JOURNAL OF CLINICAL MICROBIOLOGY, Apr. 1994, p. 1095–1098
0095-1137/94/$04.00+0
Copyright © 1994, American Society for Microbiology
Vol. 32, No. 4

Characterization by Automated DNA Sequencing of Mutations in the Gene (*rpoB*) Encoding the RNA Polymerase β Subunit in Rifampin-Resistant *Mycobacterium tuberculosis* Strains from New York City and Texas

VIVEK KAPUR,[1] LING-LING LI,[1] SERBAN IORDANESCU,[2] MARCIE R. HAMRICK,[1] AUDREY WANGER,[3] BARRY N. KREISWIRTH,[2] AND JAMES M. MUSSER[1,4]*

Section of Molecular Pathobiology, Department of Pathology, Baylor College of Medicine,[1] Clinical Microbiology Laboratory, The Methodist Hospital,[4] and Department of Pathology and Laboratory Medicine, University of Texas Medical School,[3] Houston, Texas 77030, and Tuberculosis Center, Public Health Research Institute, New York, New York 10016[2]

Received 18 October 1993/Returned for modification 7 December 1993/Accepted 3 January 1994

Automated DNA sequencing was used to characterize mutations associated with rifampin resistance in a 69-bp region of the gene, *rpoB*, encoding the β subunit of RNA polymerase in *Mycobacterium tuberculosis*. The data confirmed that greater than 90% of rifampin-resistant strains have sequence alterations in this region and showed that most are missense mutations. The analysis also identified several mutant *rpoB* alleles not previously associated with resistant organisms and one short region of *rpoB* that had an unusually high frequency of insertions and deletions. Although many strains with an identical IS*6110* restriction fragment length polymorphism pattern have the same variant *rpoB* allele, some do not, a result that suggests the occurrence of evolutionary divergence at the clone level.

Tuberculosis, a disease whose frequency declined steadily in the United States since the early 1950s, has resurged, although it was optimistically thought that it might be virtually eliminated by the end of the first decade of the 21st century (2). The yearly decline in tuberculosis incidence ended in 1984, and after several years of a plateau phase, the incidence increased from 1988 through 1992 (13). An estimated 51,700 excess tuberculosis cases have occurred in the United States through 1992 (3). Commensurate with the rise in tuberculosis frequency, an increase in the isolation of *Mycobacterium tuberculosis* strains resistant to one or more of the antibiotics commonly used as treatment has occurred (6, 8, 12). Recently, the molecular basis of rifampin resistance in *M. tuberculosis* was identified (23). Telenti et al. (23) described 15 distinct mutations present in the gene, *rpoB*, encoding the β subunit of RNA polymerase in 64 rifampin-resistant organisms recovered from several continents but absent in 56 susceptible organisms. Each mutation results in the substitution of 1 of 8 amino acids clustered in a 23-amino-acid (69-bp) region. Two mutations (CAC→TAC and TCG→TTG, resulting in His→Tyr and Ser→Leu substitutions at amino acid residues 526 and 531, respectively) accounted for 61% of resistant strains. Although organisms from nine countries were studied, only 12 U.S. isolates were examined, and their states of origin were not provided.

We report here the use of an automated DNA sequencing strategy to characterize *rpoB* mutations in a sample of 121 rifampin-resistant strains from three states. The present study was undertaken for two reasons. First, because the rifampin-resistant strains studied by Telenti et al. (23) were mainly drawn from non-U.S. sources, we believed it important to examine a large sample of strains from defined geographic regions of this country. The second purpose of our study was to examine the range of *rpoB* mutations present in *M. tuberculosis* strains classified on the basis of IS*6110* restriction fragment length polymorphism (RFLP) pattern subtypes. This question is of interest because it bears on the extent to which IS*6110* subtypes adequately mark individual resistant clones.

A sample of 128 *M. tuberculosis* strains was studied, and it included 98 strains cultured from patients in New York City, 28 strains recovered from infected individuals in Texas, and 2 strains grown from diseased New Mexicans. For the Texas sample, 15 strains were isolated from infected Houstonians, and 13 were taken from tuberculosis patients living throughout the state. The sources of isolation for the non-Houston sample were San Antonio (n = 3), McAllen (n = 2), and Brownsville, Dallas, Eagle Pass, El Paso, Galveston, Harlingen, Laredo, and Victoria (n = 1 each). The communities of Brownsville, Harlingen, McAllen, Eagle Pass, Laredo, and El Paso are located essentially on the border of the United States and Mexico formed by the Rio Grande River. Strains in the non-Houston sample were selected from rifampin-resistant organisms cultured from patients throughout the state and were chosen to enhance the geographic diversity of the Texas collection. Twenty-six of these strains were studied previously (16). Seven of the 128 strains were susceptible to rifampin.

Rifampin susceptibility testing was performed by either the radiometric (21) or the agar dilution (11) method. The IS*6110* RFLP patterns of the strains were determined according to internationally standardized guidelines (24).

A 350-bp fragment of *rpoB* was amplified by the PCR with the following synthetic oligonucleotide primers: forward, 5′-GGGAGCGGATGACCACCCA-3′; and reverse, 5′-GCGG TACGGCGTTTCGATGAAC-3′. PCR amplification of 0.5 μl of chromosomal DNA was performed with a 30-μl mixture containing 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM $MgCl_2$, 200 μM each dATP, dCTP, dGTP, and dTTP, 500 nM each primer, and 0.83 U of AmpliTaq DNA polymerase

* Corresponding author. Mailing address: Section of Molecular Pathobiology, Department of Pathology, Baylor College of Medicine, Houston, TX 77030. Phone: (713) 798-3823. Fax: (713) 798-5838.

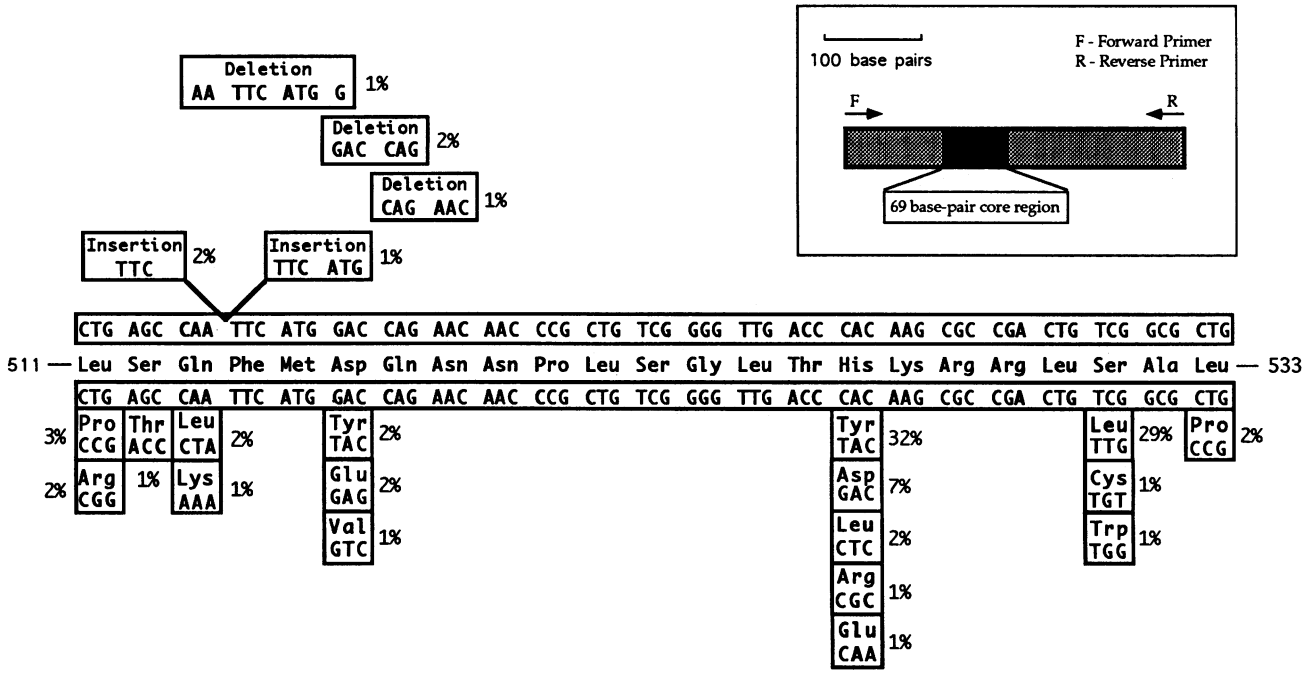

FIG. 1. Frequency of mutations in codons 511 through 533 of the *M. tuberculosis rpoB* gene in 121 rifampin-resistant isolates. The positions of insertions and deletions are illustrated above the wild-type sequence, and missense mutations are recorded below. To permit ready comparison between our data and those generated by Telenti et al. (23), we used the *rpoB* codon numbering system that they used. The codon numbers are designated on the basis of alignment of the translated *E. coli rpoB* sequence with a portion of the translated *M. tuberculosis* sequence and are not the actual *M. tuberculosis rpoB* codons.

(Perkin-Elmer Cetus, Norwalk, Conn.). The following thermocycler parameters were used: denaturation at 94°C for 4 min and 30 cycles of denaturation at 94°C for 1 min, primer annealing at 56°C for 1 min, extension at 72°C for 1 min, and a final extension at 72°C for 15 min. The unincorporated nucleotides and primers were separated from amplified DNA by filtration through Microcon 100 microconcentrators (Amicon, Inc., Beverly, Mass.). Sequencing reactions with the *Taq* DyeDeoxy terminator cycle sequencing kit (Applied Biosystems Inc.) were performed with 7 μl of PCR-amplified DNA as the template and 3.2 pmol of either the forward or the reverse primer. The unincorporated dye terminators and primers were separated from the extension products by spin column purification (Centri-Sep; Princeton Separations Inc., Adelphia, N.J.). The products were dried in a vacuum centrifuge, resuspended in 4 μl of loading buffer (5:1 deionized formamide–50 mM EDTA [pH 8.0]), heat denatured for 2 min at 90°C, and immediately loaded on an acrylamide gel in an Applied Biosystems model 373A automated DNA sequencer. Both strands were sequenced, and previously undescribed alleles were sequenced again. The data were assembled and edited with EDITSEQ, ALIGN, and SEQMAN programs (DNASTAR).

To permit ready comparison between our data and those generated by Telenti et al. (23), we used the same *rpoB* codon numbering system. The codon numbers are designated on the basis of alignment of the translated *Escherichia coli rpoB* sequence with a portion of the translated *M. tuberculosis rpoB* sequence and are not the actual *M. tuberculosis rpoB* codons.

Our analysis identified 23 distinct *rpoB* alleles associated with rifampin resistance (Fig. 1). Twelve of the mutant *rpoB* alleles were created by point mutations, two were due to insertions of 3 or 6 bp, three were made by deletion of 6 bp (two alleles) or 9 bp (one allele), and six were characterized by combinations of two point mutations (Table 1).

A total of 52 strains had one point mutation only in codon 526 (CAC; His), and 34 strains had only a TCG→TTG (Ser→Leu) mutation in codon 531. Point mutations were also identified in target region codons other than 526 and 531 (Fig. 1).

Several examples of alleles with complex mutations were identified by the analysis, and these included insertions, dele-

TABLE 1. *M. tuberculosis* isolates with more than one point mutation in the studied 69-bp region of the *rpoB* gene

Strain(s)	Codon	Nucleotide change	Amino acid substitution
TN 276 and TN 605	511	CTG→CCG	Leu→Pro
	516	GAC→GAG	Asp→Glu
TDH 794 and TDH 796	511	CTG→CCG	Leu→Pro
	516	GAC→TAC	Asp→Tyr
TN 648	526	CAC→CAA	His→Gln
	531	TCG→TGT	Ser→Cys
TDH 805 and 3750	528	CGA→CGT	Arg→Arg
	531	TCG→TTG	Ser→Leu
TDH 798	514	TTC→TTT	Phe→Phe
	531	TCG→TTG	Ser→Leu
TDH 801	511	CTG→CCG	Leu→Pro
	512	AGC→ACC	Ser→Thr

tions, and combinations of point mutations in two distinct codons. The insertions represent apparent duplication events. The insertions and deletions are clustered in one area of the 69-bp segment.

Several strains had combinations of point mutations in two noncontiguous codons; these are summarized in Table 1.

There were examples of the identification of distinct *rpoB* alleles among isolates with the same IS*6110* RFLP pattern. For example, strains with the H subtype, characterized by two IS*6110* copies, had a TCG→TTG (Ser→Leu) mutation in codon 531 (seven strains), a CTG→CCG (Leu→Pro) change in codon 533 (one strain), or a CAC→TAC (His→Tyr) mutation in codon 526 (one strain). Of the six strains with a P IS*6110* RFLP pattern (12 hybridizing bands), four had a CAC→TAC (His→Tyr) change in codon 526, but two were characterized by an *rpoB* allele with missense mutations in codons 511 (CTG→CCG; Leu→Pro) and 516 (GAC→GAG; Asp→Glu).

The two strains from New York City that had the 3-bp TTC insert and were resistant only to rifampin had the same IS*6110* RFLP pattern. Similarly, six strains with the W IS*6110* RFLP pattern (16 hybridizing bands) had a CAC→TAC (His→Tyr) mutation in codon 526. One strain that is apparently a single-band variant of the W pattern had a TCG→TTG change in codon 531 specifying a Ser→Leu amino acid substitution.

In the United States, the dramatic resurgence of tuberculosis, the spread of multidrug-resistant strains, the relatively limited number of effective antituberculosis drugs, and the person-to-person respiratory transmission route of *M. tuberculosis* together constitute a clear and present danger to society. It is therefore of signal importance to have a full understanding of the nature and extent of mutations conferring antimicrobial resistance and to develop rapid and unambiguous procedures to detect these organisms. The results of our study demonstrated that virtually every *rpoB* mutation previously associated with rifampin resistance (23) was present in strains from New York City and that the commonly occurring mutations were represented in the Texas and New Mexico samples. These data are consistent with the identification of the common *rpoB* mutations among rifampin-resistant organisms from geographically widespread countries reported by Telenti et al. (23). However, our analysis also defined several mutant *rpoB* alleles not described in the earlier study.

Our analysis found evidence of geographic variation in the frequency of occurrence of some rare mutant *rpoB* alleles, a result that may be due to local strain spread. Consistent with this hypothesis is the identification of two strains recovered from patients living in Victoria, Tex., that had the same unusual combination of two *rpoB* mutations (CTG→CGG and GAC→TAC in codons 511 and 516, respectively). The occurrence of geographic variation in the frequency of certain *rpoB* mutations may have important implications for development of molecular strategies designed to rapidly identify mutant *rpoB* alleles, for it suggests the need for relatively extensive sampling of resistant strains from several areas of the country, especially regions, such as Texas and other border states, with an abundant immigrant population. These regions may undergo increased strain flux as a consequence of passage of the North American Free Trade Agreement. It will clearly be of interest to perform an extensive study of strains recovered from patients of known ethnicity and immigration history.

One unexpected and important finding was the difference in the relative abundance of the CAC→TAC (His→Tyr) mutation in codon 526 and the TCG→TTG (Ser→Leu) change in codon 531 in our sample compared with a group of resistant strains from nine countries (23). Strains with the CAC→TAC codon 526 change accounted for about 30% of our sample, whereas organisms with this mutation represented only 12% of the sample examined by Telenti et al. (23). Similarly, only about 25% of our strains had the TCG→TTG mutation in codon 531, whereas organisms with this change accounted for 47% of the earlier sample. The data can be explained by either geographic variation in the frequency of occurrence of particular *rpoB* mutations or sample bias. Inasmuch as successful development of rapid tests to detect rifampin-resistant strains may depend in part on an understanding of the frequency distribution of *rpoB* mutations, it will be very important to characterize large samples of strains from diverse geographic sources.

The *rpoB* sequence data also identified one previously unrecognized episode of person-to-person *M. tuberculosis* transmission. The Texas strains with the unusual combination of mutations in codons 511 and 516 were initially designated as originating from patients in Victoria and San Antonio. After we identified the same *rpoB* allele in these two strains, the information was relayed to the Texas Department of Health, with the suggestion that further investigation of the two patients be undertaken to explore the possibility of epidemiologic association. The patients were found to be related as uncle and nephew.

Our analysis demonstrated that strains classified on the basis of IS*6110* RFLP patterns can have distinct *rpoB* alleles. Although there are several hypotheses that can explain this observation, we favor the notion that clonally related susceptible strains (as marked by IS*6110* RFLP patterns) have independently spawned subclones bearing distinct mutant *rpoB* alleles. A second hypothesis that can explain the data invokes convergence to IS*6110* RFLP identity by resistant strains bearing the same *rpoB* allele.

No mutation was detected in 3 of the 121 rifampin-resistant *M. tuberculosis* strains in our sample. These organisms were retested for resistance to rifampin and were confirmed to be resistant to this antibiotic. For these strains, it is possible that mutations elsewhere in *rpoB* conferred resistance. In *E. coli*, substitutions of amino acids 146 and 687, which would be located exterior to the residues encoded by the 350-bp region characterized in the *M. tuberculosis* gene, have been reported to result in rifampin resistance (14, 17, 20). Because of the presumed large size of *rpoB* in *M. tuberculosis* (9, 10) (the *E. coli* gene is ~4 kb in size), we did not attempt to characterize other gene regions. A second formal possibility to account for the lack of detected *rpoB* mutations in these three rifampin-resistant strains is that changes have occurred in one or more genes whose products participate in antibiotic permeability or metabolism.

One question not yet addressed is the relationship between mutant *rpoB* genotypes and organism phenotypes. It is well known from work with *E. coli* bearing *rpoB* mutations in the region analogous to that studied in *M. tuberculosis* that *rpoB* mutations can have pleiotropic effects and that certain *rpoB* alleles confer significantly different phenotypes (15). For example, 12 of 17 rifampin-resistant mutants of *E. coli* had an altered ability to form colonies at a temperature of 20 or 44°C. Six mutants were nonpermissive at a low temperature (20°C), four were nonpermissive at a high temperature (44°C), and two were nonpermissive at both low and high temperatures (15). In addition, many of the *E. coli rpoB* missense mutants had significantly longer doubling times at 37°C. Similarly, Ovchinnikov et al. (19) demonstrated that one *E. coli* strain with an allele (*rpoB256*) characterized by a deletion of 9 bp leading to the elimination of 3 amino acid residues at positions 531 to 533 of the wild-type sequence also had a significant decrease in

growth rate even in the absence of rifampin. It is therefore plausible that one or more of the *M. tuberculosis rpoB* mutations that we identified bestow distinct phenotypes that could be clinically relevant.

It will be crucial to determine at the molecular level why two mutations account for an unusually large percentage of all rifampin-resistant strains. Adequate explanation must account for both the position of the affected amino acids and the particular substituted residue. The likelihood of pathogen survival may be enhanced by selection of mutations that balance a high level of antibiotic resistance with minimal deleterious metabolic effects. Without a detailed understanding of the structure-function relationships of the RNA polymerase β subunit, as determined by X-ray crystallographic and other types of analyses, it is impossible to offer an adequate explanation for the large number of strains with mutations at positions 526 and 531.

In summary, our study demonstrates the feasibility and ease of using automated DNA sequencing to rapidly and unambiguously characterize *rpoB* mutations associated with rifampin resistance. One of the main advantages of this technique is that it permits the identification of all mutations in the target sequence and, as demonstrated here, successfully identifies previously uncharacterized mutant alleles. Because the automated sequencing strategy is generally applicable to any gene, it is readily adaptable for the identification of point mutations associated with resistance to other antituberculosis medications, such as isoniazid (1, 4), streptomycin (5, 7, 18), and ciprofloxacin (22) and to analysis of primary clinical specimens.

We are indebted to R. B. Ferguson, Tuberculosis/Mycology Section, Scientific Laboratory Division, New Mexico Department of Health, S. Gibson, Bacteriology/Mycology Branch, Texas Department of Health, and L. B. Elliott, Microbiological Services Division, Bureau of Laboratories, Texas Department of Health, for providing strains. We gratefully acknowledge E. Mayrand and R. Stanton of Applied Biosystems Inc. for providing DNA sequencing kits.

REFERENCES

1. **Banerjee, A., E. Dubnau, A. Quemard, V. Balasubramanian, K. Sun Um, T. Wilson, D. Collins, G. de Lisle, and W. R. Jacobs, Jr.** 1994. *inhA*, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. Science **263:**227–230.
2. **Centers for Disease Control.** 1989. A strategic plan for the elimination of tuberculosis in the United States. Morbid. Mortal. Weekly Rep. **38**(Suppl. S-3)**:**1–25.
3. **Centers for Disease Control.** 1993. Tuberculosis morbidity—United States, 1992. Morbid. Mortal. Weekly Rep. **42:**696–697, 703–704.
4. **Cole, S. T.** Personal communication.
5. **Douglass, J., and L. M. Steyn.** 1993. A ribosomal gene mutation in streptomycin-resistant *Mycobacterium tuberculosis* isolates. J. Infect. Dis. **167:**1505–1506.
6. **Ellner, J. J., A. R. Hinman, S. W. Dooley, M. A. Fischl, K. A. Sepkowitz, M. J. Goldberger, T. M. Shinnick, M. D. Iseman, and W. R. Jacobs, Jr.** 1993. Tuberculosis symposium: emerging problems and promise. J. Infect. Dis. **168:**537–551.
7. **Finken, M., P. Kirschner, A. Meier, A. Wrede, and E. C. Böttger.** 1993. Molecular basis of streptomycin resistance in *Mycobacterium tuberculosis*: alterations of the ribosomal protein S12 gene and point mutations within a functional 16S ribosomal RNA pseudoknot. Mol. Microbiol. **9:**1239–1246.
8. **Frieden, T. R., T. Sterling, A. Pablos-Mendez, J. O. Kilburn, G. M. Cauthen, and S. W. Dooley.** 1993. The emergence of drug-resistant tuberculosis in New York City. N. Engl. J. Med. **328:**521–526.
9. **Honore, N., S. Bergh, S. Chanteau, F. Doucet-Populaire, K. Eiglmeier, T. Garnier, C. Georges, P. Launois, T. Limpaiboon, S. Newton, K. Niang, P. del Portillo, G. R. Ramesh, P. Reddi, P. R. Ridel, N. Sittisombut, S. Wu-Hunter, and S. T. Cole.** 1993. Nucleotide sequence of the first cosmid from the *Mycobacterium leprae* genome project: structure and function of the rif-str regions. Mol. Microbiol. **7:**207–214.
10. **Honore, N., and S. T. Cole.** 1993. Molecular basis of rifampin resistance in *Mycobacterium leprae*. Antimicrob. Agents Chemother. **37:**414–418.
11. **Inderlied, C.** 1991. Antimycobacterial agents: in vitro susceptibility testing, spectrums of activity, mechanisms of action and resistance, and assays for activity in biological fluids, p. 136–197. *In* V. Lorian (ed.), Antibiotics in laboratory medicine. The Williams & Wilkins Co., Baltimore.
12. **Iseman, M. D.** 1993. Treatment of multidrug-resistant tuberculosis. N. Engl. J. Med. **329:**784–791.
13. **Jereb, J. A., G. D. Kelly, S. W. Dooley, G. M. Cauthen, and D. E. Snider.** 1991. Tuberculosis morbidity in the United States: final data, 1990. Morbid. Mortal. Weekly Rep. **40**(Suppl. S-3)**:**23–27.
14. **Jin, D. J., and C. A. Gross.** 1988. Mapping and sequencing of mutations in the *Escherichia coli rpoB* gene that lead to rifampicin resistance. J. Mol. Biol. **202:**45–58.
15. **Jin, D. J., and C. A. Gross.** 1989. Characterization of the pleiotropic phenotypes of rifampin-resistant *rpoB* mutants of *Escherichia coli*. J. Bacteriol. **171:**5229–5231.
16. **Kapur, V., M. R. Hamrick, L.-L. Li, S. Iordanescu, B. N. Kreiswirth, and J. M. Musser.** 1993. Exploitation of automated DNA sequencing for rapid characterization of *rpoB* mutations conferring rifampicin resistance in *Mycobacterium tuberculosis*, abstr. T12. Abstr. Am. Soc. Microbiol. Conf. Mol. Diagn. Ther., Jackson Lake Lodge, Moran, Wyo.
17. **Lisitsyn, N. A., E. D. Sverdlov, E. P. Moiseyeva, O. N. Danilevskaya, and V. G. Nikiforov.** 1984. Mutation to rifampin resistance at the beginning of the RNA polymerase beta subunit gene in *Escherichia coli*. Mol. Gen. Genet. **196:**173–174.
18. **Nair, J., D. A. Rouse, G.-H. Bai, and S. L. Morris.** 1993. The *rpsL* gene and streptomycin resistance in single and multiple drug-resistant strains of *Mycobacterium tuberculosis*. Mol. Microbiol. **10:**521–527.
19. **Ovchinnikov, Y. A., G. S. Monastyrskaya, V. V. Gubanov, V. M. Lipkin, E. D. Sverdlov, I. F. Kiver, I. A. Bass, S. Z. Mindlin, O. N. Danilevskaya, and R. B. Khesin.** 1981. Primary structure of *Escherichia coli* RNA polymerase. Nucleotide substitution in the beta subunit gene of the rifampicin resistant *rpoB255* mutant. Mol. Gen. Genet. **184:**536–538.
20. **Ovchinnikov, Y. A., G. S. Monastyrskaya, S. O. Guriev, N. F. Kalinina, E. D. Sverdlov, A. I. Gragerov, I. A. Bass, I. F. Kiver, E. P. Moiseyeva, V. N. Igumnov, S. Z. Mindlin, V. G. Nikiforov, and R. B. Khesin.** 1983. RNA polymerase rifampicin resistance mutations in *Escherichia coli*: sequence changes and dominance. Mol. Gen. Genet. **190:**344–348.
21. **Siddiqi, S., J. Libonati, and G. Middlebrook.** 1981. Evaluation of a rapid method for drug susceptibility testing of *Mycobacterium tuberculosis*. J. Clin. Microbiol. **13:**908–912.
22. **Takiff, H. E.** Personal communication.
23. **Telenti, A., P. Imboden, F. Marchesi, D. Lowrie, S. Cole, M. J. Colston, L. Matter, K. Schopfer, and T. Bodmer.** 1993. Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. Lancet **341:**647–650.
24. **van Embden, J. D. A., M. D. Cave, J. T. Crawford, J. W. Dale, K. D. Eisenach, B. Gicquel, P. Hermans, C. Martin, R. McAdam, T. M. Shinnick, and P. M. Small.** 1993. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. J. Clin. Microbiol. **31:**406–409.