Ethambutol Resistance in Mycobacterium tuberculosis: Critical Role of embB Mutations

SRINAND SREEVATSAN,1 KATHRYN E. STOCKBAUER,1 XI PAN,2 BARRY N. KREISWIRTH,2 SORAYA L. MOGHAZEH,3 WILLIAM R. JACOBS, JR.,3 AMALIO TELENTI,4 AND JAMES M. MUSSER1,5*

Section of Molecular Pathobiology, Department of Pathology, Baylor College of Medicine,1 and Clinical Microbiology Laboratory, The Methodist Hospital,2 Houston, Texas; Tuberculosis Center, Public Health Research Institute, New York, New York;2 Department of Microbiology and Immunology and Howard Hughes Medical Research Institute, Albert Einstein College of Medicine, Bronx, New York3; and Institute for Medical Microbiology, University of Bern, Bern, Switzerland4

Received 28 March 1997/Returned for modification 1 May 1997/Accepted 12 May 1997

Ethambutol [(S,S’)-2,2’-(ethylenedimino)di-1-butanol; EMB], is a first-line drug used to treat tuberculosis. To gain insight into the molecular basis of EMB resistance, we characterized the 10-kb embCAB locus in 16 EMB-resistant and 3 EMB-susceptible genetically distinct Mycobacterium tuberculosis strains from diverse localities by automated DNA sequencing and single-stranded conformation polymorphism analysis. All 19 organisms had virtually identical sequences for the entire 10-kb region. Eight EMB-resistant organisms had mutations located in codon 306 of embB that resulted in the replacement of the wild-type Met residue with Ile or Val. Automated sequence analysis of the S’ region (1,892 bp) of embB in an additional 69 EMB-resistant and 30 EMB-susceptible M. tuberculosis isolates from diverse geographic localities and representing 70 distinct IS6110 fingerprints confirmed the unique association of substitutions in amino acid residue 306 of EmbB with EMB resistance. Six other embB nucleotide substitutions resulting in four amino acid replacements were uniquely found in resistant strains. Sixty-nine percent of epidemiologically unassociated EMB-resistant organisms had an amino acid substitution not found in susceptible strains, and most (89%) replacements occurred at amino acid residue 306 of EmbB. For strains with the Met306Leu or Met306Val replacements EMB MICs were generally higher (40 μg/ml) than those for organisms with Met306Ile substitutions (20 μg/ml). The data are consistent with the idea that amino acid substitutions in EmbB alter the drug-protein interaction and thereby cause EMB resistance.

Ethambutol [(S,S’)-2,2’-(ethylenedimino)di-1-butanol; EMB] is one of the primary drugs used in combination with isoniazid, rifampin, streptomycin, and pyrazinamide to treat tuberculosis. Recent estimates suggest that in the United States EMB resistance occurs in 2.3 and 3.8% of patients with new and recurrent tuberculosis, respectively (3). A tabulation of 63 surveys of resistance to antituberculosis drugs performed between 1985 and 1994 demonstrated that the rate of acquired EMB resistance was as high as 13.7% in some countries (5).

Although several hypotheses have been advanced to explain the mechanism of action of EMB, most studies have implicated a detrimental alteration of the mycobacterial cell wall structure (1, 7, 9, 16, 20, 28, 29, 32, 38). Takayama and Kilburn (32) showed that EMB inhibited the transfer of arabinogalactan into the cell wall of Mycobacterium smegmatis, a process that led to the accumulation of trehalose mono- and dimycolates in the medium. EMB was also shown to inhibit the transfer of [14C]glucose into the α-arabinosyl residue of arabinogalactan (29). On this basis, it was proposed that deprivation of transfer of arabinogalactan to the cell wall led to the accumulation of mycolic acid in M. smegmatis, a result consistent with the finding that EMB treatment causes bacterial decumpling and morphological changes (16).

Despite the evidence implicating arabinosyltransferases as EMB targets, only recently has insight into the molecular genetics of arabinan biosynthesis been obtained. Belanger et al. (2) identified a two-gene locus (embA/IB) in Mycobacterium avium that encodes arabinosyltransferases mediating polymerization of arabinose into arabinogalactan. More recently, three genes encoding a putative EMB target (M. smegmatis) were cloned, sequenced, and characterized by molecular genetic strategies (36). These genes are organized as an operon, and because two of the three genes are homologous to embA in M. avium, they were designated embC, embA, and embB (36). The corresponding Mycobacterium tuberculosis genes (embCAB) were then identified from an ordered cosmids library, sequenced, and shown to be organized as a 10-kb operon. Preliminary studies demonstrated the occurrence of missense changes in embB codon 306 in a sample of EMB-resistant, but not EMB-susceptible, M. tuberculosis isolates (36). The present study was undertaken to more fully investigate the association of EMB resistance, MICs, and sequence variation in the 10-kb embCAB operon in a diverse set of recent clinical isolates of M. tuberculosis.

MATERIALS AND METHODS

Bacterial strains. The study is based on a sample of 118 M. tuberculosis isolates recovered in diverse localities in the United States, Mexico, Europe, Yemen, the Philippines, Japan, and India. The sample included 85 EMB-resistant and 33 EMB-susceptible organisms. The bacteria were recovered from diseased patients with a variety of distinct clinical manifestations, including pulmonary and extrapulmonary infections. IS6110 genotyping and genetic group assignment. Strain growth and DNA isolation were done in laboratories equipped with biosafety level 3 facilities. Isolation of chromosomal DNA and IS6110 typing were performed by previously described methods (37). Recent data have demonstrated that all M. tuberculosis isolates can be assigned to one of three genetic groups on the basis of the

* Corresponding author. Mailing address: Section of Molecular Pathobiology, Department of Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030. Phone: (713) 798-4198. Fax: (713) 798-4595. E-mail: jmuesser@path bcm.tmc.edu.
polymorphisms present at codon 463 of the gene (katG) encoding catalase-peroxidase and codon 95 of the gene (gyrA) encoding the A subunit of DNA gyrase (31). The group designations used are as follows: group 1, katG63 CTG (Leu) plus gyrA495 ACC (Thr); group 2, katG63 CGG (Arg) plus gyrA495 ACC (Thr); group 3, katG63 CGG (Arg) plus gyrA495 ACC (Ser).

**EMB susceptibility testing.** The strains studied were initially tested for EMB susceptibility in routine diagnostic laboratories either by the BACTEC radiometric method (7.5 μg/ml) or by agar diffusion with Middlebrook 7H10 medium (5 μg/ml). To determine EMB MICs, susceptibility testing was performed by agar diffusion as described previously (26). EMB (Sigma Chemical Co., St. Louis, Mo.) was incorporated into 7H10 agar (Difco, Detroit, Mich.) at the concentration of 5% CO2. Each plate was checked weekly, and the susceptibility results were recorded between weeks 3 and 4.

**Characterization of the embC locus.** A combination of automated DNA sequencing and single-stranded conformation polymorphism (SSCP) analysis was used to analyze nucleotide variation in the 10-kb embC region. First, a 7.5-kb region of the embCAB region was sequenced in stages by using 30 primers. This region contains the distal one-third of embC and all of the embB and embB genes. The complete primer list is available by request from J.M.M. A GeneAmp System 9600 thermocycler (Perkin-Elmer Corp, Foster City, Calif.) was used for all target amplifications, with the following parameters: annealing temperature of 65°C for 9 s, extension at 72°C for 82 s, and a denaturation step at 94°C for 54 s for a total of 30 cycles. Each reaction was preceded by an initial denaturation step at 94°C for 60 s and was terminated with a final extension step at 72°C for 5 min. DNA sequencing reactions were performed with the Taq DyeDecay terminator cycle sequencing kit by using AmpliTaq DNA polymerase FS (Applied Biosystems, Inc., Foster City, Calif.). The forward, reverse, and internal primers were used for sequencing. Sequence data generated with an ABI 377 automated instrument were assembled and edited electronically with EDITSEQ, ALIGN, and MEGALIGN programs (DNASTAR, Madison, Wis.) and were compared with a published sequence of embCAB (36) and a sequence generated from M. tuberculosis H37Rv.

Nineteen M. tuberculosis (16 EMB-resistant and 3 EMB-susceptible) isolates were analyzed by sequencing the 7.5-kb region of embCAB except for a 75-bp segment located at the beginning of embB. This region has an inferred stem-loop structure (36), and sequence data could not be obtained for all 19 isolates. The 5’ region (1,892 bp) of embB was analyzed in an additional 69 EMB-resistant and 30 EMB-susceptible M. tuberculosis isolates from diverse geographic localities and representing 70 IS6110 fingerprints.

Sequence variation in the proximal two-thirds of embB in 20 EMB-resistant organisms was studied by targeted SSCP analysis with fluorescent-labelled primers (35). Two regions that were thought to potentially have resistance-associated mutations were investigated. One region containing the embC promoter was analyzed with primers TE26 (5’-GGTCCAGAAGCCGCAGGAC-3’) and TE19 (5’-CTAGTGTGATTAGAGCCGG-3’), and a second region conserved to the embB EMB resistance-determining region was studied with primers TE19 (5’-GCCAGAGGTGTTAGATCCA-3’) and TE20 (5’-TGAAGATGCGCCGATGATC-3’).

**RESULTS**

**Alterations in the embCAB locus among 19 strains sequenced for 7.5 kb.** Automated DNA sequencing analysis of 7.5 kb of the 10-kb embCAB locus in 16 EMB-resistant and 3 EMB-susceptible isolates identified a total of four polymorphic nucleotide sites, each located in different regions of the gene locus. For example, a polymorphism (G→C) was identified in embC in two EMB-resistant and one EMB-susceptible isolate. This change would produce a Val→Leu substitution in amino acid 981 of EmbC. These same three isolates also had a G→A synonymous substitution at nucleotide 2898 of embA. One resistant isolate also had a missense mutation in codon 5 of the gyrA gene (GTG→ATG) resulting in an Ile→Met alteration. Impor-

tantantly, eight epidemiologically independent EMB-resistant strains had one of three distinct missense mutations in codon 306 of embB (ATG, Met) resulting in substitution with Ile (ATA and ATC) or Val (GTG). Hence, the sequence data for the 7.5-kb region of the embCAB operon clearly suggested that missense mutations associated with resistance were overrepresented in embB codon 306.

**Variation in the 5’ segment of embB among 69 EMB-resistant and 30 EMB-susceptible isolates.** Because the sequencing results suggested that mutations in codon 306 of embB were important in EMB resistance, we next sequenced a 1,892-bp embB region in 69 EMB-resistant and 30 EMB-susceptible organisms. This sample included 51 EMB-resistant and 25 EMB-susceptible strains judged to be epidemiologically unassociated because they had distinct IS6110 patterns. All 30 EMB-susceptible organisms had the same wild-type sequence. In striking contrast, many of the 69 EMB-resistant strains had one of eight different missense mutations in this 1,892-bp region. Three mutations occurred in individual codons, including 285 (TTC→TTA, Phe→Leu), 330 (TTC→GTC, Phe→Val), and 630 (ACC→ATC, Thr→Ile). Five distinct missense mutations were identified in codon 306 (ATG, Met): GTC (Val), CTG (Leu), ATA (Ile), ATC (Ile), and ATT (Ile) (Fig. 1). When only epidemiologically unassociated organisms are considered, 69% of EMB-resistant bacteria had an amino acid change in the region of EmbB studied, and most replacements (89%) occurred at position 306.

The analysis included two pairs of EMB-resistant and -susceptible isolates with two different IS6110 fingerprints. The resistant isolates had either an EmbB Met306Ile or a Met306Val alteration, whereas the paired susceptible isolates with the same IS6110 type had identical wild-type embB sequences.

**SSCP analysis of the proximal two-thirds of embC.** The DNA sequence data and recent molecular genetic studies of EMB-resistant mutants of M. smegmatis (36) suggested that mutations frequently associated with EMB resistance were located in embB. To minimize cost, we chose to characterize the proximal two-thirds of embC by SSCP analysis rather than automated DNA sequencing. Analysis of two targeted regions in 20 EMB-resistant organisms identified only 1 isolate with an aberrant profile by SSCP analysis. Subsequent DNA sequencing confirmed that this strain had an ACC→ATC substitution resulting in a Thr270Ile replacement. Because this strain also had an EmbB Met306Val substitution, the relevance of the EmbC alteration to EMB resistance is unclear.

**Correlation of embB mutations with EMB MICs.** We next studied the relationship of embB mutations with the level of resistance to EMB. Because of the relatively large number of strains whose embB sequences were characterized, the EMB MIC was not analyzed for all organisms. Rather, we studied a representative subset of strains that together contained a broad array of embB mutations and IS6110 subtypes. MICs were determined for 36 resistant and 4 susceptible organisms by plating on 7H10 agar containing EMB at a range of concentrations between 0 and 40 μg/ml. The MICs for the genetically characterized M. tuberculosis isolates are presented in Table 1. In contrast to EMB-susceptible organisms with a wild-type embB codon 306, for all isolates with EMB position 306 substitutions, EMB MICs were 20 μg/ml or greater. Interestingly, regardless of the genetic background, as indexed by IS6110 typing and katG463gyrA95 group, for bacteria with a Met306Ile replacement the MIC was virtually always 20 μg/ml, whereas for organisms with a Met306Leu or Met306Val change the MIC was 40 μg/ml. In addition, for one organism with a Phe330Val amino acid substitution the MIC was 40 μg/ml, and for two epidemiologically related strains with a Thr360Ile change the EMB MIC exceeded 40 μg/ml, the highest concentration of drug tested.

Nucleotide sequence accession numbers. The mutations identified in this study have been deposited in EMBL under accession no. U68480.
DISCUSSION

The recent demonstration that mutations in the \textit{M. tuberculosis} \textit{embCAB} operon are associated with EMB resistance prompted a more thorough examination of sequence variation in this operon in genetically distinct susceptible and resistant organisms. In particular, the aim of this study was to determine if polymorphisms were uniquely present in EMB-resistant versus EMB-susceptible \textit{M. tuberculosis} isolates. Our results documented two findings about sequence variation in this 10-kb region. First, there is restricted variation in \textit{embCAB} in natural populations of \textit{M. tuberculosis} recovered from diverse geographic sources. This result confirmed data generated by study of several other antimycobacterial agent target genes including \textit{katG} (12, 14, 22–24), \textit{rpoB} (14, 15, 23, 34), \textit{strA} (8, 22, 25), \textit{rrs} (12, 23), and \textit{gyrA} (15, 23, 33, 38). The restricted natural variation in \textit{embCAB} is also consistent with the hypothesis that \textit{M. tuberculosis} is evolutionarily new, perhaps having arisen as recently as 15,000 to 20,000 years ago (13, 30, 31). Second, the results documented the unique and common association of \textit{EmbB} amino acid residue 306 substitutions in EMB resistance. The only reasonable hypothesis to explain the occurrence of five distinct mutant codons resulting in three different amino acid replacements at \textit{EmbB} position 306 of EMB-resistant organisms is that they have arisen by positive Darwinian selection in the course of drug therapy. This interpretation is strongly supported by the observation that pairs of organisms with the same IS6110 types, but differing by amino acid replacements at \textit{EmbB} position 306, were resistant to EMB. The results of gene transfer experiments also support this thesis. Transformation of susceptible \textit{M. smegmatis} with a multicopy plasmid carrying a mutant \textit{M. smegmatis} \textit{embB} gene (Ile303Phe) resulted in more than a 500-fold increase in the EMB MIC (36). Taken together, we believe that it is likely that amino acid substitutions at \textit{EmbB} position 306 mediate EMB resistance, rather than act simply as surrogate markers for drug-resistant organisms.

The region of \textit{EmbB} containing residue Met 306 is highly conserved in \textit{M. tuberculosis}, \textit{M. avium}, \textit{Mycobacterium leprae}, and \textit{M. smegmatis} (2, 36). Moreover, on the basis of alignments of inferred amino acid sequences, a Met residue would be present at position 306 in these four species. Biochemical studies and sequence alignments have suggested that \textit{EmbB} is a glycosyltransferase (2). Telenti et al. (36) hypothesized that \textit{EmbB} is an integral membrane protein with 12 transmembrane domains and a C-terminal globular region of about 400 amino acids with a predicted location in the periplasm. Our data are fully consistent with the idea that the structural alterations that we identified detrimentally affect the interaction

FIG. 1. DNA sequencing chromatograms showing the wild type and five mutants of \textit{embB} codon 306. The region of each chromatogram depicting \textit{embB} codon 306 is outlined. The wild-type codon in EMB-susceptible isolates is ATG (Met), and the five mutant codons found in EMB-resistant isolates are GTG (Val), CTG (Leu), ATA (Ile), ATT (Ile), and ATC (Ile). The chromatograms were generated on the basis of sequence data obtained with an Applied Biosystems model 377 DNA sequencing instrument.
between EMB, a putative arabinose analog (18), and EmbB, likely to be a glycosyltransferase. The identification of several mutant EmbB variants provides an important resource for biochemical and structural studies designed to more fully investigate the molecular mechanism of action of EMB. For the ~30% of EMB-resistant isolates lacking EmbB changes, alternative mechanisms of resistance must be operating. Although we currently do not understand the molecular basis for resistance in these isolates, we note that overexpression of EmbB proteins was documented to mediate EMB resistance in M. smegmatis (36).

Strains with the IS6110 subtype, designated H strains (two copies of IS6110) (17), warrant special comment. Our analysis showed that these organisms have wild-type embA4 and embB genes. However, they also all have a polymorphism in embC that results in a Val981Leu substitution in EmbC. Because we observed that organisms with this embC change were sometimes judged to be EMB susceptible but at other times were judged to be EMB resistant by several independent diagnostic laboratories, we investigated the MICs for five H strains containing wild-type embCAB sequences except for the change producing the Val981Leu replacement. Interestingly, the MIC for all five H organisms was 15 μg/ml, a value consistent with EMB resistance. We speculate that H strains with this amino acid change may be reported as susceptible or resistant, depending on the method used for susceptibility testing, slight variations in growth conditions, or other subtle changes in testing protocols. It is well known that EMB activity differs in various media due to variable absorption and inactivation (10, 11). However, until additional data are available, it is prudent to be cautious about concluding that the Val981Leu change and other amino acid replacements located outside of EmbB position 306 are causally involved in EMB resistance.

The data generated from the EMB susceptibility testing indicate a strong nonrandom association between certain amino acid substitutions at EmbB position 306 and the level of resistance to EMB. Analogous data have been presented for amino acid substitutions in the fluoroquinolone resistance-determining region of the A subunit of RNA gyrase (33, 39), changes in the beta subunit of RNA polymerase that confer resistance to rifampin and related drugs (4, 21, 27), and mutations conferring streptomycin resistance (6, 19). The association of particular amino acid substitutions with the EMB MIC implies that position 306 contains critical structure-function information. Additional studies are required to test this proposition directly.

ACKNOWLEDGMENTS

This research was supported by Public Health Service grant AI-57004 (to J.M.M.) and Swiss National Science Foundation grant 31-55030.9.

The ongoing support of A. M. Ginsberg is gratefully appreciated.

REFERENCES


