easier differentiation between blue and white

* Corresponding author.
† Present address: Department of Veterinary Sciences, University of Nebraska, Lincoln, NE 68583-0905.

TABLE 1. Bacterial strains

Strain	Description	Source
E. coli DH5α	F⁻ φ80d*lacZ*ΔM15 *endA1 recA1 hsdR17 glnV thi-1* λ⁻ *gyrA96 relA1* Δ(*lacZ YA-argF*)U169	Bethesda Research Laboratories
M. smegmatis		
$mc^2 1$	ATCC 607 *M. smegmatis*	W. Jones, CDC[a]
$mc^2 6$	Orange-rough variant of $mc^2 1$.	17
$mc^2 155$	Efficient plasmid transformation (Ept) mutant of $mc^2 6$	35
$mc^2 22$	White-rough variant of $mc^2 1$	This work
$mc^2 23$	Orange-smooth variant of $mc^2 1$	This work
$mc^2 31$	ATCC 27204 *M. smegmatis* SN2	W. Jones, CDC
$mc^2 32$	ATCC 27199 *M. smegmatis* L1	W. Jones, CDC
$mc^2 687$	Integration of pYUB215 into $mc^2 6$ by homologous recombination at *asd*	This work

[a] CDC, Centers for Disease Control.

colonies. Casamino Acids and diaminopimelate (DAP; a mixture of DD, LL, and *meso* isomers of DAP) were added at 40 and 100 μg/ml, respectively, when necessary.

DNA manipulations. All enzyme reactions were carried out under the conditions recommended by the manufacturers (Promega, New England Biolabs, Boehringer Mannheim, and Bethesda Research Laboratories). Isolation of plasmid DNA and physical manipulations of this DNA were carried out for *E. coli* as described previously (33). Isolation of mycobacterial total chromosomal DNA was carried out by the beadbeater method (17). Transformations of plasmids into *E. coli* and *M. smegmatis* were carried out by electroporation (8, 17).

Construction of pYUB215. The cloning of the *M. smegmatis* aspartate semialdehyde dehydrogenase (*asd*) gene which was used in the construction of pYUB215 will be described elsewhere (7). The *asd* gene is carried on a 5-kb *Eco*RI fragment which has two *Pst*I sites in the middle of the gene and two *Pst*I sites flanking the gene. The *asd* gene was cloned into the *Eco*RI site of pBluescript II KS⁺ (Stratagene, La Jolla, Calif.), and the resulting plasmid was designated pYUB114. The plasmid pYUB114 was digested with *Pst*I and gel purified. The gel-purified DNA fragment was then ligated to a 1.4-kb *Pst*I cassette which contains the Tn*903* aminoglycoside phosphotransferase (*aph*) gene (Pharmacia). Transformants with this ligation were screened for replacement of the central 1.6-kb *Pst*I fragment that contains the *asd* gene. The resulting plasmid, designated pYUB205, has a deletion of the *asd* gene replaced by *aph* (Δ*asd*::*aph*). The plasmid pYUB205 was cut with *Eco*RI, and the large fragment was isolated from an agarose gel. This fragment was then ligated to pYUB174 which had been cut with *Eco*RI and treated with calf intestinal phosphatase (Boehringer Mannheim, Indianapolis, Ind.). The plasmid pYUB174 which contains the β-galactosidase gene was constructed by placing the previously isolated mycobacteriophage L1 promoter (2) upstream of a truncated β-galactosidase gene to allow expression of β-galactosidase in mycobacteria. The resulting plasmid, pYUB215, contains Δ*asd*::*aph* and β-galactosidase driven from the L1 promoter.

Southern blot analyses. Digests of total chromosomal DNA were electrophoresed on 0.8% agarose gels and transferred as previously described (37) to Biotrans nylon membranes (ICN, Irvine, Calif.). Hybridizations were carried out at 65°C overnight. The blots were first washed at room temperature twice for 30 min in 2× SSPE (1× SSPE is 0.18 M NaCl, 10 mM $NaPO_4$, and 1 mM EDTA [pH 7.7]) and 0.2% sodium dodecyl sulfate (SDS) and then for 1 h at 65°C in 0.2× SSPE and 0.5% SDS. After washing, the blots were exposed to film overnight (XAR 5; Eastman Kodak Co., Rochester, N.Y.). DNA probes were prepared by using the nick translation system from Bethesda Research Laboratories.

Construction of $mc^2 687$. To integrate pYUB215 into the chromosome of *M. smegmatis*, the plasmid was transformed by electroporation into $mc^2 6$. This transformation was then plated on M-ADC plates containing kanamycin and X-Gal (M-ADC-KX). Blue colonies were picked and screened by Southern analysis to establish integration of pYUB215. One of these recombinants was designated $mc^2 687$.

Isolation of β-galactosidase mutants of $mc^2 687$. Ten individual colonies were picked from an $mc^2 687$ plate and grown in 5 ml of M-ADC-TW broth with kanamycin until saturation. Dilutions of these 10 independent cultures were plated on M-ADC-KX plates such that approximately 1,000 colonies per plate were obtained. The number of white colonies divided by the total number of colonies present represents the frequency of isolation of mutations in the β-galactosidase gene of $mc^2 687$.

Isolation of IS*1096* and sequence analyses. This system facilitates isolation of the insertion element once it is observed. The insertion element which is inserted in the β-galactosidase gene of the integrated plasmid, pYUB215, is recovered by cutting total chromosomal DNA with *Eco*RI to completion. Digestion with *Eco*RI will free the original pYUB174 plasmid that Δ*asd*::*aph* was cloned into which contains a copy of the insertion element. To isolate IS*1096*, *Eco*RI-digested DNA was self ligated at a DNA concentration of less than 5 ng/μl, transformed into *E. coli* DH5α, and selected by ampicillin resistance.

A single clone containing IS*1096* was isolated in this manner and designated pYUB209. The IS*1096* insertion in pYUB209 was localized by restriction analysis within the β-galactosidase gene, between two *Hpa*I sites. To facilitate sequencing of the insertion element, the *Hpa*I fragment was cloned in both orientations into the *Sma*I site of pGEM7Zf⁺ (Promega). Several deletions were constructed in pYUB209 by using the opportune sites *Bst*XI and *Bam*HI within IS*1096* and the polylinker of pGEM7Zf⁺. These deletions allowed sequencing in both directions from these sites into IS*1096*. IS*1096* and the junction between IS*1096* and pYUB209 were sequenced completely on both strands by using synthetic oligonucleotides in addition to these deletions. Sequencing reactions were carried out by using Sequenase Version 2.0 (U.S. Biochemical Corp.) on double-stranded DNA templates (20).

Nucleotide sequence accession number. The GenBank accession number for the nucleotide sequence of insertion element IS*1096* is M76495.

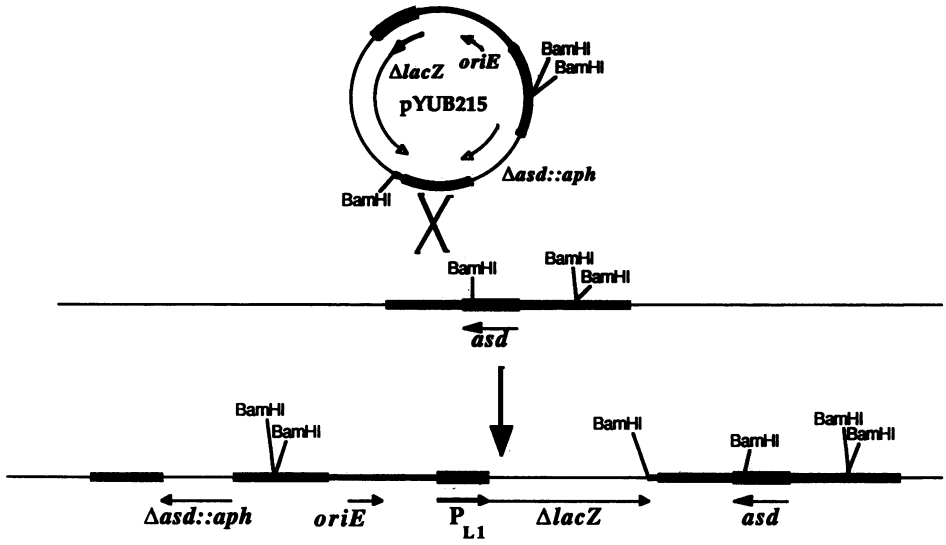


FIG. 1. Construction of a genetic screen for homologous recombination in *M. smegmatis*. The plasmid pYUB215 contains a truncated *lacZ* gene from *E. coli* which is driven from a mycobacteriophage L1 promoter, an *E. coli* origin which does not function in mycobacteria, the *E. coli* ampicillin resistance gene which is not useful in mycobacteria, and the Δ*asd*::*aph* region which consists of a 5-kb fragment of the *M. smegmatis* chromosome where the central 1.6-kb *Pst*I DNA fragment that contains the *asd* gene has been replaced by the *aph* gene from Tn*903*. Homologous recombination may occur between the *asd* flanking regions on pYUB215 and the homologous regions on the *M. smegmatis* chromosome. The recombination event shown here is the same as that observed in mc^{2}687.

RESULTS

Construction of a chromosomal transposon trap for *M. smegmatis*. The construction that served as an efficient transposon trap was designed initially to allow isolation of a mutant of *M. smegmatis* containing a deletion-insertion mutation of the *asd* gene. We cloned the *M. smegmatis asd* gene by complementing an *asd* mutant of *E. coli* and determined the nucleic acid sequence (7). By using this gene, a plasmid, pYUB215, was constructed containing the following: (i) the *aph* gene cloned between two *Pst*I sites within the *asd* region that created a 1.6-kb deletion (Δ*asd*::*aph*), (ii) approximately 2 kb of *M. smegmatis* DNA flanking each side of the *asd* gene, (iii) the *E. coli* β-galactosidase gene fused to a mycobacteriophage promoter, and (iv) the ColE1 origin of replication and β-lactamase gene from pBR322 (6). This plasmid should provide an ideal system for allelic exchange in mycobacteria since it cannot replicate autonomously in mycobacteria, has a positive selection for the desired mutation (*aph*), and allows for differentiation between single-crossover recombination events and allelic exchange by using the β-galactosidase gene.

When the plasmid pYUB215 was initially transformed into mc^{2}6, blue and kanamycin-resistant colonies were obtained at a frequency of 10^{-5} to 10^{-6} per transformant obtained from 1 μg of the control plasmid pMV262 (38) on M-ADC-KX agar containing DAP and Casamino Acids. No significant numbers of white colonies were obtained above a background of 1 to 5 colonies, attributed to spontaneous kanamycin-resistant mutants. The blue colonies obtained were thought to be generated by a single-crossover recombination event in which pYUB215 recombined into the homologous DNA sequences located on either side of the *aph* gene (Fig. 1). Southern analysis of eight independent transformants confirmed that pYUB215 had integrated on either side of the *aph* gene, and one such transformant, mc^{2}687, was used for further analysis (Fig. 2). These results may indicate that the *asd* gene is essential in mycobacteria, since allelic exchange experiments in mycobacteria have previously been successful (15), or that double-crossover events occur at an extremely low frequency in the *asd* region of *M. smegmatis*. The necessity of having simultaneous double-crossover events can be obviated by constructing the chromosomal Δ*asd*::*aph* mutant in a two-step process. The strain mc^{2}687 has already undergone a single-crossover event to integrate pYUB215; thus, a second recombination event between the two alleles in the chromosome may produce the desired mutant. We reasoned that it should be possible to screen for intrachromosomal recombination by the loss of β-galactosidase when the plasmid sequences are looped out during this event. The strain mc^{2}687 was grown in M-ADC-TW with kanamycin, and aliquots were plated on M-ADC agar, with the addition of kanamycin and X-Gal, and screened for the production of white colonies.

Identification of insertion elements by loss of β-galactosidase activity. White mutants of mc^{2}687, having lost β-galactosidase activity, were obtained at a frequency of 8×10^{-5} per cell. *Bam*HI digests of chromosomal DNA from these clones were compared with mc^{2}687 by Southern analysis (Fig. 2). The banding patterns of each of the eight clones shown may be explained by insertion of an approximately 2-kb fragment into the β-galactosidase region. The inserted fragment contains a *Bam*HI site, and thus all clones show a shift of the largest band, containing β-galactosidase, and an additional smaller band which ranges in size from approximately 2 to 5 kb. The fragment containing the β-galactosidase gene was isolated by cutting total genomic DNA with *Eco*RI and cloning into *E. coli* as described above. The resulting plasmid from one of these clones was used for restriction and sequence analysis.

Characterization of IS*1096*. The approximate positions of eight insertions into the β-galactosidase gene are shown in Fig. 3. From the distribution of insertions in the β-galactosidase gene, IS*1096* appears to transpose in a random fashion. IS*1096* was also used as a probe in Southern analysis

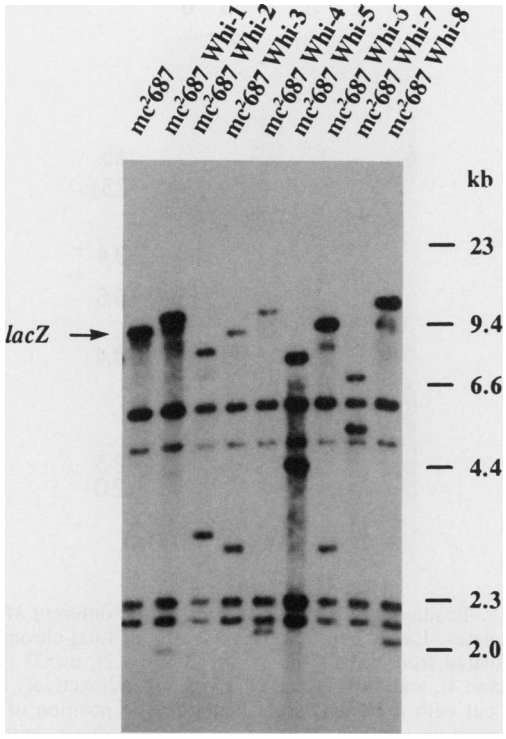


FIG. 2. Southern analysis of white mutants of mc²687. All lanes are *Bam*HI digests of total chromosomal DNA from mc²687 and white-colony mutants of mc²687 probed with pYUB215. From the top of the blot, in the mc²687 lane, the second and fifth bands are flanking *Bam*HI fragments to the integrated pYUB215 and the first, third, and fourth bands correspond to internal fragments from pYUB215 (first) or the chromosomal *asd* gene (third and fourth). This pattern is consistent with that expected for homologous recombination as shown in Fig. 1. The *lacZ* gene-containing fragment is the largest band (indicated by an arrow) at approximately 9.5 kb. All white mutants show a shift in the size of the *lacZ* band and an extra smaller band. This smaller band added to the *lacZ* band results in a fragment of approximately 11.5 kb which can be explained by insertion of an approximately 2.2-kb insertion element, which contains a *Bam*HI site, into the *lacZ* gene.

TABLE 2. Determination of IS*1096* transposition frequency

Culture no.	Frequency of occurrence of white colonies[a]
1	4.04×10^{-5}
2	1.29×10^{-4}
3	8.03×10^{-5}
4	7.61×10^{-5}
5	4.31×10^{-5}
6	1.47×10^{-4}
7	4.26×10^{-5}
8	3.94×10^{-5}
9	4.11×10^{-5}
10	7.92×10^{-5}
Mean	7.2×10^{-5}

[a] Frequency of loss of β-galactosidase activity in cultures from 10 independent colonies of mc²687. Frequencies are expressed as the number of white colonies observed for every one blue colony, which represents the number of insertions into β-galactosidase for every normal mc²687 cell.

comparing the white clones isolated with the parent strain mc²687. It was observed that 15 bands were present in both strains, with the white clones showing an additional band which varies in size with the varying position of the insertion into the β-galactosidase gene (7).

To more accurately determine the frequency of IS*1096* transpositions in *M. smegmatis*, 10 individual colonies of mc²687 were grown to stationary phase in M-ADC-TW broth with kanamycin. Approximately 1,000 cells (as determined by A_{600}) from each culture were plated on 20 M-ADC-KX agar plates to screen for white colonies. The average frequency of loss of β-galactosidase activity in these cultures was 7.2×10^{-5} per cell (Table 2). This number includes only those colonies which proved to be truly negative for β-galactosidase activity when retested on X-Gal-containing plates.

IS*1096* is unique to *M. smegmatis*. Southern analysis of *Streptomyces coelicolor*, *E. coli* K-12, *Bacillus subtilis*, and 12 strains of mycobacteria was performed by using the internal IS*1096* fragment as a probe (Fig. 4). IS*1096* is present only in *M. smegmatis*. It is important to note that the pathogenic strains of mycobacteria, *M. tuberculosis*, *M. bovis*, *M. avium*, and *M. leprae*, lack the insertion element.

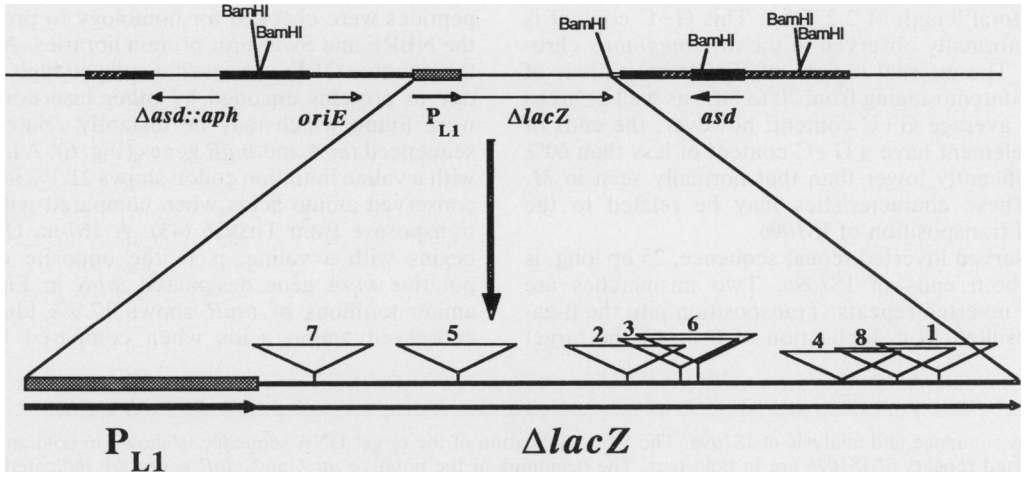


FIG. 3. Distribution of IS*1096* insertions in the *lacZ* gene. Approximate positions of eight IS*1096* transpositions into the pYUB215 β-galactosidase gene as determined from Southern analysis are shown. *Bam*HI fragments shown in Fig. 2 were plotted against standards to allow estimation of sizes. Restriction mapping of IS*1096* gave the orientation of the *Bam*HI site within the insertion element, allowing accurate approximation of the positions of the various insertions into the *lacZ* gene.

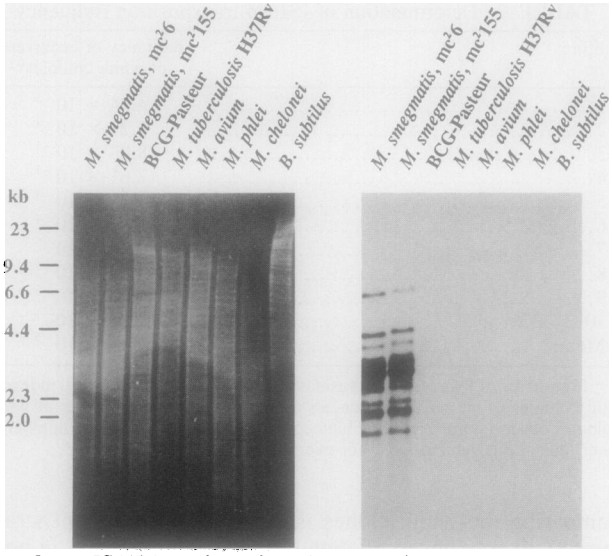


FIG. 4. IS*1096*, an insertion element unique to *M. smegmatis*. An agarose gel and Southern blot of total chromosomal DNA from various strains of mycobacteria and a single strain of *B. subtilis* digested to completion with *Pst*I are shown. The blot was probed with an internal fragment of IS*1096* (see Fig. 6) which should hybridize to only one arm of the insertion element. DNA concentrations appear to be relatively the same, as seen from the agarose gel. Only the *M. smegmatis* strains hybridize to the IS*1096* probe.

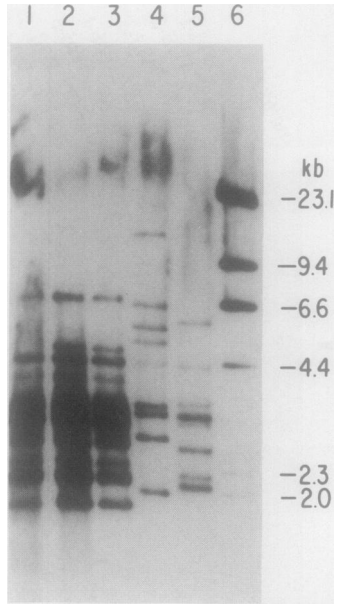


FIG. 5. Positions of IS*1096* in the genome of different *M. smegmatis* isolates. Lanes 1 to 5 are *Pst*I digests of total chromosomal DNA isolated from mc²6 (lane 1), mc²22 (lane 2), mc²23 (lane 3), mc²31 (lane 4), and mc²32 (lane 5). Lane 6 is radioactively labelled λ DNA cut with a *Hin*dIII size standard. The position of IS*1096* varies a great deal in the various *M. smegmatis* isolates, with one or possibly no conserved bands across all isolates.

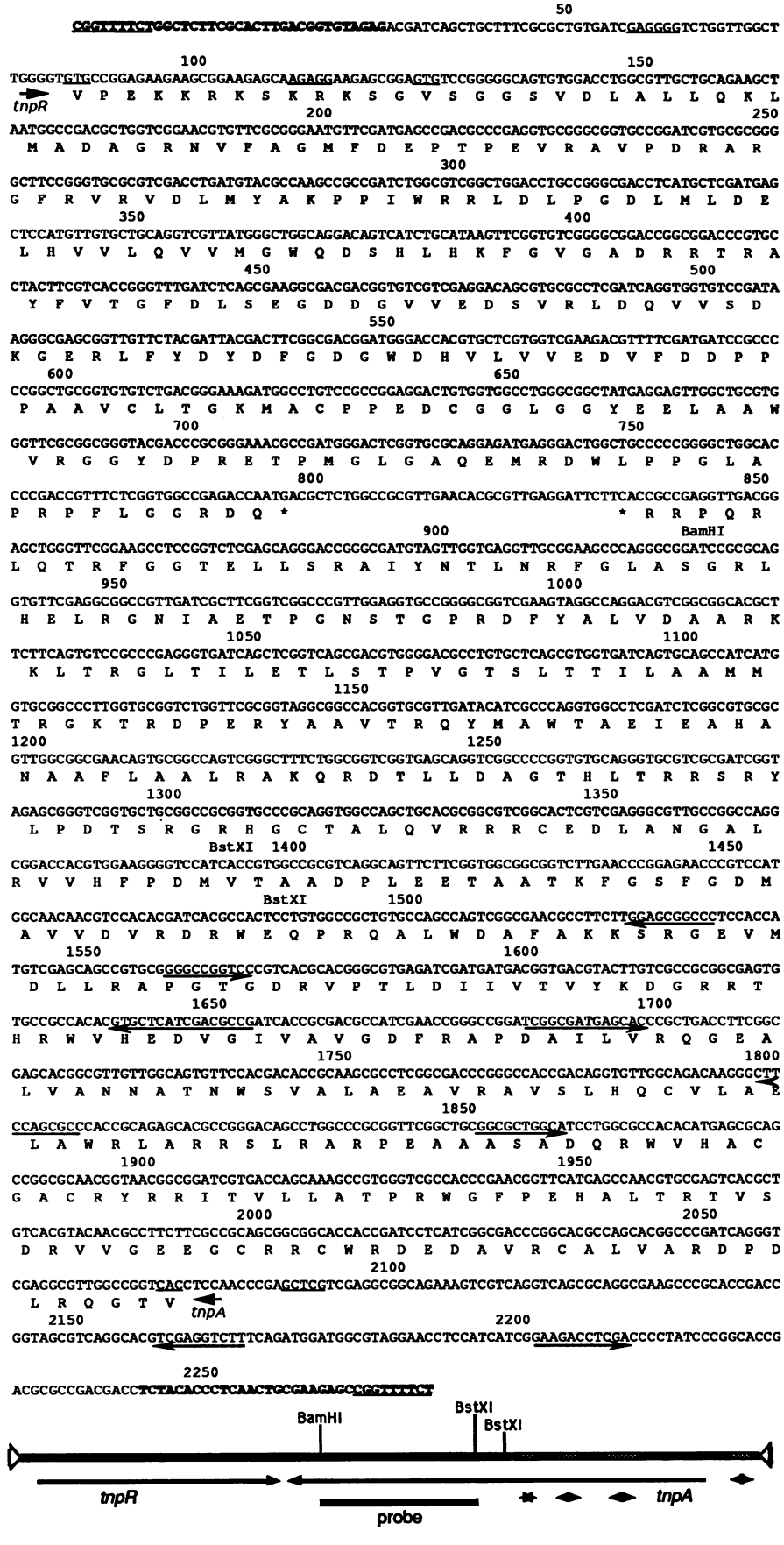
Several different *M. smegmatis* isolates, including the three morphotypes of ATCC 607, were studied by Southern analysis to determine the degree of variation in restriction pattern when probed with IS*1096* (Fig. 5). There is a great deal of variability in the restriction patterns of the different isolates of *M. smegmatis*. Only one of the bands appears in all of the isolates tested, and the number of insertion elements found ranges from 8 to 16. Both mc²22 and mc²23 have an additional copy of IS*1096* not present in mc²6 which may be responsible for their observed phenotypic variation. However, mc²155, a high-efficiency transformation mutant of mc²6, shows the same pattern as mc²6.

Nucleotide sequence of IS*1096*. Sequence analysis revealed that the guanine plus cytosine (G+C) content in IS*1096* is 67% over its total length of 2,275 bp. This G+C content is equal to that normally observed in the *M. smegmatis* chromosome (45). The internal regions of IS*1096* have areas of higher G+C content ranging from 70 to 80% as well as areas similar to the average G+C content; however, the ends of the insertion element have a G+C content of less than 60% which is significantly lower than that normally seen in *M. smegmatis*. These characteristics may be related to the mechanism of transposition of IS*1096*.

A well-conserved inverted repeat sequence, 25 bp long, is observed at both ends of IS*1096*. Two mismatches are present in the inverted repeats. Transposition into the β-galactosidase resulted in a duplication of 8 bp of the target DNA sequences on both sides of the insertion point. This duplication is consistent with the mechanism of transposition of most insertion elements. Three sets of inverted repeats approximately 9 bp in length, the positions of which are shown in Fig. 6, were found in the right end of IS*1096*. There is also a fourth set of inverted repeats which is seen in the very 3′ end of the insertion element which corresponds to a region just upstream of a large open reading frame (ORF).

Identification of putative functional ORFs. ORF analysis of the sequence of IS*1096* revealed the presence of 13 ORFs which are longer tnan 100 amino acids (aa). All 13 predicted peptides were checked for homology to proteins present in the NBRF and Swissprot protein libraries. Although none of the putative ORFs observed displayed high levels of homology to proteins encoded by other insertion elements, two were found which may be distantly related to previously sequenced *tnpA* and *tnpR* genes (Fig. 6). A large 414-aa ORF with a valine initiation codon shows 21.1% identical and 66% conserved amino acids when compared with 185 aa of the transposase from Tn*3926* (43). A 167-aa ORF, which also begins with a valine, is in the opposite direction to the putative *tnpA* gene designated *tnpR* in Fig. 6. The 84-aa amino terminus of *tnpR* shows 17.9% identical and 69% conserved amino acids when compared with the amino

FIG. 6. DNA sequence and analysis of IS*1096*. The 8-bp duplication of the target DNA sequence is shown in bold and underlined text. The 25-bp inverted repeats of IS*1096* are in bold text. The beginning of the putative *tnpA* and *tnpR* genes are indicated by labelled short arrows. Possible initiation codons and ribosomal binding sites are indicated by underlined text. Pairs of arrows in opposite orientation designate inverted repeats which were found throughout the sequence. There are one set of inverted repeats in the putative promoter region of *tnpA* and three sets in the putative coding region of *tnpA*. The internal probe which was used for Southern analysis in Fig. 4 and 5 is shown as a solid bar between the *Bam*HI and *Bst*XI sites in the graphic representation of the sequence.


tnpR
tnpA
BamHI
BstXI
BstXI
BamHI
BstXI
tnpR
probe
tnpA

terminus of the Tn*1000* resolvase (31). The 72-aa carboxy terminus shows 16.7% identical and 67% conserved amino acids when compared with the Tn*552* resolvase (32). These levels of homology are on the borderline of what would be considered significant homology (29); however, the proteins encoded by these ORFs are of a similar size and significantly more similar than the other ORFs found on IS*1096* to standard transposase and resolvase proteins. This may indicate that these ORFs do encode proteins which are involved in the transposition of IS*1096* but are distantly related to the well-characterized proteins known to be involved in transposition of other transposable elements. It may be possible that the mechanism of transposition of IS*1096* is significantly different from that of previously studied insertion elements.

DISCUSSION

We have isolated an insertion element, IS*1096*, from *M. smegmatis* that is of potential use for insertional mutagenesis of mycobacteria. IS*1096* was discovered as the result of attempts to isolate mycobacterial mutants in *asd*. A mycobacterial *asd* mutant may be a useful strain for the maintenance of recombinant plasmids in vivo, as shown in salmonella (27). The plasmid pYUB215 has the advantage of containing the β-galactosidase gene as a marker for the plasmid sequences, which allows screening for allelic exchange during intrachromosomal recombination by the observation of white colonies on X-Gal. In other bacteria, intrachromosomal recombination between large regions of homology occurs at frequencies which range from 1 in 100 to 1 in 1,000 (11, 26, 34). It is possible that the frequency of intrachromosomal recombination in *M. smegmatis* is much lower than that normally seen and will not be observed above the background loss of β-galactosidase activity because of insertion of IS*1096*. However, no allelic exchange was observed when pYUB215 was originally introduced into *M. smegmatis*, an event that should be independent of intrachromosomal recombination frequencies. Although there have been attempts at isolating mutants involved in DAP biosynthesis in other gram-positive bacteria (46), they have not been successful. This may point toward differences in the requirements for DAP or the ability to transport DAP across the cell wall of gram-positive bacteria (46) as opposed to *E. coli* and *Salmonella* spp. (*asd* mutants are readily obtained for *E. coli* and *Salmonella* spp. [27]). Therefore, we favor the interpretation that the *asd* gene is essential in mycobacteria. This would make it impossible to select for intrachromosomal recombination producing the chromosomal *asd* mutant in the absence of a complementing *asd* gene. We are currently testing this hypothesis by constructing the chromosomal *asd* mutant in a merodiploid strain.

Since the β-galactosidase gene is stably maintained in the chromosome, it allows observation of low-frequency events such as chromosomal rearrangements or mutagenesis due to insertion elements similar to IS*1096*. Thus, the utilization of this marker system may have a general application as a genetic trap for insertion elements. It remains to be seen whether this system will allow isolation of insertion elements in other species of mycobacteria. A small percentage (~1 to 10%) of the white mutants isolated from mc^{2}687 could not be explained by the insertion of IS*1096* (7). It is possible that there is another smaller insertion element in *M. smegmatis* which could explain these mutants. The use of the *asd* genes from other mycobacteria in this type of system could allow isolation of insertion elements from any mycobacteria for which there is an efficient transformation system. Useful insertion elements for the development of mutagenesis systems for mycobacteria should be of a relatively small size, have a high frequency of transposition, not be present in the bacterial strains where they will be used for mutagenesis, and exhibit no site or regional specificity.

The only insertion element other than IS*1096* which has been shown to transpose in mycobacteria is IS*6100* from *M. fortuitum*. IS*6100* has three limitations which may make it a secondary choice for mutagenesis of mycobacteria: the transposase present in this element has significant homology to the transposase of Tn*1696* (22) which does not transpose in a random fashion (23); the frequency of transposition was fairly low (10^{-6} to 10^{-7}) and transposition occurs only by formation of a cointegrate structure which is not resolved after transposition (22); and the insertion element is fairly large, approximately 4 kb, an inconvenient size for the design of transposon delivery systems. Of the insertion elements isolated thus far in mycobacteria, IS*1096* shows the greatest promise for the development of a transposon delivery system since it appears to have most of the desired characteristics.

IS*1096* transposes in a random fashion which rivals that of the most random elements seen in gram negatives (4). The frequency of transposition of IS*1096* is high (10^{-5}). The size of IS*1096*, 2.2 kb, is very manageable. Its small size should enable development of transposon delivery systems where the size of the fragment which can be delivered is limited. There are, however, a few concerns which could affect the usefulness of this element in mycobacterial mutagenesis systems. Since all of the transposition events observed thus far have been into the β-galactosidase gene, it is possible that IS*1096* has some preference for this gene over others (possibly because of the lower G+C content of the β-galactosidase gene compared to that of the mycobacterial chromosome in general), and, thus, it may have some regional preference. However, if this were the case, one would expect to observe regional specificity within β-galactosidase itself. A second concern is that IS*1096* has not yet been tested in the pathogenic mycobacteria where its random transposition mechanism would be most useful for studies of virulence. There is, however, no reason to believe that, if an insertion element functions in one species of mycobacteria, it will not function in the other mycobacteria. Furthermore, the fact that it is absent from pathogenic mycobacteria will facilitate analysis of mutants which are isolated in these strains. At present, studies are ongoing to determine whether IS*1096* will transpose in other species of mycobacteria and develop an efficient transposon delivery system to facilitate efficient mutagenesis of mycobacteria. It is expected that IS*1096* will be a valuable tool in the study of mycobacterial molecular genetics.

ACKNOWLEDGMENTS

We thank Nigel Grindley and Graham Hatfull for discussion and critical review of the manuscript as well as assistance with sequence analysis, Jeanie Dubnau, Ruth McAdam, Nancy Connell, and Ken Stover for helpful discussions and gifts of bacterial DNA, and Torin Weisbrod for excellent technical assistance.

This work was supported by grants AI23545, AI26941, AI27235, and AI26170 from the National Institutes of Health, the Rockefeller Foundation, the Immunology of Tuberculosis Committee of the World Health Organization/UNDP Special Programme for Vaccine Development, the Immunology and Chemotherapy Committees of the UNDP/World Bank/World Health Organization Special Programme for Research and Training in Tropical Diseases, and National Cancer Institute Training Grant 5T32CA09173.

ADDENDUM IN PROOF

By using a plasmid-based transposon trap, a different insertion element, IS*6120*, has been identified in the same *M. smegmatis* strain, mc^{2}155 (C. Guilhot, B. Gicquel, J. Davies, and C. Martin, Mol. Microbiol, in press). IS*1620* clearly differs from IS*1096*, as it is 1,486 bp in length and appears to be present in three copies in the *M. smegmatis* chromosome. Work is under way to explain why the different traps preferentially yielded different insertion elements.

REFERENCES

1. **Aldovini, A., and R. A. Young.** 1991. Humoral and cell-mediated immune responses to live recombinant BCG-HIV vaccines. Nature (London) **351**:479–482.
2. **Barletta, R. G., D. D. Kim, S. B. Snapper, B. R. Bloom, and W. R. Jacobs, Jr.** Identification of expression signals of the mycobacteriophages Bxb1, L1, and TM4 using *Escherichia-Mycobacterium* shuttle plasmids pYUB75 and pYUB76 designed to create translational fusions to the *lacZ* gene. J. Gen. Microbiol., in press.
3. **Belisle, J. T., L. Pascopella, J. Imamine, P. J. Brennan, and W. R. Jacobs, Jr.** 1991. Isolation and expression of a gene cluster responsible for biosynthesis of the glycopeptidolipid antigens of *Mycobacterium avium*. J. Bacteriol. **173**:6991–6997.
4. **Berg, D. E.** 1989. Transposon Tn*5*, p. 185–210. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
5. **Bernardini, M. L., J. Mounier, H. D'Hauteville, M. Coquis-Rondon, and P. J. Sansonetti.** 1989. Identification of *icsA*, a plasmid locus of *Shigella flexneri* that governs bacterial intra- and intercellular spread through interaction with F-actin. Proc. Natl. Acad. Sci. USA **86**:3867–3871.
6. **Bolivar, F., R. Rodriguez, P. J. Greene, M. Betlach, H. L. Heyneker, H. W. Boyer, J. Crosa, and S. Falkow.** 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2**:95–113.
7. **Cirillo, J. D., B. R. Bloom, and W. R. Jacobs, Jr.** Unpublished data.
8. **Dower, W. J., J. F. Miller, and C. W. Ragsdale.** 1988. High efficiency transformation of *E. coli* by high voltage electroporation. Nucleic Acids Res. **16**:6127–6145.
9. **Eisenach, K. D., M. D. Cave, J. H. Bates, and J. T. Crawford.** 1990. Polymerase chain reaction amplification of a repetitive DNA sequence specific for *Mycobacterium tuberculosis*. J. Infect. Dis. **161**:977–981.
10. **Fields, P. I., E. A. Groisman, and F. Heffron.** 1989. A Salmonella locus that controls resistance to microbicidal proteins from phagocytic cells. Science **243**:1059–1061.
11. **Fisher, S. H., C. J. Bruton, and K. F. Chater.** 1987. The glucose kinase gene of Streptomyces coelicolor and its use in selecting spontaneous deletions for desired regions of the genome. Mol. Gen. Genet. **206**:35–44.
12. **Green, E. P., M. L. V. Tizard, J. Thompson, D. J. Winterbourne, J. J. McFadden, and J. Hermon-Taylor.** 1989. Sequence and characteristics of IS*900*, an insertion element identified in a human Crohn's disease isolate of *Mycobacterium paratuberculosis*. Nucleic Acids Res. **17**:9063–9073.
13. **Grosskinsky, C. M., W. R. Jacobs, Jr., J. E. Clark-Curtiss, and B. R. Bloom.** 1989. Genetic relationships among *Mycobacterium leprae*, *Mycobacterium tuberculosis*, and candidate leprosy vaccine strains determined by DNA hybridization: identification of an *M. leprae*-specific repetitive sequence. Infect. Immun. **57**:1535–1541.
14. **Hopwood, D. A., M. J. Bibb, K. F. Chater, T. Kieser, C. J. Bruton, H. M. Kieser, D. J. Lydiate, C. P. Smith, J. M. Ward, and H. Schrempf.** 1985. Genetic manipulation of streptomyces: a laboratory manual. John Innes Foundation, Norwich, Great Britain.
15. **Husson, R. N., B. E. James, and R. A. Young.** 1990. Gene replacement and expression of foreign DNA in mycobacteria. J. Bacteriol. **172**:519–524.
16. **Isberg, R. R., and S. Falkow.** 1985. Genetic analysis of bacterial virulence determinants in *Bordetella pertussis* and the pathogenic yersinia. Curr. Top. Microbiol. Immunol. **118**:1–11.
17. **Jacobs, W. R., Jr., G. V. Kalpana, J. D. Cirillo, L. Pascopella, R. A. Udani, W. D. Jones, Jr., R. G. Barletta, and B. R. Bloom.** 1991. Genetic systems for the mycobacteria. Methods Enzymol. **204**:537–555.
18. **Jacobs, W. R., Jr., M. Tuckman, and B. R. Bloom.** 1987. Introduction of foreign DNA into mycobacteria using a shuttle phasmid. Nature (London) **327**:532–536.
19. **Kalpana, V. G., B. R. Bloom, and W. R. Jacobs, Jr.** 1991. Insertional mutagenesis and illegitimate recombination in mycobacteria. Proc. Natl. Acad. Sci. USA **88**:5433–5437.
20. **Kraft, R., J. Tardiff, K. S. Krauter, and L. A. Leinwand.** 1988. Using mini-prep plasmid DNA for sequencing double stranded templates with Sequenase. BioTechniques **6**:544–547.
21. **Lee, M. H., L. Pascopella, W. R. Jacobs, Jr., and G. F. Hatfull.** 1991. Site-specific integration of mycobacteriophage L5: integration-proficient vectors for *Mycobacterium smegmatis*, BCG, and *M. tuberculosis*. Proc. Natl. Acad. Sci. USA **88**:3111–3115.
22. **Martin, C., J. Timm, J. Rauzier, R. Gomez-Lus, J. Davies, and B. Gicquel.** 1990. Transposition of an antibiotic resistance element in mycobacteria. Nature (London) **345**:739–743.
23. **Martinez, E., and F. de la Cruz.** 1988. Transposon Tn*21* encodes a *RecA*-independent site-specific integration system. Mol. Gen. Genet. **211**:320–325.
24. **McAdam, R. A., P. W. M. Hermans, D. van Soolingen, Z. F. Zainuddin, D. Catty, J. D. A. van Embden, and J. W. Dale.** 1990. Characterization of a *Mycobacterium tuberculosis* insertion sequence belonging to the IS3 family. Mol. Microbiol. **4**:1607–1613.
25. **McFadden, J. J., P. D. Butcher, R. Chiodini, and J. Hermon-Taylor.** 1987. Crohn's disease-isolated mycobacteria are identical to *Mycobacterium paratuberculosis*, as determined by DNA probes that distinguish between mycobacterial species. J. Clin. Microbiol. **25**:796–801.
26. **Morse, M. L., E. M. Lederberg, and J. Lederberg.** 1956. Transductional heterogenotes in *Escherichia coli*. Genetics **41**: 758–779.
27. **Nakayama, K., S. M. Kelly, and R. Curtiss III.** 1988. Construction of an *asd*$^+$ expression-cloning vector: stable maintenance and high level expression of cloned genes in a salmonella vaccine strain. Bio/Technol. **6**:693–697.
28. **Patel, R. J., J. W. U. Fries, W. F. Piessens, and D. F. Wirth.** 1990. Sequence analysis and amplification by polymerase chain reaction of a cloned DNA fragment for identification of *Mycobacterium tuberculosis*. J. Clin. Microbiol. **28**:513–518.
29. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85**:2444–2448.
30. **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.
31. **Reed, R. R., G. I. Shibuya, and J. A. Steitz.** 1982. Nucleotide sequence of γδ resolvase gene and demonstration that its gene product acts as a repressor of transcription. Nature (London) **300**:381–383.
32. **Rowland, S., and K. G. H. Dyke.** 1990. Tn*552*, a novel transposable element from *Staphylococcus aureus*. Mol. Microbiol. **4**:961–975.
33. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
34. **Segall, A. M., and J. R. Roth.** 1989. Recombination between homologies in direct and inverse orientation in the chromosome of salmonella: intervals which are nonpermissive for inversion formation. Genetics **122**:737–747.
35. **Snapper, S., L. Lugosi, A. Jekkel, R. Melton, T. Kieser, B. R. Bloom, and W. R. Jacobs, Jr.** 1988. Lysogeny and transformation of mycobacteria: stable expression of foreign genes. Proc. Natl. Acad. Sci. USA **85**:6987–6991.

36. **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.
37. **Southern, E. M.** 1975. Detection of specific sequences amongst DNA fragments separated by gel electrophoresis. J. Mol. Biol. **908**:503–517.
38. **Stover, C. K., V. F. de la Cruz, T. R. Fuerst, J. E. Burlein, L. A. Benson, L. T. Bennet, G. P. Bansal, J. F. Young, M. H. Lee, G. F. Hatfull, S. B. Snapper, R. G. Barletta, W. R. Jacobs, Jr., and B. R. Bloom.** 1991. New use of BCG for recombinant vaccines. Nature **351**:456–460.
39. **Taylor, R. K., C. Manoil, and J. Mekalanos.** 1989. Broad-host-range vectors for delivery of Tn*phoA*: use in genetic analysis of secreted virulence determinants of *Vibrio cholerae*. J. Bacteriol. **171**:1870–1878.
40. **Thangaraj, H. S., F. I. Lamb, E. O. Davis, P. J. Jenner, L. H. Jeyakumar, and M. J. Colston.** 1990. Identification, sequencing, and expression of *Mycobacterium leprae* superoxide dismutase, a major antigen. Infect. Immun. **58**:1937–1942.
41. **Thierry, D., A. Brisson-Noël, V. Vincent-Lévy-Frébault, S. Nguyen, J. Guesdon, and B. Gicquel.** 1990. Characterization of a *Mycobacterium tuberculosis* insertion sequence, IS*6110*, and its application in diagnosis. J. Clin. Microbiol. **28**:2668–2673.
42. **Thierry, D., M. Cave, K. Eisenach, J. Crawford, J. Bates, B. Gicquel, and J. Guesdon.** 1990. IS*6110*, an IS-like element of *Mycobacterium tuberculosis*. Nucleic Acids Res. **18**:188.
43. **Turner, A. K., and J. Grinsted.** 1989. DNA sequence of the transposase gene of the class II transposon, Tn*3926*. Nucleic Acids Res. **17**:1757.
44. **Way, J. C., M. A. Davis, D. Morisato, D. E. Roberts, and N. Kleckner.** 1984. New Tn*10* derivatives for transposon mutagenesis and for construction of *lacZ* operon fusions by transposition. Gene **32**:369–379.
45. **Wayne, L. G., and W. M. Gross.** 1968. Base composition of deoxyribonucleic acid isolated from mycobacteria. J. Bacteriol. **96**:1915–1919.
46. **Yeh, P., A. M. Sicard, and A. J. Sinskey.** 1988. General organization of the genes specifically involved in the diaminopimelate-lysine biosynthetic pathway of *Corynebacterium glutamicum*. Mol. Gen. Genet. **212**:105–111.