JOURNAL OF BACTERIOLOGY, Nov. 1994, p. 6738–6743
0021-9193/94/$04.00+0

Vol. 176, No. 21

Isolation of Ribonucleotide Reductase from *Mycobacterium tuberculosis* and Cloning, Expression, and Purification of the Large Subunit

FUDE YANG, GUIZHEN LU, AND HARVEY RUBIN*

Infectious Disease Section, Department of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania

Received 14 July 1994/Accepted 31 August 1994

Ribonucleotide reductase, an allosterically regulated, cell cycle-dependent enzyme catalyzing a unique step in the synthesis of DNA, the reduction of 2′-ribonucleotides to 2′-deoxyribonucleotides, was purified 500-fold from *Mycobacterium tuberculosis* Erdman strain through cell disruption, ammonium sulfate fractionation, and dATP-Sepharose affinity column chromatography. As in eucaryotes and certain bacteria and viruses, the *M. tuberculosis* enzyme consists of two nonidentical subunits, R1 and R2, both of which are required for activity. R1 has a molecular mass of 84 kDa, as identified by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and photoaffinity labeling with dATP. The amino acid sequences of the N-terminal peptide and two internal peptides were determined, and a partial R1 gene was isolated by PCR with primers designed from these amino acid sequences. Additional coding sequences were isolated by screening size-selected libraries, and a full-length form of *M. tuberculosis* R1 was generated by PCR amplification of high-molecular-weight *M. tuberculosis* DNA and expressed in *Escherichia coli*. This coding sequence is 2,169 nucleotides long and contains no introns. The predicted molecular mass of R1 from the DNA sequence is 82,244 Da. Recombinant *M. tuberculosis* R1, purified to homogeneity, was biochemically active when assayed with extracts of *M. tuberculosis* enriched for R2.

Tuberculosis in all of its manifestations remains, globally, the leading cause of death from a single infectious agent (20). Studies from two urban centers indicate that between 30 and 40% of new cases are the result of recent infection, not 10% as had been thought prior to the publication of these epidemiological investigations. Furthermore, recently transmitted cases accounted for almost two-thirds of drug-resistant tuberculosis (1, 19). Clearly, new approaches to the development of antituberculous therapy are necessary. Inactivation of ribonucleotide reductase (RR), the cell cycle-regulated, allosteric enzyme that catalyzes the reduction of nucleoside diphosphates (NDPs) to deoxynucleoside diphosphates (dNDPs) (14), may be a particularly attractive target for new antituberculous agents. This enzymatic activity is the first step in DNA synthesis and has therefore been recognized as a primary target in the design of cancer chemotherapeutic agents. Furthermore, RR is gaining wide acceptance as a target for antiviral agents (13) and possibly even antiparasitic chemotherapy (16). There are compelling rationales for identifying *Mycobacterium tuberculosis* RR as a potential drug target: (i) the reduction of NDPs cannot be bypassed by a complementary activity arising either in the bacteria or in the host and (ii) there are ample data indicating that inhibition of RR in a variety of mycobacterial species substantially alters the growth patterns of the organisms. For example, studies in the 1960s and 1970s showed that *Mycobacterium smegmatis* cultured in iron-depleted media displayed altered, elongated morphology with decreased DNA synthesis and increased activity of DNA repair enzymes (32). When grown in the presence of the radical scavenger hydroxyurea, *M. smegmatis* contained a decreased DNA/protein ratio, with an increase in DNA polymerase and ATP-dependent DNAase activities, measured in crude extracts (31). The authors speculated that their results were consistent with the inhibition of mycobacterial RR. More directly, hydroxyurea prevented the growth of *M. smegmatis* at 200 μg/ml and is partially inhibitory for growth of the organism at 50 μg/ml (31). Other investigations have shown that two heterocyclic hydrazone inhibitors of RR have MICs against the virulent H37Rv strain of *M. tuberculosis* of 80 and 43 μM/liter (18).

RR derived from eucaryotes, *Plasmodium falciparum*, certain viruses, and *Escherichia coli* is a two-subunit, allosterically regulated enzyme with an $\alpha 2\beta 2$ quaternary structure (7). Substrate and effector binding sites have been localized to R1, the large subunit. The predominant catalytic mechanism is dependent upon the formation of a tyrosyl radical stabilized by a dinuclear iron center located in R2 (14). Each subunit is inactive when assayed individually. The genes encoding the two subunits are located on an operon in *E. coli* with the large-subunit gene (*nrdA*) 5′ to the small-subunit gene (*nrdB*) (22, 25).

As part of an investigation of the regulation of mycobacterial growth and DNA synthesis by RR, we report the purification of this enzyme from the Erdman strain of *M. tuberculosis* as well as the cloning, expression, and biological activity of the large subunit.

MATERIALS AND METHODS

Materials. [5-^{3}H]CDP, [8,5-^{3}H]GDP, [8-^{3}H]ADP, and [α-^{32}P]dATP were purchased from Amersham. All cold NDPs and nucleoside triphosphates (NTPs) were from Sigma. Sepharose 4B was purchased from Pharmacia. Phenylboronate Sepharose (PBA-60) was purchased from Amicon. The molecular weight markers for sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) were purchased from Bio-Rad. Irradiated *M. tuberculosis* (Erdman strain) cell paste and *M. tuberculosis* DNA were provided by P. J. Brennan, Department of Microbiology, Colorado State University.

Purification of *M. tuberculosis* RR. Twenty grams of *M. tuberculosis* cell paste was washed once with 100 ml of 50 mM Tris-HCl–5 mM $MgCl_2$–0.1 mM dithiothreitol, pH 7.6 (buffer A), resuspended in 200 ml of buffer A containing 2 mM phenylmethylsulfonyl fluoride, and subjected to two rounds of

* Corresponding author. Mailing address: Department of Medicine, University of Pennsylvania, Philadelphia, PA 19104. Phone: (215) 662 6475. Fax: (215) 662 7842. Electronic mail address: (bitnet) harvey @slack.med.upenn.edu.

disruption in a prechilled French press. The cell debris was removed by centrifugation at 23,000 × *g* for 30 min. The supernatant was precipitated by addition of 10% streptomycin sulfate in buffer A to a final concentration of 0.5%. The resulting suspension was stirred for an additional 10 min, and the precipitate was removed by centrifugation (23,000 × *g*, 20 min). Solid ammonium sulfate was slowly added to the supernatant to 60% saturation with stirring. After the addition was completed, the suspension was stirred for 10 min and the precipitate was collected by centrifugation (23,000 × *g*, 20 min) and resuspended in 15 ml of buffer A. The suspension was dialyzed against the same buffer for 5 h with one buffer change. The dialysate (referred to hereafter as partially purified enzyme) was centrifuged at 13,800 × *g* for 5 min and then applied onto a dATP-Sepharose column (1.0 by 3.0 cm) at room temperature in small aliquots. dATP-substituted Sepharose gel was prepared essentially according to the published method (2). The column was then washed with 10 column volumes of buffer A. RR was eluted with 10 ml of buffer A containing 10 mM ATP, concentrated to 200 μl with Centriprep-10 (Amicon), and stored at −70°C (the product is referred to hereafter as highly purified enzyme).

RR activity assay. For the RR assay, the method of Steeper and Steuart (21) was modified to directly separate the deoxyribonucleotide product from the reaction mixture over a phenylboronate agarose (PBA-60) gel (12). The reaction mixture, made up in a final volume of 100 μl of 60 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid; pH 7.6) buffer, contained 8 mM magnesium acetate, 8.75 mM NaF, 0.05 mM $FeCl_3$, 25 mM dithiothreitol, and various amounts of effector and ^{3}H-NDP substrate. The reaction was started by the addition of the enzyme (either partially purified or highly purified), carried out at 37°C, and stopped by heating in a boiling water bath for 3 min. The denatured protein was removed by centrifugation. The supernatant was diluted with an equal volume of 50 mM Tris-HCl buffer (pH 8.5) containing 50 mM magnesium chloride and applied onto a PBA-60 column (0.5 by 6.0 cm) which was preequilibrated with the same buffer. The column was then washed with 5 ml of the same buffer. The quantity of deoxyribonucleotide was determined by liquid scintillation. The column was regenerated by a washing with 10 ml of 50 mM sodium citrate buffer (pH 6.5) and double-deionized water. All assays were carried out in triplicate.

Photoaffinity labeling of *M. tuberculosis* RR with [α-^{32}P] dATP. Partially purified RR (30 μg) or pure RR (3 μg) in 20 ml of buffer A was mixed with 16 pmol of [α-^{32}P]dATP (3,000 Ci/mmol) in the presence or absence of 5 mM ATP or 2.5 mM CDP, and the mixture was incubated on ice for 5 min. The mixture was placed as a drop on Parafilm on dry ice and irradiated for 30 min with a UVP Inc. UV minerallight model UVGL-58 lamp. After irradiation, the protein was precipitated with 5% trichloroacetic acid and washed twice with buffer A containing 5% trichloroacetic acid. The protein was then dissolved in loading buffer and analyzed on 12% SDS slab gels. The stained and dried gels were autoradiographed at room temperature for 5 h.

N-terminal and internal amino acid sequence analysis. Highly purified *M. tuberculosis* RR (30 μg) was subjected to preparative SDS-PAGE (12% polyacrylamide gel) and blotted onto a polyvinylidene difluoride membrane (Immobilon-P; Millipore) in 12.5 mM Tris–95 mM glycine–10% MeOH, pH 8.6, at 4°C (100 V, 1 h). The membrane was washed with double-distilled water and stained for 5 min with 0.25% Coomassie blue R250 in 40% MeOH and destained for 10 min with 50% MeOH. The membrane was vacuum dried, and the protein band corresponding to *M. tuberculosis* R1 was submitted for N-terminal and internal sequence analysis to The Wistar Institute Protein Core Facility (Philadelphia, Pa.).

TABLE 1. Purification of RR from *M. tuberculosis* Erdman strain[a]

Purification step	Total protein (mg)	Sp act (U)	Total activity (U)	Protein recovery (%)	Activity recovery (%)
Crude extract	560			100	
60% $(NH_4)_2SO_4$	240	2	480	43	100
dATP-Sepharose	0.08	1,000	80	0.033	17

[a] Data are based on 20 g of cell paste.

Isolation of a partial sequence of the *M. tuberculosis* R1 gene. PCR with primers designed on the basis of internal amino acid sequences was carried out in a total volume of 100 μl which contained 0.25 μg of *M. tuberculosis* genomic DNA, 100 pmol of primers, all four dNTPs (each at 0.2 mM), 10 ml of 10× PCR buffer (Perkin-Elmer), and 2.5 U of *Taq* polymerase. The reaction was carried out in 20 cycles of the following program: 20 s at 94°C, 30 s at 45°C, and 60 s at 72°C. The PCR product was purified from an agarose gel by using Qiaex silica gel particles (Qiagen) according to the manufacturer's protocol.

Expression and activity of recombinant *M. tuberculosis* R1 produced in *E. coli*. The R1 gene was isolated from high-molecular-weight *M. tuberculosis* DNA by PCR using the following primers that contained the indicated *Nhe*I cloning sites (underlined): N primer, 5'-AAAAAA<u>GCTAGC</u>CCCAC CGTGATCGCCGAGCCCGTAGCCTC; and C primer, 5'-A AAAAA<u>GCTAGC</u>CTACAGCATGCAGGA. The PCR mixture, in a total volume of 100 μl, contained 0.25 μg of *M. tuberculosis* genomic DNA, 100 pmol of each primer, all four dNTPs (each at 0.2 mM), and 2.5 U of *Taq* polymerase. The reaction was carried out in 30 cycles of the following program: 20 s at 94°C, 20 s at 55°C, and 90 s at 72°C. The PCR product was gel purified with Qiaex silica gel particles, digested with *Nhe*T, phenol extracted, and precipitated with ethanol. The cloning vector containing the heat-inducible p_L promoter described previously (17) was prepared by digestion with *Nhe*T, treated with alkaline phosphatase, phenol extracted, and precipitated with ethanol. A 28-ng sample of *M. tuberculosis* R1 DNA prepared as described above was ligated with *Nhe*T-digested pZMs (15) (15 ng) in a final volume of 10 μl containing 400 U of T4 DNA ligase and 1 ml of 10× ligation buffer at 16°C overnight. The ligation mix was then used to transform N4830 (Pharmacia, Piscataway, N.J.) competent cells and plated onto Luria-Bertani agar supplemented with ampicillin. *M. tuberculosis* R1 was expressed by heat induction at 42°C. The purification of the recombinant R1 was essentially the same as that of the wild-type R1 from *M. tuberculosis*.

Nucleotide sequence accession number. The GenBank/EMBL accession number for the DNA sequence of *M. tuberculosis* R1 is L34407.

RESULTS AND DISCUSSION

Purification of *M. tuberculosis* RR. The results of the purification scheme are summarized in Table 1. Twenty grams of cell paste yielded 80 μg of protein with a specific activity of 1,000 U (nanomoles of product per milligram of protein per hour). RR activity was not detected in the crude extract; however, it was detectable in the 60% ammonium sulfate fraction and was stimulated by addition of ATP and inhibited by dATP. On the basis of this finding and the observation that mammalian RR as well as *E. coli* RR was purified by dATP affinity chromatog-

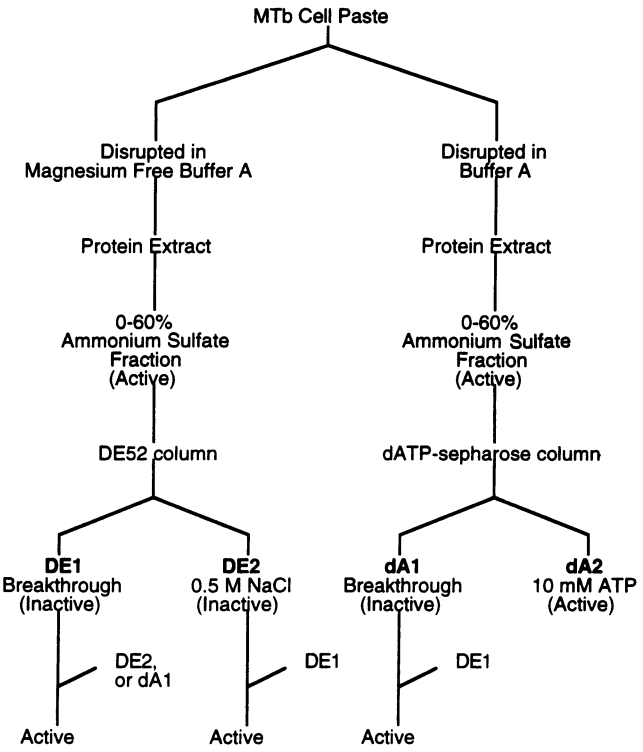

FIG. 1. Scheme for assays of RR activity. MTb, *M. tuberculosis*.

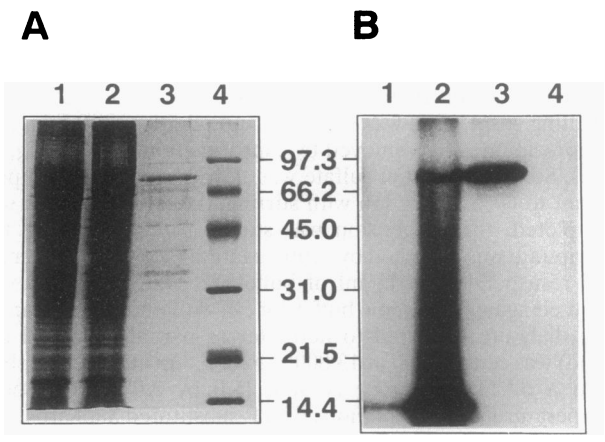

FIG. 2. SDS-PAGE analysis of dATP photoaffinity-labeled *M. tuberculosis* RR. (A) Coomassie blue-stained gel. Lanes 1 and 2, dATP-labeled partially purified RR in the presence of 5 mM ATP and, 2.5 mM CDP, respectively; lane 3, dATP-labeled highly purified RR in the presence of 2.5 mM CDP; lane 4, molecular mass markers (in kilodaltons). (B) Autoradiogram of the same gel as in panel A.

raphy (6, 24), *M. tuberculosis* RR was purified 500-fold with a dATP-Sepharose affinity column (Table 1).

RR activity found in the 60% ammonium sulfate fraction was resolved into two components by DE52 column chromatography with Mg^{2+}-free buffer A. The two fractions, one in the breakthrough fraction (DE1) and a second in the 0.5 M NaCl fraction (DE2), lacked RR activity when assayed individually. The breakthrough fraction of dATP affinity chromatography (dA1), which contained no RR activity but was rich in R2, was able to restore RR activity to DE1 but not to the DE2 fraction, indicating that DE1 contains R1 (Fig. 1).

The enzyme was stable throughout the purification. However, activity decreased after 1 month of storage at −70°C if the concentration of the protein was lower than 1 mg/ml. The partially purified enzyme was stable throughout the 4-h incubation during the activity assay in the presence of substrate and effectors.

SDS-PAGE of the dATP-Sepharose affinity-purified material showed one major band with a molecular mass of 84,000 Da. This band was specifically labeled by [α-^{32}P]dATP in the presence of 2 mM CDP and was completely inhibited by 5 mM ATP (Fig. 2), which provided additional evidence that the protein was R1 (3).

Activity of *M. tuberculosis* RR. *M. tuberculosis* RR utilized all four ribonucleoside diphosphates as substrates. The reduction of CDP and UDP could be detected in 60% ammonium sulfate precipitate, whereas reduction of ADP and GDP required the use of the dATP affinity-purified material. Maximum activity (2 nmol of dCDP/h/mg of protein) of partially purified enzyme for CDP reduction was obtained in the presence of 6 mM ATP. In the presence of dGTP (6 mM) and ATP (3 mM), 1.8 μg of the highly purified enzyme reduced 50 pmol of dADP in 3 h. The same amount of dGDP was produced by equal concentrations of highly purified enzyme in the presence of dTTP (1.5 mM) and ATP (3 mM). The reduction of all four NDPs was inhibited by dATP.

Identification of the gene encoding *M. tuberculosis* R1. Sufficient quantities of purified enzyme were generated to obtain N-terminal and internal amino acid sequence data in order to design PCR primers. The results of the amino acid determinations, with the corresponding positions, are shown in Table 2.

A fragment of 908 bp of R1 gene was isolated by PCR using primers corresponding to peptide 2 [5′-GA(G/A)TTCTTCCA (G/A)AC] and peptide 3 (5′-GCGTAGGTGTCGATGAT). The 906-bp fragment was used to probe *Eco*RI-digested high-molecular-weight *M. tuberculosis* DNA. Two bands, 1.1 and 2 kb, were observed on the Southern blot. Two size-selected libraries were generated in lambda ZAP II, one containing inserts of 1.1 kb and one containing inserts of 2.0 kb. Plaques were screened with the 908-bp fragment, positive plaques were picked, and the plasmid containing the insert was rescued. The 2-kb fragment contained 548 bp of coding region including a potential C terminus, 358 bp of which overlapped with the 908-bp probe. The 1.1-kb fragment contained coding region 5′ to that contained within the 2-kb fragment but did not extend all the way to the N terminus. The N-terminal 522-bp fragment

TABLE 2. N-terminal and internal sequences of *M. tuberculosis* R1

Peptide	Source	Sequence
1	N terminus	P-1TVIAEPVASGAHASYSGGPGETDYHALNA-30
2	Internal sequence 1	E-358FFQTLAELQFESGYPYIMFEDTVN-382
3	Internal sequence 2	I-655DTYAAATQHVDQG-669

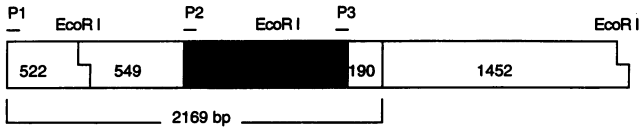

FIG. 3. Organization of the cloning strategy for *M. tuberculosis* R1. The initial 908-bp fragment was generated by PCR from internal amino acid sequence data. Two *Eco*RT fragments of 1.1 and 2 kb provided all but the N-terminal region, which was subsequently obtained as a PCR product by using the results of amino acid analysis of the N terminus and an internal site. See text for details.

was isolated by PCR using primers corresponding to peptides 1 [CCCACCGT(G/C)ATCGCCGAGCC(C/G)GT] and 2 (AG GGTCTGGAAGAACTC). Peptide 1 was the sequence determined from N-terminal analysis of highly purified *M. tuberculosis* R1 and therefore may represent a processed form of R1. In this regard, R1 proteins with heterogeneous N termini and identical activities have been isolated from *E. coli*, suggesting that the N terminus does not play a central role in either the catalytic or the regulatory activity (23). Figure 3 summarizes the organization of the cloning strategy.

The nucleotide sequence of the 2,169-bp R1 gene (Fig. 4) encodes a protein of 723 amino acids with a calculated molecular mass of 82,244 Da. The coding region is 59% G+C, with the third position of the codon 70% GC rich. The 3′ noncoding region is 63% GC rich.

The alignment of *M. tuberculosis* R1 with human R1 and *E. coli* R1 is shown in Fig. 5. The derived gene product contains five cysteines that are highly conserved and which are thought to be essential. Cys-187 and Cys-424 (Cys-225 and Cys-462 in *E. coli*) correspond to the cysteines proposed to be involved in the active site (8, 10, 11). Cys-718 and Cys-721 (Cys-754 and Cys-759 in *E. coli*) align with the C-terminal consensus sequence -Cys-x-x(-x-x)-Cys-x(-x)-COOH that has been proposed to be involved in shuttling electrons from thioredoxin to the active-site cysteines (9). Cys-398 could be aligned with Cys-439 in *E. coli*, which was suggested to function as the radical in R1 that initiates catalysis by abstraction of the 3′ H from the substrate (11).

In addition to the alignment of the five cysteines, there are two regions that are highly conserved among mammals, *E. coli*, *P. falciparum* (16), and *M. tuberculosis*, namely, I-476GLG-479 and K-698TLYY-702. The function of these consensus areas is not clear; however, the recent X-ray structure of *E. coli* R1 indicates that Y-730 and Y-731, which correspond to *M. tuberculosis* Y-701 and Y-702, respectively, may be involved in the radical transfer reaction (26). Protein alignment analysis together with the result of NDP reduction activity strongly suggests that *M. tuberculosis* R1 belongs to RR class I (4).

```
CCC ACC GTG ATC GCC GAG CCC GTA GCC TCC GGC GCG CAC GCC TCT TAC TCT GGG      54
Pro Thr Val Ile Ala Glu Pro Val Ala Ser Gly Ala His Ala Ser Tyr Ser Gly      18

GGG CCG GGC GAA ACG GAC TAT CAC GCG CTG AAC GCG ATG CTG AAC CTG TAC GAC     108
Gly Pro Gly Glu Thr Asp Tyr His Ala Leu Asn Ala Met Leu Asn Leu Tyr Asp      36

GCG GAC GGC AAG ATC CAG TTC GAC AAG GAT CGG GAA GCA GCC CAC CAG TAC TTT     162
Ala Asp Gly Lys Ile Gln Phe Asp Lys Asp Arg Glu Ala Ala His Gln Tyr Phe      54

TTG CAG CAT GTC AAT CAG AAC ACG GTC TTC TTC CAT AAT CAG GAC GAG AAG CTC     216
Leu Gln His Val Asn Gln Asn Thr Val Phe Phe His Asn Gln Asp Glu Lys Leu      72

GAC TAC CTG ATC CGC GAG AAT TAC TAC GAG CGT GAG GTT CTC GAC CAG TAC TCG     270
Asp Tyr Leu Ile Arg Glu Asn Tyr Tyr Glu Arg Glu Val Leu Asp Gln Tyr Ser      90

CGC AAC TTC GTC AAG ACG CTG CTA GAC CGC GCC TAC GCC AAA AAG TTC CGG TTT     324
Arg Asn Phe Val Lys Thr Leu Leu Asp Arg Ala Tyr Ala Lys Lys Phe Arg Phe     108

CCG ACG TTT TTG GGT GCG TTC AAG TAC TAC ACC TCC TAC ACG CTG AAA ACC TTT     378
Pro Thr Phe Leu Gly Ala Phe Lys Tyr Tyr Thr Ser Tyr Thr Leu Lys Thr Phe     126

GAC GGG AAG CGC TAT CTG GAG CGC TTC GAG GAC CGC GTG GTC ATG GTG GCG CTA     432
Asp Gly Lys Arg Tyr Leu Glu Arg Phe Glu Asp Arg Val Val Met Val Ala Leu     144

ACG TTG GCC GCC GGC GAT ACC GCA CTT GCC GAG CTG CTG GTC GAC GAG ATC ATC     486
Thr Leu Ala Ala Gly Asp Thr Ala Leu Ala Glu Leu Leu Val Asp Glu Ile Ile     162

GAC GGC CGC TTC CAG CCC GCC ACA CCG ACG TTT TTG AAT TCT GGC AAG AAG CAG     540
Asp Gly Arg Phe Gln Pro Ala Thr Pro Thr Phe Leu Asn Ser Gly Lys Lys Gln     180

CGC GGG GAG CCC GTG AGC TGT TTT TTG CTT CGC GTC GAA GAT AAC ATG GAG TCG     594
Arg Gly Glu Pro Val Ser Cys Phe Leu Leu Arg Val Glu Asp Asn Met Glu Ser     198

ATC GGA CGG TCG ATC AAC TCC GCG CTG CAG CTA TCC AAG CGT GGC GGG GGA GTG     648
Ile Gly Arg Ser Ile Asn Ser Ala Leu Gln Leu Ser Lys Arg Gly Gly Gly Val     216

GCG TTG CTG CTG ACC AAC ATT CGC GAG CAC GGC GGC GCC ATC AAG AAC ATC GAG     702
Ala Leu Leu Leu Thr Asn Ile Arg Glu His Gly Gly Ala Ile Lys Asn Ile Glu     234

AAC CAG TCC TCG GGC GTC ATC CCC ATC ATG AAG TTG CTG GAG GAT GCG TTC TCC     756
Asn Gln Ser Ser Gly Val Ile Pro Ile Met Lys Leu Leu Glu Asp Ala Phe Ser     252

TAC GCC AAC CAG CTG GGC GCT CGT CAA GGT GCC GGC GCG GTG TAC CTG CAC GCC     810
Tyr Ala Asn Gln Leu Gly Ala Arg Gln Gly Ala Gly Ala Val Tyr Leu His Ala     270

CAT CAC CCC GAC ATC TAC CGA TTC CTG GAC ACC AAG CGT GAG AAC GCC GAC GAG     864
His His Pro Asp Ile Tyr Arg Phe Leu Asp Thr Lys Arg Glu Asn Ala Asp Glu     288

AAG ATC CGG ATC AAG ACG CTG AGT CTG GGG GTG GTG ATC CCC GAC ATC ACC TTC     918
Lys Ile Arg Ile Lys Thr Leu Ser Leu Gly Val Val Ile Pro Asp Ile Thr Phe     306

GAG TTG GCC AAG CGC AAC GAT GAC ATG TAC CTG TTC TCG CCC TAC GAT GTC GAG     972
Glu Leu Ala Lys Arg Asn Asp Asp Met Tyr Leu Phe Ser Pro Tyr Asp Val Glu     324

CGG GTC TAC GGT GTG CCG TTC GCT GAC ATC TCG GTC ACC GAG AAG TAC TAC GAA    1026
Arg Val Tyr Gly Val Pro Phe Ala Asp Ile Ser Val Thr Glu Lys Tyr Tyr Glu     342

ATG GTC GAT GAC GCG CGC ATC CGC AAG ACC AAG ATC AAG GCA CGG GAG TTC TTC    1080
Met Val Asp Asp Ala Arg Ile Arg Lys Thr Lys Ile Lys Ala Arg Glu Phe Phe     360

CAG ACG CTG GCC GAG CTG CAG TTC GAG TCC GGC TAC CCC TAT ATC ATG TTC GAA    1134
Gln Thr Leu Ala Glu Leu Gln Phe Glu Ser Gly Tyr Pro Tyr Ile Met Phe Glu     378

GAC ACC GTC AAT CGC GCT AAT CCA ATT GAT GGC AAG ATC ACG CAC AGC AAC CTG    1188
Asp Thr Val Asn Arg Ala Asn Pro Ile Asp Gly Lys Ile Thr His Ser Asn Leu     396

TGC TCG GAG ATC CTG CAA GTG TCT ACG CCG TCA TTG TTC AAC GAG GAC TTG TCG    1242
Cys Ser Glu Ile Leu Gln Val Ser Thr Pro Ser Leu Phe Asn Glu Asp Leu Ser     414

TAT GCC AAA GTG GGC AAA GAC ATT TCG TGC AAC CTG GGG TCG CTG AAC ATC GCC    1296
Tyr Ala Lys Val Gly Lys Asp Ile Ser Cys Asn Leu Gly Ser Leu Asn Ile Ala     432

AAG ACG ATG GAC TCG CCG GAC TTC GCG CAG ACG ATC GAG GTG GCG ATC CGC GCG    1350
Lys Thr Met Asp Ser Pro Asp Phe Ala Gln Thr Ile Glu Val Ala Ile Arg Ala     450

TTG ACC GCG GTG AGG CAC CAA ACC CAT ATC AAG TCG GTG CCC TCA ATC GAG CAG    1404
Leu Thr Ala Val Arg His Gln Thr His Ile Lys Ser Val Pro Ser Ile Glu Gln     468

GGC AAC AAC GAC TCC CAC GCG ATC GGG CTA GGA CAG ATG AAC CTG CAC GGC TAC    1458
Gly Asn Asn Asp Ser His Ala Ile Gly Leu Gly Gln Met Asn Leu His Gly Tyr     486

CTG GCC CGG GAA CGC ATC TTC TAC GGA TCC GAC GAA GGC ATC GAC TTC ACC AAC    1512
Leu Ala Arg Glu Arg Ile Phe Tyr Gly Ser Asp Glu Gly Ile Asp Phe Thr Asn     504

ATC TAC TTC TAT ACG GTG CTG TAT CAC GCG TTG CGG GCA TCC AAC CGC ATC GCG    1566
Ile Tyr Phe Tyr Thr Val Leu Tyr His Ala Leu Arg Ala Ser Asn Arg Ile Ala     522

ATC GAA CGC GGC ACG CAC TTC AAG GGT TTC GAG CGG TCC AAG TAC GCG TCC GGG    1620
Ile Glu Arg Gly Thr His Phe Lys Gly Phe Glu Arg Ser Lys Tyr Ala Ser Gly     540

GAA TTC TTC GAC AAG TAC ACC GAC CAG ATT TGG GAG CCG AAG ACC CAG AAG GTA    1674
Glu Phe Phe Asp Lys Tyr Thr Asp Gln Ile Trp Glu Pro Lys Thr Gln Lys Val     558

CGC CAG CTG TTC GCC GAC GCC GGC ATC CGC ATC CCA ACG CAG GAC GAC TGG CGT    1728
Arg Gln Leu Phe Ala Asp Ala Gly Ile Arg Ile Pro Thr Gln Asp Asp Trp Arg     576

CGG CTC AAG GAG TCG GTG CAA GCG CAC GGC ATC TAC AAC CAG AAC CTG CAG GCG    1782
Arg Leu Lys Glu Ser Val Gln Ala His Gly Ile Tyr Asn Gln Asn Leu Gln Ala     594

GTG CCG CCG ACC GGG TCG ATT TCC TAC ATC AAC CAT TCG ACG TCG TCG ATT CAC    1836
Val Pro Pro Thr Gly Ser Ile Ser Tyr Ile Asn His Ser Thr Ser Ser Ile His     612

CCG ATC GTG TCG AAG GTC GAG GTC CGC AAG GAA GGC AAG ATC GGG CGG GTC TAC    1890
Pro Ile Val Ser Lys Val Glu Val Arg Lys Glu Gly Lys Ile Gly Arg Val Tyr     630

TAC CCG GCG CCG TAT ATG ACC AAC GAC AAC CTG GAG TAC TAC GAA GAC GCC TAC    1944
Tyr Pro Ala Pro Tyr Met Thr Asn Asp Asn Leu Glu Tyr Tyr Glu Asp Ala Tyr     648

GAG ATC GGT TAC GAG AAG ATC ATC GAC ACC TAC GCG GCG GCC ACC CAG CAT GTG    1998
Glu Ile Gly Tyr Glu Lys Ile Ile Asp Thr Tyr Ala Ala Ala Thr Gln His Val     666

GAT CAA GGG CTT TCG CTG ACG TTG TTC TTC AAA GAC ACC GCC ACC ACC CGC GAC    2052
Asp Gln Gly Leu Ser Leu Thr Leu Phe Phe Lys Asp Thr Ala Thr Thr Arg Asp     684

GTG AAC AAG GCG CAG ATT TAC GCC TGG CGC AAG GGG ATC AAG ACG CTG TAC TAC    2106
Val Asn Lys Ala Gln Ile Tyr Ala Trp Arg Lys Gly Ile Lys Thr Leu Tyr Tyr     702

ATC CGG CTG CGG CAG ATG GCG TTG GAG GGC ACC GAG GTC GAG GGT TGC GTG TCC    2160
Ile Arg Leu Arg Gln Met Ala Leu Glu Gly Thr Glu Val Glu Gly Cys Val Ser     720

TGC ATG CTG TAG                                                           2172
Cys Met Leu ***                                                           723
```

FIG. 4. DNA sequence and derived amino acid sequence of *M. tuberculosis* R1. Underlined residues correspond to the peptides identified in the amino acid analysis (Table 2).

```
          1          10         20         30         40         50
Human     Mhv....iKRDGrqERvmfDKItsriqklcyGLnmdfvdpaQitmkviqgLYsGvtTvel
E.coli    MnqnllvtKRDGstERinlDKIhrvldwaafGLhnvsi..sQvelrshiqfYDGikTsdi
MTb                ptvia.epvasgahasysGgpgft...dyhalnamlnlYDadgkiqf
                   1         10         20         30         40

            60         70         80         90        100        110
Human     DtlaaetAAtLttkh.PDYaiLAARIAvsnLhKetKkvFsdvmedlynYinphnGKhspm
E.coli    hetiikAAAdlisrdaPDYQyLAARLAiFHLrKkaygqFeppalydhvvkmvemGKYdnh
MTb       Dkdreaahqyfl.......Qhvnqntvf FH.nqdeKldy...lirenyYerevldqYsrn
               50                 60         70         80         90

            120        130        140        150        160        170
Human     vaKstLDivlAnKDRlnsaIiyDRDFsY.nYfgfKTLErsYLL..kinGkvaErpQhM..
E.coli    lledyteeefkqmDtF...IdhdRDmtf.sYaavKqLEGKYLvqnrvtGEiyEsaQflyi
MTb       fvKtlLDrayAkKfRF...ptflgaFkYytsytlKTfdGKryL......ErfEdrvvMva
               100        110        120        130        140

                  180        190        200        210        220
Human     Lmrvsv..gihkEdidAaiETY.nllSerwFthasPT.LfnaGTnRPQLSSCFLLsmkdD
E.coli    LvaAclfsnyprEtrlqyvkrfyDavStfkislpTPi.msgvrTptrQFSSCVLiec.gD
MTb       LtlA........AgdtAlaEllvDeiidgrFqpaTPTfLnsgkkqRgepvSCFLLrv.eD
                     150        160        170        180        190

            230        240        250        260        270        280
Human     SiEgIydtlkqcaliSKsaGGIGvavscIRAtGSyIaGtngnSnGlvPmlrvynntaryv
E.coli    SldSInadssaivkyvsqraGIGinagrIRAlGSpIrGgEafhtGcIPfyKhfqtAvksc
MTb       nmESIgrsinsalqlSKrgGGvallltnIRehGgaIkniEnqSsGvIPimKlledAfsya
               200        210        220        230        240        250

            290        300        310        320        330        340
Human     dQGGnkRpGAfAiyLepWHLDIfeFLDLKkNtGkEeqRaRdlffalwIPDLfmkRvetNq
E.coli    sQGGv.RgGAatlfypmWHLeveslLvLKnNrGvEgnRvRhmdyGVqInkLmytRllkge
MTb       nQlGa.RqGAgAvyLhahHpDIyrFLDtKrenadEkiRiktlslGVvIPDitfelakrNd
               260        270        280        290        300        310

            350        360        370        380        390        400
Human     DwsLmcPnecPGLdeVw...gEEFEKLYasYEKqgRvRKvv.KAqqLwyaiiEsQtETGt
E.coli    DitLFSPsDVPGLYdafFADQEEFErLYtkYEKDdsIRKqrvKAvELFslmmqerasTGr
MTb       DmyLFSPyDVervYgVpFADisvtEKyYemvd.DaRIRKtkiKArEfFqtlaElQfEsGy
               320        330        340        350        360        370

                  410        420        430                   440        450
Human     PYmlykDsCNRkSn.qqnlGTIKCSNLCtEIveyTSk.............dEvAvCNLASL
E.coli    iYIqnvDhCNthSPfDpaiaPvrqSNLClEIalpTKP....lndvndenGEiAlCtLsaf
MTb       PYImfeDtvNRanPiD...GkIthSNLCsEIlqvstPslfnedlsyakvGkdisCNLgSL
               380        390        400        410        420

                  460        470        480        490        500
Human     alnmyvtse..htydfkkLaEvtkVvVRnLnki.IDinyYPVPeAclsNkrhrpIGIGVq
E.coli    ........Nlgainnl deLeElailAVRALdAl.lDyqdYPiPaAkrGamgrrtlGIGVi
MTb       ........Niakdmdspdfaqtie VAiRALtAvrhqthiksVPsieqGNndshaIGlGqm
                        440        450        460        470        480

            510        520        530        540        550        560
Human     gLAdafilmrypfesaEAqlLnkqiFETIyYgALeAScdLAKEQGp.yetyEGSpvsKGIL
E.coli    NfAyYLAndgkrYsdgsAnnLThktFEaIqYylLkASNeLAKEQGa.cpwfnettYAKGIL
MTb       NLhgYLARerifYgsdEgidfTniyFyTvlYhALrASNriAiErGthfkgfErSkYAsGef
               490        500        510        520        530        540

                        570          580        590        600        610
Human     ..........qy.....Dmwnvtptd L.w..DWkvLKEkIakyGIRNSlLiApMPTaStaQ
E.coli    ..........PiptykkDLdtiAnepLhy..DWeaLrESIktHGlRNStLsAlMPsetsSQ
MTb       fdkytdqiwePktqkvrqLfadAgiriptqdDWrrLKESvqaHGIyNqnLqAvpPTgSiSy
               550        560        570        580        590        600

            620        630        640        650        660        670
Human     IlgnneSIEPytSniytRrvlsGefqivnphlLkdlterglwhEeMknqiiacNgsiqsip
E.coli    IsnaTngIEPprgyVsikaskdG........ILRqvvPd...yEhl.......hDayellw
MTb       InhsTsSIhPivSkVevRkegk.........IGRvyyPa....pyMt......NDnleyye
               610        620                630                640

                  680        690        700        710          720        730
Human     EiPddlkqlykyvweisqktvLkmaaergaFIDQSQSlNihiaepny..GKlTsmhfygwK
E.coli    EmPgn.............dgyLqlvgImqkfIDQSiSaNtnydpsrFPSGKvpmqqll..K
MTb       daye.................igyekIidtyaaatQhvdqglsltlFfkdtaTtrdvn..K
                           650        660        670        680

                        740        750        760        770        780
Human     qgl.......KTgmYYLrtrpaaNpiQftLnkeklkdkekvskeeeekerntaamvcslen
E.coli    dlltAykfGvKT.lYYq......NtrdgAedaqdd.....................lvpsiq
MTb       aqiyAwrkGiKT.lYYi......rlrQmALegte.........................
          690        700        710

                  790
Human     rDeC..lmCgs
E.coli    dDGCesgaCki
MTb       veGC..vsCml
                  720723
```

FIG. 5. Alignment of human, *E. coli*, and *M. tuberculosis* (MTb) R1 sequences. Cysteine residues that are highly conserved and considered essential are indicated (boldface).

There is no open reading frame in the 1,452 bp following the stop codon. In particular, no coding sequences corresponding to R2 were identified. Therefore, the genetic organization of the *M. tuberculosis* RR system is different from that of *E. coli*,

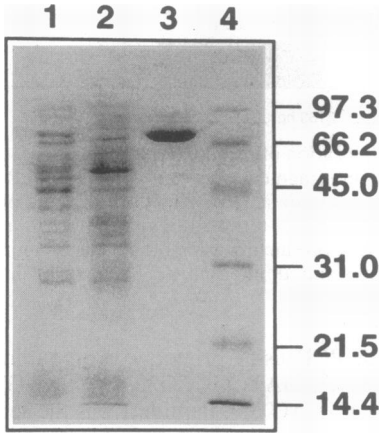

FIG. 6. Expression of recombinant *M. tuberculosis* R1 in *E. coli*. Lanes 1 and 2, whole-cell extract following heat induction or from uninduced cultures, respectively; lane 3, purified recombinant R1; lane 4, molecular mass markers (in kilodaltons).

which is organized as an operon with 182 bp separating the carboxy terminus of R1 and the amino terminus of R2 (4, 25).

Expression and activity of recombinant *M. tuberculosis* R1. Recombinant *M. tuberculosis* R1 was expressed in *E. coli* by using a heat-induced expression system (17). Recombinant *M. tuberculosis* was soluble (Fig. 6) and had the same molecular weight as R1 purified from *M. tuberculosis*, indicating little or no glycosylation. The recombinant R1 could also be photoaffinity labeled by [α-^{32}P]dATP in the presence of CDP. The activity of purified recombinant *M. tuberculosis* R1 assayed with dA1 was comparable to that of partially purified wild-type *M. tuberculosis* RR, indicating the authenticity of the recombinant gene product. (Fig. 7).

Very little is known about the biochemistry of DNA replication in mycobacteria. The mean generation time for *M. tuberculosis* is 24 h, compared with 3 h for *M. smegmatis* and 1.3 h for *E. coli*. Genomic DNA is replicated in approximately 10 h in *M. tuberculosis*, whereas the comparable times for *M. smegmatis* and *E. coli* are 1.8 and 1 h, respectively (30). The activity of *M. tuberculosis* R1 reported above reflects the

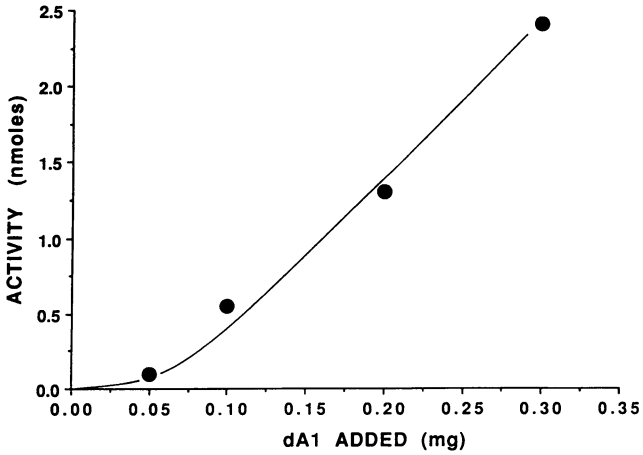

FIG. 7. CDP reduction by recombinant *M. tuberculosis* R1 (5 μg) assayed with increasing amounts of partially purified *M. tuberculosis* R2.

turnover of an enzyme system not optimized with saturating amounts of R2 and therefore cannot yet be used in evaluating the long doubling time of this organism. A series of papers by Wheeler (27–29) described de novo and scavenging pathways for purines and pyrimidines in *Mycobacteria avium*, *Mycobacterium microti*, and *Mycobacterium leprae*; however, the molecular characterization of the enzymes in these pathways has not yet been accomplished.

The *recA* gene, a regulatory component of the SOS system of response to nucleic acid damage and the essential element in homologous recombination, has been cloned and extensively studied in mycobacteria (5). While RR is induced by many of the same set of stimuli that induce the SOS response, regulation of expression and activity of RR in *E. coli* does not involve the products of the *recA*, *recB*, *recC*, or *lexA* genes. The availability of cloned RR subunit genes will allow research to progress in the important but up to now underinvestigated area of regulation of DNA replication in the mycobacteria.

ACKNOWLEDGMENTS

We thank Zhi mei Wang for helpful discussions and Jerry Salem for help with computer analysis.

This work was supported by Apollon Inc., Malvern, Pa.

REFERENCES

1. **Alland, D., G. E. Kalkut, A. R. Moss, R. A. McAdam, J. A. Hahn, W. Bosworth, E. Drucker, and B. R. Bloom.** 1994. Transmission of tuberculosis in New York City. N. Engl. J. Med. **330:**1710–1716.
2. **Berglund, O., and F. Eckstein.** 1972. Synthesis of ATP- and dATP substituted Sepharoses and their application in the purification of phage-T4-induced ribonucleotide reductase. Eur. J. Biochem. **28:** 492–496.
3. **Caras, I. W., and D. W. Martin.** 1982. Direct photoaffinity labeling of an allosteric site of subunit protein M1 of mouse ribonucleotide reductase by dATP. J. Biol. Chem. **257:**9505–9512.
4. **Carlson, J., J. A. Fuchs, and J. Messing.** 1984. Primary structure of the *Escherichia coli* ribonucleoside diphosphate reductase. Proc. Natl. Acad. Sci. USA **81:**4294–4297.
5. **Colston, M. J., and E. O. Davis.** 1994. Homologous recombination, DNA repair, and mycobacterial *recA* genes, p. 217–226. *In* Barry R. Bloom (ed.), Tuberculosis: pathogenesis, protection and control. American Society for Microbiology, Washington, D.C.
6. **Engstrom, Y., S. Eriksson, L. Thelander, and M. Akerman.** 1979. Ribonucleotide reductase from calf thymus. Purification and properties. Biochemistry **18:**2841–2948.
7. **Eriksson, S., and B.-M. Sjoberg.** 1989. Ribonucleotide reductase, p. 190–215. *In* G. Herve (ed.), Allosteric enzymes. CRC Press, Boca Raton, Fla.
8. **Mao, S. S., T. P. Holler, J. M. Bollinger, G. X. Yu, M. I. Johnston, and J. Stubbe.** 1992. Interaction of C225SR1 mutant subunit of ribonucleotide reductase with R2 and nucleoside diphosphates: tales of a suicidal enzyme. Biochemistry **31:**9744–9751.
9. **Mao, S. S., T. P. Holler, G. X. Yu, J. M. Bollinger, S. Booker, M. I. Johnston, and J. Stubbe.** 1992. A model for the role of multiple cysteine residues involved in ribonucleotide reduction: amazing and still confusing. Biochemistry **31:**9733–9743.
10. **Mao, S. S., M. I. Johnston, J. M. Bollinger, and J. Stubbe.** 1989. Mechanism-based inhibition of a mutant *Escherichia coli* ribonucleotide reductase (C225S) by its substrate CDP. Proc. Natl. Acad. Sci. USA **86:**1485–1489.
11. **Mao, S. S., G. X. Yu, D. Chalfoun, and J. Stubbe.** 1992. Characterization of C439S R1, a mutant of *E. coli* ribonucleotide diphosphate reductase: evidence that C439 is a residue essential for nucleotide reduction and C439SR1 is a protein possessing novel thioredoxin-like activity. Biochemistry **31:**9752–9759.
12. **Moore, E. C., D. Peterson, L. Y. Yang, C. Y. Yeung, and N. F. Neff.** 1974. Separation of ribonucleotides and deoxyribonucleotides on columns of borate covalently linked to cellulose. Application to the assay of ribonucleotide diphosphate reductase. Biochemistry **13:**2904–2907.
13. **Moss, N., R. Deziel, J. Adams, N. Aubry, M. Bailey, M. Baillet, P. Beaulieu, J. DiMaio, J.-S. Duceppe, J.-M. Ferland, J. Gauthier, E. Ghiro, S. Goulet, L. Grenier, P. Lavallee, C. Lepine-Frenette, R. Plante, S. Rakhit, F. Soucy, D. Wernic, and Y. Guidon.** 1993. Inhibition of herpes simplex virus type 1 ribonucleotide reductase by substituted tetrapeptide derivatives. J. Med. Chem. **36:**3005–3009.
14. **Reichard, P.** 1993. From RNA to DNA, why so many ribonucleotide reductases? Science **260:**1773–1777.
15. **Rubin, H., M. Plotnnick, Z.-M. Wang, X. Liu, Q. Zhong, N. M. Schechter, and B. S. Cooperman.** 1994. Conversion of α1-antichymotrypsin into a human neutrophil elastase inhibitor: demonstration of variants with different rate constants, stoichiometries of inhibition and complex stabilities. Biochemistry **33:**7627–7633.
16. **Rubin, H., J. S. Salem, L.-S. Li, F.-D. Yang, S. Mama, Z.-M. Wang, A. Fisher, C. S. Hamann, and B. S. Cooperman.** 1993. Cloning, sequence determination, and regulation of the ribonucleotide reductase subunits from *Plasmodium falciparum*: a target for antimalarial therapy. Proc. Natl. Acad. Sci. USA **90:**9280–9284.
17. **Rubin, H., Z. Wang, E. B. Nickbarg, S. McLarney, N. Naidoo, O. L. Schoenberger, J. Johnson, and B. S. Cooperman.** 1990. Cloning, expression and biological activity of native and site-directed altered human alpha 1-antichymotrypsin. J. Biol. Chem. **265:**1199–1207.
18. **Schaper, K.-J., J. K. Seydel, M. Rosenfeld, and J. Kazda.** 1986. Development of inhibitors of mycobacterial ribonucleotide reductase. Lepr. Rev. **57(Suppl. 3):**254–264.
19. **Small, P. R., P. C. Hopewell, S. P. Singh, A. Paz, J. Parsonnet, D. C. Ruston, G. F. Schecter, C. L. Daley, and G. K. Schoolnik.** 1994. The epidemiology of tuberculosis in San Francisco. N. Engl. J. Med. **330:**170–1709.
20. **Snider, D. E., M. Raviglione, and A. Kochi.** 1994. Global burden of tuberculosis, p. 3–12. *In* B. R. Bloom (ed.), Tuberculosis: pathogenesis, protection and control. American Society for Microbiology, Washington, D.C.
21. **Steeper, J. R., and C. D. Steuart.** 1970. A rapid assay for CDP reductase activity in mammalian cell extracts. Anal. Biochem. **34:** 123–130.
22. **Sun, L., B. A. Jacobson, B. S. Dien, F. Srienc, and J. A. Fuchs.** 1994. Cell cycle regulation of the *Escherichia coli nrd* operon: requirement for a *cis*-acting upstream AT-rich sequence. J. Bacteriol. **176:**2415–2416.
23. **Thelander, L.** 1973. Physicochemical characterization of ribonucleoside diphosphate reductase from *Escherichia coli*. J. Biol. Chem. **248:**4591–4601.
24. **Thelander, L., B. M. Sjoberg, and S. Eriksson.** 1978. Ribonucleoside diphosphate reductase (*Escherichia coli*). Methods Enzymol. **51:**227–237.
25. **Tuggle, C. K., and J. A. Fuchs.** 1986. Regulation of the operon encoding ribonucleotide reductase in *Escherichia coli*: evidence for both positive and negative control. EMBO J. **5:**1077–1085.
26. **Uhlin, U., and H. Eklund.** 1994. Structure of ribonucleotide reductase protein R1. Nature (London) **370:**533–539.
27. **Wheeler, P. R.** 1987. Biosynthesis and scavenging of purines by pathogenic mycobacteria including *Mycobacterium leprae*. J. Gen. Microbiol. **133:**2999–3011.
28. **Wheeler, P. R.** 1987. Enzymes for purine synthesis and scavenging in pathogenic mycobacteria and their distribution in *Mycobacterium leprae*. J. Gen. Microbiol. **133:**3013–3018.
29. **Wheeler, P. R.** 1990. Biosynthesis and scavenging of pyrimidines by pathogenic mycobacteria. J. Gen. Microbiol. **136:**189–201.
30. **Wheeler, P. R., and C. Ratledge.** 1994. Metabolism of *Mycobacterium tuberculosis*, p. 353–388. *In* B. R. Bloom (ed.), Tuberculosis: pathogenesis, protection and control. American Society for Microbiology, Washington, D.C.
31. **Winder, F. G., and D. S. Barber.** 1973. Effects of hydroxyurea, nalidixic acid and zinc limitation on DNA polymerase and ATP-dependent deoxyribonuclease activities of *Mycobacterium smegmatis*. J. Gen. Microbiol. **76:**189–196.
32. **Winder, F. G., and M. S. McNulty.** 1970. Increased DNA polymerase activity accompanying decreased DNA content in iron-deficient *Mycobacterium smegmatis*. Biochim. Biophys. Acta **209:** 578–586.