Biochemical and Structural Characterization of the Secreted Chorismate Mutase (Rv1885c) from *Mycobacterium tuberculosis* H₃₇R_v: an *AroQ Enzyme Not Regulated by the Aromatic Amino Acids

Sook-Kyung Kim,¹ Sathyavelu K. Reddy,^{1†} Bryant C. Nelson,² Gregory B. Vasquez,¹ Andrew Davis,¹ Andrew J. Howard,^{1‡} Sean Patterson,¹ Gary L. Gilliland,^{1§} Jane E. Ladner,¹ Prasad T. Reddy^{1*}

Biochemical Science Division¹ and Analytical Chemistry Division,² Chemical Science and Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899

Received 30 March 2006/Accepted 24 July 2006

The gene Rv1885c from the genome of *Mycobacterium tuberculosis* H₃₇R_v encodes a monofunctional and secreted chorismate mutase (*MtCM) with a 33-amino-acid cleavable signal sequence; hence, it belongs to the *AroQ class of chorismate mutases. Consistent with the heterologously expressed *MtCM having periplasmic destination in *Escherichia coli* and the absence of a discrete periplasmic compartment in *M. tuberculosis*, we show here that *MtCM secretes into the culture filtrate of *M. tuberculosis*. *MtCM functions as a homodimer and exhibits a dimeric state of the protein at a concentration as low as 5 nM. *MtCM exhibits simple Michaelis-Menten kinetics with a K_m of 0.5 ± 0.05 mM and a k_{cat} of 60 s^{-1} per active site (at 37°C and pH 7.5). The crystal structure of *MtCM has been determined at 1.7 Å resolution (Protein Data Bank identifier 2F6L). The protein has an all alpha-helical structure, and the active site is formed within a single chain without any contribution from the second chain in the dimer. Analysis of the structure shows a novel fold topology for the protein with a topologically rearranged helix containing Arg₁₃₄. We provide evidence by site-directed mutagenesis that the residues Arg₄₉, Lys₆₀, Arg₇₂, Thr₁₀₅, Glu₁₀₉, and Arg₁₃₄ constitute the catalytic site; the numbering of the residues includes the signal sequence. Our investigation on the effect of phenylalanine, tyrosine, and tryptophan on *MtCM shows that *MtCM is not regulated by the aromatic amino acids. Consistent with this observation, the X-ray structure of *MtCM does not have an allosteric regulatory site.

Chorismate occupies a central place in the shikimate pathway for the biosynthesis of the aromatic amino acids: phenylalanine, tyrosine, and tryptophan (16). Chorismate mutase (CM) (EC 5.4.99.5) catalyzes the conversion of chorismate to prephenate by Claisen rearrangement. CM accelerates the rearrangement by a factor of 10^6 compared to the uncatalyzed rearrangement at 25°C (1). Several biochemical and structural studies on CM have been performed due to the diverse nature of CMs (6, 24, 26, 29, 51, 58). Besides CM, the other enzymes that utilize chorismate are chorismate lyase, isochorismate synthase, anthranilate synthase, and *p*-aminobenzoate synthase (16). Subsequent to the conversion of chorismate to prephenate by CM, prephenate dehydratase and prephenate dehydrogenase catalyze the biosynthesis of phenylalanine and tyrosine, respectively. Anthranilate synthase and a battery of other enzymes catalyze the biosynthesis of tryptophan from chorismate. CMs exist in multiple forms both functionally and structurally. Some examples of the monofunctional CMs from *Bacillus subtilis* (13) and *Serratia rubidaea* (55) lack allosteric control and exhibit simple Michaelis-Menten kinetics. Other examples of the monofunctional CMs from *Saccharomyces cerevisiae* (49)

and *Aspergillus nidulans* (25) exhibit complex multi-end product effector control. Tryptophan is a potent allosteric activator, and phenylalanine and tyrosine are inhibitors of this class of CMs. On three occasions, CMs have evolved to become one catalytic component of a bifunctional protein: CM-prephenate dehydratase (8, 10), CM-prephenate dehydrogenase (8, 21), and CM-3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (20). The CMs from *Escherichia coli* (29) and *S. cerevisiae* (58) are all α -helical structures whereas the CMs from *B. subtilis* (6, 26) and *Thermus thermophilus* (17) have α/β barrel structures.

The shikimate pathway enzymes are essential for the biosynthesis of phenylalanine, tyrosine, tryptophan, and a number of other aromatic compounds in bacteria, fungi, algae, and plants, but this pathway is absent in mammals. One can take advantage of the nonoccurrence of CMs in humans and try to develop antimicrobial drugs to combat dreaded human pathogens such as *Mycobacterium tuberculosis*. With this intention, we gathered from the annotation of *M. tuberculosis* genome two putative genes for CM: Rv1885c and Rv0948c (5, 7). The Rv1885c gene product has about 25% identity with the monofunctional CM of *Erwinia herbicola* (56). The Rv0948c gene product has similarity to the amino terminus of some CM-prephenate dehydratases. We reported earlier that Rv1885cCM has neither prephenate dehydratase nor prephenate dehydrogenase activities (S. Patterson, S.-K. Kim, G. Vasquez, S. K. Reddy, and P. T. Reddy, Am. Soc. Microbiol. Biodefense Res. Meet., abstr. 173, 2004). The Rv1885c CM is synthesized with an amino terminal signal sequence which is cleaved off upon expression in *E. coli* with the mature protein

* Corresponding author. Mailing address: National Institute of Standards and Technology, Mail stop 831.2, Bldg. 227, Room B244, Gaithersburg, MD 20899. Phone: (301) 975-4871. Fax: (301) 975-5449. E-mail: prasad.reddy@nist.gov.

† Present address: Department of Zoology, Sri Venkateswara University, Tirupati, Andhra Pradesh, India.

‡ Present address: Illinois Institute of Technology, Chicago, IL 60616.

§ Present address: Centocor, Inc., 145 King of Prussia Rd., Radnor, PA 19087.

beginning with Asp₃₄. Two reports on the same theme of Rv1885c CM have appeared recently (42, 48). According to a terminology coined earlier for the secreted CMs of AroQ class (4) and Rv1885c CM (48), we will refer to this *M. tuberculosis* CM as *MtCM to infer its secretory property. In our study we have furthered the knowledge on the catalytic mechanism by crystal structure analysis and site-directed mutagenesis of the catalytic site residues. We have identified Arg₄₉, Lys₆₀, Arg₇₂, Thr₁₀₅, Glu₁₀₉, and Arg₁₃₄ in the mature protein (numbering of amino acids includes the signal sequence) as the critical residues for catalysis. The crystal structure analysis of *MtCM shows that the protein folds into an all α -helical bundle similar to the structures of *E. coli* CM (EcCM) (29) and *S. cerevisiae* CM (ScCM) (58). The catalytic site of *MtCM is formed within a single chain of the dimer with no contribution from the second chain, similar to the catalytic site of ScCM but different from the catalytic site of EcCM, which includes residues from both the chains. The coordinates for the structure of *MtCM were released first from our laboratory with the Protein Data Bank (PDB) identifier 2F6L. While this article was in preparation, we noticed a similar structural analysis of *MtCM by Ökvist et al. (40) (PDB 2FP1/2FP2) and later by Qamra et al. (43) (PDB 2AO2). We report here that the *MtCM is secreted from *M. tuberculosis* into the culture medium. Our search for *MtCM homologues in other mycobacterial species revealed its existence in pathogenic species such as *Mycobacterium bovis* and *Mycobacterium avium*. Perhaps, *MtCM may have some implications in virulence of the human pathogen *M. tuberculosis*. Sasso et al. (48) discussed the available evidence for the idea that *MtCM may be involved in pathogenesis.

MATERIALS AND METHODS

Materials. Restriction endonucleases, a Quick ligase kit, T4 polynucleotide kinase, and calf intestinal alkaline phosphatase were purchased from New England Biolabs, Beverly, MA. SeaKem GTG Agarose and NuSieve GTG Agarose were purchased from FMC BioProducts, Rockland, ME. The PCR product purification kit, gel extraction kit, and plasmid purification kit were from QIAGEN, Valencia, CA. DEAE cellulose (DE52) was from Whatman laboratories. A Big Dye dideoxy terminator cycle sequencing kit was purchased from Applied Biosystems, Inc. (Perkin Elmer Cetus, Foster City, CA). A QuikChange II XL site-directed mutagenesis kit was purchased from Stratagene, La Jolla, CA. L-Malic acid and NAD were obtained from Sigma Biochemicals, St. Louis, MO. Middlebrook 7H9 broth base was obtained from Difco laboratories, Detroit, MI. Oligonucleotides were obtained from Operon, QIAGEN, Valencia, CA. Crystal Screen Kit I was purchased from Hampton Research, Laguna Niguel, CA.

***E. coli* strains and plasmids.** The expression vector pRE1 for native protein production based on the λ P_L promoter was described previously (45). The *E. coli* strains for λ P_L-based expression are C600 λ lysogen (r^- m⁺ λ cI⁺) and MZ1 (λ c1857s [59]). C600 λ lysogen and MZ1 were kindly provided by Donald Court of the National Cancer Institute, Frederick, MD. The *E. coli* strains NovaBlue and BL21(DE3) and the expression vector pET15b, for the N-terminal histidine-tagged fusion expression based on isopropyl- β -D-thiogalactopyranoside (IPTG) induction, were obtained from Novagen, Madison, WI. NovaBlue {*endA1 hsdR17*(r_{K12}^- m_{K12}⁺) *supE44 thi-1 recA1*gyrA96 *relA1 lacF*[*proA*⁺*B*⁺*lacI*^q*Z* Δ M15::Tn10] (Tet^r)} was used for cloning, and BL21(DE3) [*F*⁻ ompT *hsdS_B*(r_B^- m_B⁻) *gal dcm* (DE3)] was used for expression of the cloned gene. The DE3 designation indicates that the host is a lysogen of λ DE3 that carries a chromosomal copy of the T7 RNA polymerase gene under control of the *lacUV5* promoter.

DNA procedures. Ten milliliters of Luria-Bertani (LB) medium, supplemented with 100 μ g of ampicillin/ml, was inoculated with a colony of C600 (λ cI⁺) or NovaBlue harboring a plasmid. Cells were grown overnight at 37°C. Plasmids were isolated by using a QIAGEN mini plasmid purification kit. Digestion of DNA with restriction enzymes was performed according to the manufacturer's

recommendation. DNA fragments were separated by electrophoresis on SeaKem GTG agarose or NuSieve GTG agarose and used for ligation. Ligation of DNA fragments was performed using a Quick ligase kit. Competent cells of the *E. coli* strains used here were prepared by the Hanahan method (15).

Cloning of Rv1885c and production of *MtCM in *E. coli*. The full-length Rv1885c gene coding sequence was amplified by PCR using the forward primer 5'-GGAATTCATATGCTTACCGTCCACGTGAG (with NdeI restriction recognition sequence in bold letters) and the reverse primer 5'-CGCGGATCC TCAGGCCGCGGTAGGGC (with BamHI restriction recognition sequence in bold letters). Amplification conditions were as follows: 96°C for 60 s for initial melting of DNA, followed by 25 cycles of amplification, with each cycle consisting of melting at 96°C for 45 s, annealing at 50°C for 45 s, and polymerization at 72°C for 60 s. Polymerization was continued at the end for 10 min at 72°C. A total of 200 ng of *M. tuberculosis* H₃₇R_v chromosomal DNA (kindly provided by John Belisle and Patrick Brennan, Colorado State University) and 100 ng of primers were used in the amplification. The Rv1885c gene beginning at codon 34 was similarly amplified using the forward primer 5'-GGAATTCATATGGACGGC ACCAGCCAGTTAGCC and the same reverse primer (restriction recognition sequence in bold letters). The amplified DNA was digested with NdeI and BamHI and cloned into the respective sites of the plasmid pRE1 (45). A recombinant was isolated from C600 λ lysogen and introduced into the strain MZ1 for native protein production (45). Similarly, the NdeI and BamHI fragments were also cloned into the respective sites of the expression vector pET15b, and a recombinant was isolated from *E. coli* NovaBlue. The pET15b recombinant was introduced into *E. coli* BL21(DE3) for *MtCM production based on the IPTG induction (52).

For protein production trials using the pRE1 recombinant (45), *E. coli* MZ1 harboring the recombinant plasmid pRE1 was grown in 50 ml of LB medium, supplemented with ampicillin (50 μ g/ml), at 32°C in a water bath shaker to an A₆₀₀ of $\approx$ 0.4. To induce *MtCM synthesis, 50 ml of LB medium kept at 65°C was added to the culture and immediately shifted to a 42°C shaking water bath. The culture was further supplemented with ampicillin. A time course of induction was carried out by harvesting 25 ml of induced cells at hourly intervals for up to 4 h. We observed that 2 h of induction gave optimal production of the mature protein from the full-length clone. The protein was produced to $\approx$ 5% of the total cellular protein and was completely soluble. We would like to point out that growth of the cells after induction was hampered by the production of the Rv1885c gene product, suggesting the toxic nature of the gene product. For protein production using the pET15b recombinant, *E. coli* BL21(DE3) harboring the recombinant plasmid was grown in LB medium at 37°C to an A₆₀₀ of $\approx$ 0.6. *MtCM synthesis was induced with 30 μ M IPTG at 25°C overnight.

Purification of *MtCM from cellular extract for crystallization. For purification of the protein produced from the full-length clone, the MZ1/pRE1 recombinant culture was scaled up to 4 liters of induced cells. Cells were harvested by centrifugation and washed with 25 mM Tris-HCl buffer, pH 7.5. Cells (5 g of wet weight) were suspended in 50 ml of 50 mM Tris-HCl, pH 7.5, 1 mM dithiothreitol [DTT], 1 mM EDTA, and 1 mM phenylmethylsulfonyl fluoride (lysis buffer). Cell extract was prepared by passing the cell suspension through a French press twice and by centrifugation at 100,000 $\times$ g for 1 h. The supernatant was loaded onto a 40-ml DE52 column and eluted with a linear NaCl gradient generated from 200 ml of the lysis buffer and 200 ml of the lysis buffer containing 0.5 M NaCl. Fractions corresponding to the overproduced protein, as judged by the Coomassie stain on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels, were concentrated and further purified on a 480-ml Sephadex-G75 superfine column (2.6 by 100 cm) equilibrated with the lysis buffer supplemented with 100 mM NaCl. Again, the purest protein fractions corresponding to the overproduced protein were pooled and concentrated. The final yield of the protein originating from the full-length clone was 10 mg, i.e., 2.5 mg/liter culture with a purity of $\approx$ 95%. For the purpose of crystallization, the protein was further purified on a 5-ml hydroxylapatite column. The protein was dialyzed against 20 mM potassium phosphate buffer, pH 7.4, containing 1 mM DTT. The protein was loaded and eluted with a phosphate buffer gradient generated from 50 ml each of 20 mM and 300 mM phosphate buffer, pH 7.4. Pure protein was dialyzed against 50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 mM DTT, and 100 mM sodium chloride and concentrated to $\approx$ 10 mg/ml.

Isolation of the periplasmic proteins. Periplasmic proteins were isolated by the method of Neu and Heppel (38) with the slight modification described in the QIAGEN handbook. Induced cells from a 100-ml culture of BL21(DE3)/pET15b recombinant (300 mg of cells) were suspended in 30 ml of 20% sucrose–30 mM Tris-HCl, pH 8.0, at 24°C, and EDTA was added to a final concentration of 1 mM. The suspension was mixed in a rotary shaker at 150 rpm for 10 min at 24°C and centrifuged at 10,000 $\times$ g for 15 min at 4°C. The cell pellet was suspended in 30 ml of ice-cold 5 mM MgSO₄ and kept on a rotary shaker at 150 rpm for 10

min at 4°C. The osmotic-shocked cell suspension was centrifuged at 10,000 × *g* for 20 min at 4°C, and the supernatant containing the periplasmic proteins was removed and buffered with 1 M Tris-HCl, pH 8.0, to a 30 mM final concentration. The osmotic-shocked cell pellet was suspended in 30 ml of 30 mM Tris-HCl, pH 8.0 and passed through a French press. This represents cytoplasmic proteins. Total cellular extract and the cytoplasmic membranes were prepared from unshocked cells.

Small-scale purification of *MtCM from the periplasmic proteins. Periplasmic proteins were isolated as described above from a 100-ml culture of *E. coli* BL21(DE3) harboring the pET15b recombinant. The periplasmic fluid was concentrated to ~200 µl in a Millipore centrifugal tube with a 5,000-molecular-weight cutoff and chromatographed on a 210-ml Biosep SEC-3000 high-pressure liquid chromatography (HPLC) column (Phenomenex, Torrance, CA). The column was equilibrated and eluted with 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT, and 100 mM NaCl. The purity of *MtCM isolated from the periplasm by a single chromatography step equaled that observed for the protein purified from the cellular extract as described above by a three-step procedure. Therefore, we purified all the mutant variants of *MtCM from the periplasm.

Growth of *M. tuberculosis*, isolation of culture filtrate, and preparation of cell extract. A clinical isolate of *M. tuberculosis* H₃₇R_v and a laboratory strain of *M. tuberculosis* H₃₇R_v were grown in Middlebrook 7H9 medium supplemented with glycerol, albumin-dextrose complex, and Tween 80 (23) in a roller bottle in a biosafety level 3 facility. After 1 week (early in the exponential phase of growth) and 3 weeks (late in the exponential phase of growth) at 37°C, cell cultures were centrifuged at 5,000 × *g*. The culture supernatant (filtrate) was filtered through a Millipore membrane filter with 0.2-µm-pore-size cutoff. The filter has low affinity for proteins. The culture filtrate was dialyzed overnight against 50 mM Tris-HCl, pH 7.5, to remove the medium constituents except albumin although we found that the growth medium has no effect on CM activity of the purified protein.

Cell extract of the laboratory strain of *M. tuberculosis* H₃₇R_v was prepared from a 100-ml culture in the late exponential phase of growth. Cells were collected by centrifugation and washed twice with lysis buffer (50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 100 mM NaCl). The cell pellet was suspended in 0.5 ml of lysis buffer in a microcentrifuge tube, and glass beads (100 µm size) equivalent to a 0.5-ml volume were added. Cell extract was prepared in a bead beater by subjecting the cell-bead suspension to four 45-s pulses by intermittent cooling. Unbroken cells were centrifuged off at 14,000 rpm for 5 min in a microcentrifuge. The cell extract was filtered through a membrane filter (0.2-µm-pore-size cutoff) for safe handling. One hundred microliters of the cell extract was diluted 80-fold (cell disruption under these conditions is >80%) with the lysis buffer in order to normalize the activity per unit volume of culture filtrate/cell extract. Separately, another 100 µl of the cell extract was diluted with the growth medium to test the effect of growth medium on CM and malate dehydrogenase (MDH) activities. MDH, a citric acid cycle cytoplasmic enzyme (44), was assayed as a control for cell lysis. Various volumes of the dialyzed culture filtrate and the diluted cell extract were assayed for CM and MDH.

MDH assay. The enzyme was assayed in a 1-ml volume containing 60 mM glycine buffer, pH 9.0, 1 mM NAD⁺, and various volumes of the dialyzed culture supernatant. The reaction was started with 20 mM L-malate (sodium salt), and the reduction of NAD⁺ was followed at 340 nm (12). The reaction was carried out at 24°C.

Site-directed mutagenesis. Mutagenesis of Arg₄₉Ala, Lys₆₀Ala, Asp₆₉Ala, Arg₇₂Ala, Thr₁₀₅Ala, Glu₁₀₉Ala, Glu₁₀₉Gln, and Arg₁₃₄Ala was accomplished using a QuikChange II XL Site-Directed Mutagenesis Kit. The numbering of the amino acid residues includes the amino terminal signal sequence. The sequence of the oligonucleotide pairs for each mutagenesis reaction is as follows: for Arg₄₉Ala, 5'-GCCGCCGCTGAGGCGTTGGAGGTCGCC and 5'-GGCGACC TCCAACGCCTCAGCGCGGCC; for Lys₆₀Ala, 5'-GGTGGCAGCCTTCGCG TGGCGTGCTCAGC and 5'-GCTGAGCACGCCACGCGAAGGCTGCCACC; for Asp₆₉Ala, 5'-CTGCCCATTTAGGCTTCGGGCCGAGTC and 5'-GACTC GGCCGGAAGCCTCAATGGGCGAG; for Arg₇₂Ala, 5'-CATTGAGGATCC GCGCAGTCGAACAGCAACTCG and 5'-CGAGTTGCTGTTCGACTGCG CCGGAATCCTCAATG; for Thr₁₀₅Ala, 5'-CAGATTGCGCCGCCGAGGCA ATCGAG and 5'-CTCGATTGCCTCGCGCGCGCAATCTG; for Glu₁₀₉Ala, 5'-CACCGAGGCAATCGGTACAGCCGGTTC and 5'-GAACCGGCTGTAC CCGATTGCTCGGTG; for Glu₁₀₉Gln, 5'-CACCGAGGCAATCAGTACAGC GCGTTTC and 5'-GAACCGGCTGTACTGGATTGCTCGGTG; for Arg₁₃₄Ala, 5'-CTATCGGCATCGGCATCGGCGATCGAA and 5'-GTCTGATCGCGGATGC CGATGCCGATAG. Nucleotide bases in bold represent the mutation. Oligonucleotides were purified by HPLC and phosphorylated with polynucleotide kinase (46). The mutagenesis efficiency was ~90%. The template plasmid DNA for mutagenesis

was the pET15b/Rv1885c recombinant. For each variant, the Rv1885c coding region was verified by DNA sequencing using primers T7P (5'-TAATACGACTCACTAT AGGG) and T7T (5'-GCTAGTTATTGCTCAGCGG).

CM assays and kinetic studies. CM was assayed as previously described (9). Since the commercial preparation of chorismate is rather impure (~70%), we used pure chorismate (>99% pure; generous gift from Marcia Holden, Biochemical Science Division, National Institute of Standards and Technology) isolated from a CM auxotroph of *E. coli* which overproduces chorismate and exports it into the culture medium (14). Reaction volumes of 0.4 ml of chorismate (typically, 1 mM) in 50 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 0.1 mg/ml bovine serum albumin, and 10 mM β-mercaptoethanol were incubated at 37°C for 5 min. The reaction was started with 5 pmol of *MtCM in a 10-µl volume (i.e., 185 ng of *MtCM equivalent to a 12.5 nM final concentration of the dimer based on the molecular mass of 36,948 Da). Any variation in the amount of *MtCM used in the assay is specified in the figure legends. The reaction was allowed to proceed at 37°C and was terminated after 1 min to 5 min (depending on the mutation in the protein used in the assay) with 0.4 ml of 1 N HCl. After a further incubation at 37°C for 10 min to convert prephenate, which is formed in the enzymatic reaction, to phenylpyruvate, 0.8 ml of 2.5 N NaOH was added. The absorbance of the phenylpyruvate chromophore was read at 320 nm. We set up a blank with no enzyme for every reaction to account for the nonenzymatic conversion of chorismate to prephenate and added the enzyme after NaOH addition. The A₃₂₀ for the blank varied from 0.1 to 0.3, depending on the concentration of chorismate and duration of the reaction. The wild-type and the Glu₁₀₉Gln mutant enzymes were assayed over a broad pH range using a universal buffer (33).

Molecular mass determination of *MtCM. We determined the molecular mass of *MtCM by liquid chromatography-mass spectrometry and of the mutant forms via matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry. Mass spectra were collected and analyzed using an Applied Biosystems Voyager-DE STR Biospectrometry Workstation (Foster City, CA).

Crystallization. Crystallization conditions were surveyed by the hanging drop vapor diffusion method using a Hampton Research Crystal Screen Kit I. *MtCM was purified from the pRE1/Rv1885c recombinant and adjusted to a concentration of 8.3 mg/ml in 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT, and 100 mM sodium chloride. The refined conditions used a well solution of 0.1 M morpholineethanesulfonic acid, pH 6.5, and 30% polyethylene glycol 8000. The crystals grew as thin plates and grew best between 25°C and 30°C. Seeding was used to reduce the number of crystals in each drop, and the polyethylene glycol 8000 was lowered to 14% to 18%.

Data collection. Diffraction data were collected at Industrial Macromolecular Crystallography Association Collaborative Access Team beamline 17ID-B at the Advanced Photon Source of Argonne National Laboratory (Argonne, IL) using a Mar Research charge-coupled-device detector. The wavelength for the data collection was 1 Å. For data collection at 100 K, the crystals were dipped in a solution made by mixing the well solution and 50% polyethylene glycol in a 1:1 ratio. The data were processed using X-GEN (19), and the statistics are shown in Table 1.

Structure determination. The structure of *MtCM was solved by molecular replacement using PHASER (35) with the partially refined structure of CM from *Yersinia pestis* (unpublished data). The asymmetric unit includes a dimer of *MtCM. REFMAC5 (37) was used to refine the model, and RESOLVE (54) was used to iteratively rebuild the model to remove bias. The refining and rebuilding were repeated. The final refinement statistics are shown in Table 1. XtalView (36) was used to view the model graphically and to build portions not built by RESOLVE. The stereochemistry was checked with PROCHECK (28) and with MolProbity (31). The Ramachandran plot in PROCHECK has 96.0% in the most favored region and 4% in the additional allowed region. The clash score from MolProbity for all atoms is 4.3, and for atoms with B values less than 40, the clash score is 2.56. Eleven residues in chain A and 11 residues in chain B of the dimeric *MtCM were modeled with two alternative conformations. No interpretable density was found for the first two residues, 34 and 35, of chain B, and the corresponding residues in chain A are not well ordered either (the numbering of the residues includes the 33-amino-acid signal sequence which is cleaved off in the mature protein).

Other methods. Protein concentration was determined according to the method of Lowry et al. using bovine serum albumin as the calibration standard (32) with the appropriate buffer in which the protein was purified as the blank. DNA sequence was determined by the dideoxy sequencing method (47) as adapted for the Applied Biosystems model 3100 Genetic Analyzer.

TABLE 1. Data Collection and refinement statistics

Parameter ^a	Value for *MtCM
Diffraction data	
Space group	P2 ₁
Cell dimensions	
<i>a</i> (Å)	42.89
<i>b</i> (Å)	72.93
<i>c</i> (Å)	62.06
α (°)	90.0
β (°)	103.85
γ (°)	90.0
Resolution limit (Å)	1.7
No. of measured intensities	296,227
No. of unique reflections	40,279
Mean redundancy	7.4
<i>R</i> _{merge} (overall/high-resolution shell)	0.067/0.381
% Completeness (overall/high-resolution shell)	98.6/97.7
Avg <i>I</i> / σ (overall/high-resolution shell)	15.3/2.3
Refinement	
Resolution limits (Å)	20–1.7
<i>R</i> factor (95% of the data)	0.182
<i>R</i> _{free} (5% of the data)	0.222
No. of water molecules	363
Bond length RMS deviation (Å)	0.019
Bond angle RMS deviation (°)	1.84
Avg <i>B</i> (main chain/side chain) (Å ²)	18.7/21.4
Avg <i>B</i> for water (Å ²)	29.3

^a RMS, root mean square.

RESULTS AND DISCUSSION

***MtCM is secreted in *M. tuberculosis* H₃₇R_v.** It is established that the Rv1885c gene product is synthesized with a signal peptide, which is cleaved off upon heterologous expression in *E. coli*, and transported into the periplasmic space (42, 48; Patterson et al., Am. Soc. Microbiol. Biodefense Res. Meet., abstr. 173, 2004) (Fig. 1, lane 4). One difference in the cellular architecture of the gram-negative *E. coli* and the gram-positive *M. tuberculosis* is that there is no discrete periplasmic space in *M. tuberculosis*, although a pseudoperiplasmic space might exist (39). Therefore, we asked the question, where is the mature *MtCM destined in its native environment? In order to probe for the cellular location of *MtCM, we tested the *M. tuberculosis* culture filtrate for CM activity. As shown in Fig. 2A, a reliable volume-dependent linear activity for CM was observed in the culture filtrate of a laboratory strain of *M. tuberculosis* (a superimposable result was observed for the culture filtrate from a clinical isolate). The CM activity in the culture filtrate from the cells in the late exponential phase of growth (3-week-old culture) was approximately twofold higher than that observed with the culture filtrate from the cells in the early exponential phase of growth (1-week-old culture). We wondered if the detection of CM in the culture filtrate was due to lysis of mycobacterial cells. If cell lysis were to occur, then the activity for CM in the culture filtrate cannot be attributed solely to *MtCM because a second gene (Rv0948c) for a cytoplasmic CM exists in *M. tuberculosis* (5, 7, 42, 48). We therefore performed two control experiments using the diluted cell extract prepared as described in Materials and Methods. First, we assayed the diluted cell extract for CM. The results in Fig. 2A clearly show that CM activity in the cell extract from late exponential phase cells is $\approx$ 10-fold lower than that observed in

the culture filtrate from the same batch of cells. Since CM activity in the cell extract is substantially lower than that present in the culture filtrate, we can attribute the CM activity in the culture filtrate as arising due to *MtCM but not due to the cytoplasmic Rv0948c CM. It has been reported (42) and we also observed (unpublished results) that the catalytic rate constant of Rv0948c CM is $\approx$ 1/50 that of *MtCM. Admittedly, the interpretation of the results would have been difficult if we had found equal or higher activity for CM in the cell extract than in the culture filtrate. Further, as a second control for cell lysis, a cytoplasmic Krebs cycle enzyme (44), MDH, was assayed in the culture filtrate as well as in the diluted cell extract by following the reduction of NAD⁺ to NADH by a sensitive spectrophotometric assay. MDH was detected only in the cell extract but not in the culture filtrate (Fig. 2B). This control experiment clearly shows that *M. tuberculosis* cells are not lysed under the growth conditions described. The results presented in this section demonstrate for the first time that *MtCM is secreted out of the cytoplasm and through the unusual architecture of the mycobacterial cell wall (3) into the culture medium. Secreted CMs have been shown to occur in the periplasm of *Salmonella enterica* serovar Typhimurium (4), *Pseudomonas aeruginosa* (4), and *E. herbicola* (56). Indeed, a periplasmic biosynthetic pathway for phenylalanine was demonstrated in *P. aeruginosa* (4). While it is intriguing to speculate about such a periplasmic pathway in *M. tuberculosis*, the lack of a discrete periplasmic compartment in this organism and the secretion of *MtCM from the cells into the culture medium preclude such a pathway. The biological significance of this highly active and secreted CM is discussed later.

Determination of the quaternary structure of *MtCM.

*MtCM protein was purified to near homogeneity as described in Materials and Methods (Fig. 3, lane 4). Purification of the

FIG. 1. SDS-PAGE (13%) analysis of the production and demonstration of the periplasmic location of *MtCM. Lane 1, molecular mass markers; lane 2, total extract of the induced cells (30 µg) of *E. coli* BL21(DE3) harboring the full-length clone (amino acids 1 to 199) in the expression vector pET15b; lane 3, extract of the induced cells after removal of the periplasmic fluid (25 µg); lane 4, periplasmic fluid (5 µg). Relevant molecular mass markers are indicated in kilodaltons on the left. The *MtCM monomer is shown with an arrow.

FIG. 2. (A) Demonstration of CM activity in the culture filtrate and cell extract of *M. tuberculosis* H₃₇R_v. The culture filtrate and cell extract were assayed for CM activity. The activity was determined at 37°C for 5 min in the presence of 2 mM chorismate. A 0.1-ml solution of 8 mM chorismate was used in a final reaction volume of 0.4 ml. The A₃₂₀ value, after subtraction of the absorbance observed with the control, is recoded on the y axis. The effect of growth medium on purified *MtCM activity was tested and found to have no effect on either the nonenzymatic or enzymatic conversion of chorismate to prephenate. Early exponential phase cells are 1-week-old cells; late exponential phase cells are 3-week-old cells. ■, culture filtrate from early exponential phase cells; ▲, culture filtrate from late exponential phase cells; ●, cell extract from late exponential phase cells. CM activity levels in the culture filtrate from late exponential phase cells of a clinical isolate and a laboratory strain of *M. tuberculosis* were identical. (B) Assay for MDH in the culture filtrate and cell extract of *M. tuberculosis* H₃₇R_v. MDH was assayed as described in Materials and Methods. ■, culture filtrate from late exponential phase cells; ▲, cell extract from late exponential phase cells.

protein by molecular sieve chromatography on a Sephadex G-75 superfine column showed an elution profile corresponding to a molecular mass of ca. 36,000 Da. Based on the monomeric molecular mass of 18,474 Da, the number of subunits in the isolated protein is two, clearly showing a dimeric form of the protein (molecular mass calibration figure is not shown). All the naturally occurring CMs thus far identified function either as dimers or trimers. A genetically engineered CM from *E. coli* functions as a dimer (29), whereas a *B. subtilis* CM functions as a trimer (6, 26). Although a naturally occurring monomeric CM has not been discovered yet, a topological redesign of *Methanococcus jannaschii* dimeric CM yielded a highly active monomer (34). Sasso et al. (48) systematically

FIG. 3. SDS-PAGE (13%) analysis of the production and purification of *MtCM. Lane 1, extract of the induced cells (30 μg) of *E. coli* MZ1 harboring the control plasmid pRE1; lane 2, extract of the induced cells (30 μg) harboring the full-length clone (amino acids 1 to 199) in pRE1; lane 3, extract of the induced cells (30 μg) harboring the truncated clone (amino acids 34 to 199) in pRE1 (note that the mature protein is not expressed in this *E. coli* background whereas Sasso et al. [48] have successfully expressed the mature protein in a thioredoxin reductase-deficient strain of *E. coli*); lane 4, purified *MtCM (10 μg); lane 5, molecular mass markers. The protein band corresponding to the *MtCM monomer is shown with an arrow. Relevant molecular mass markers are indicated in kilodaltons on the right.

investigated the quaternary structure of *MtCM by gel filtration chromatography at various protein concentrations ranging from 61 μM down to 1 μM and reported a dimeric nature of the protein in this concentration range. We performed size exclusion chromatography on a 480-ml Sephadex G-75 superfine column using 18.5 μg of protein in 1 ml (500 nM concentration based on the molecular mass of 36,948 Da) and 2.5 μg of protein in 1 ml (67.6 nM). Based on the activity analysis for CM, the protein was eluted with a peak activity at an identical elution volume corresponding to its dimeric molecular mass. Further, as the protein was eluted with a dilution factor of about 15-fold from the column, *MtCM exists in the dimeric form at a concentration as low as ≈5 nM. This observation suggests a tight association of the monomers despite a small interface between the monomers in the crystal structure of the dimer.

Kinetic properties of *MtCM. Enzymatic assays were performed with the protein purified from the periplasm to see how efficiently *MtCM catalyzes the conversion of chorismate to prephenate. We determined the K_m for chorismate at a protein concentration of 27.5 nM as well as at 8 nM (data not shown). *MtCM is active as chorismate mutase with a k_{cat} of $60 \pm 4 \text{ s}^{-1}$ per active site at 37°C and pH 7.5. *MtCM exhibited a simple hyperbolic response to chorismate at concentrations in the range of 0.1 mM to 5 mM, characteristic of Michaelis-Menten kinetics. The K_m for chorismate was calculated from Lineweaver-Burk plots (data not shown). We observed a small but reproducible variation in the K_m depending on the protein concentration used. A lower K_m of $0.50 \pm 0.05 \text{ mM}$ was obtained with a 27.5 nM protein concentration (11 pmol) whereas a K_m of $0.67 \pm 0.05 \text{ mM}$ was obtained with an 8 nM protein concentration (3.2 pmol).

Sasso et al. (48) and Prakash et al. (42) reported a K_m of 0.18 mM and 1.2 mM, respectively, with a sixfold difference. These differences in the K_m among the three reports may be reconciled by the manner in which the enzyme was purified, the temperature of the reaction assay, and whether the assay was performed as a “stop assay” by measuring phenylpyruvate (converted from prephenate) at 320 nm (42) or as an “online assay” by following the disappearance of chorismate at 274 nm (48), and, to an extent, how much protein was used in the assay as we point out here. We purified *MtCM as an untagged native protein and determined the K_m by the stop assay at 37°C. Sasso et al. (48) used a heat denaturation step to remove most of the *E. coli* proteins, followed by ion exchange chromatography, and the assay was performed online at 30°C. Prakash et al. (42) purified the protein with a carboxyl terminal His₆ tag and performed the stop assay at 30°C. These differences among the three protocols for purification of the protein, the enzyme concentration used in the assay, and the type and conditions of the assay may explain the observed variations for the K_m . While the pH 7.5 of the assay appears to be constant among the three protocols, a slight variation in the pH might have also contributed to the differences in the K_m in view of a detailed study by Sasso et al. (48), who have shown a 20-fold variation in the K_m values as a function of pH values between 5 and 8. This can be rationalized by the existence of a glutamate residue (Glu₁₀₉) that needs to be protonated for optimum activity, as we have studied in detail and described later.

***MtCM is not regulated by the aromatic amino acids.** In order to have a stringent control on the biosynthesis of tyrosine, phenylalanine, and tryptophan and their intracellular pool, some CMs have evolved to have feedback regulation of CM activity by these aromatic amino acids. Controversy exists in the recent literature with respect to the effect of the aromatic amino acids on *MtCM. The first report by Sasso et al. (48) finds no effect on the enzyme activity by the aromatic amino acids. A subsequent report by Prakash et al. (42) shows a complex regulation pattern of *MtCM by the aromatic amino acids, with activation of the enzyme at 100 μ M concentration and inhibition at 500 μ M and 800 μ M concentrations of the effectors. Prakash et al. (42) suggested that the binding of these ligands causes a conformational change in the protein that leads to the inaccessibility of the active site. To substantiate this suggestion, they performed limited proteolysis of *MtCM by trypsin in the presence of the ligands. At the concentrations of the ligands (100 μ M to 800 μ M) where they have observed activation/inhibition of the enzyme activity, there is no protection of *MtCM by trypsin. They observed protection of the enzyme from trypsin digestion only at 3,000 μ M and 5,000 μ M concentrations of the ligands, concentrations too high to be physiologically relevant.

In order to resolve this issue and classify *MtCM with regard to its allosteric regulation, we tested the effect of phenylalanine, tyrosine, and tryptophan on *MtCM activity under various conditions. We performed a number of experiments using 3.2 pmol of the protein with 2 mM chorismate as well as using 5 pmol and 11 pmol of the protein with 1 mM chorismate. We observed no effect on the activity by these aromatic amino acids other than an experimental variation of $\pm 5\%$ (a representative set of data is shown in Fig. 4). Our detailed investigation on the effect of the aromatic amino acids on *MtCM

FIG. 4. Effect of aromatic amino acids on the activity of *MtCM. Phenylalanine, tyrosine, and tryptophan were preincubated at the indicated concentrations with 3.2 pmol of *MtCM at 37°C for 5 min in 0.21 ml. The reaction was started with 0.2 ml of 4 mM chorismate, which was preincubated at 37°C for 5 min, to a final concentration of 2 mM. The reaction was allowed to proceed for 5 min. The final concentration of the Tris-HCl buffer was 50 mM, pH 7.5. An appropriate blank without the enzyme was included for every reaction, and the A_{320} value was subtracted from the A_{320} value with the enzyme. Tyrosine was used only up to a 1 mM final concentration since the solubility of this amino acid is poor at neutral pH. The catalytic rate constant, k_{cat} , of *MtCM without any aromatic amino acid present in the assay is 62 s⁻¹.

clearly shows no feedback regulation of the enzyme, consistent with the results of Sasso et al. (48) and in disagreement with the results of Prakash et al. (42). Our data generated from the stop assay is in agreement with the results of Sasso et al. (48), who performed the online assay. Prakash et al. (42) performed the stop assay, and so did we. Therefore, the type of the assay, the amount of protein used in the assay, and the temperature of the reaction for determining the effect of aromatic amino acids on *MtCM activity are not the obvious reasons for this discrepancy. However, we found a couple of differences in the composition of the reaction medium described by Prakash et al. (42). The investigators used 10 mM Tris-HCl, pH 7.5, in their assay, whereas we used 50 mM Tris-HCl, pH 7.5, in our assay, as recommended by the procedure described by Davidson and Hudson (9). The low 10 mM concentration of the buffer may not be sufficient to keep the pH constant upon the addition of chorismic acid and the aromatic amino acids. We carried out extensive measurements on the activity of the wild type and the Glu₁₀₉Gln mutant *MtCM as a function of pH as shown later (see Fig. 8). Our results show how strongly the activity of the wild-type enzyme can be inhibited at alkaline pH. Sasso et al. (48) have also documented how dramatically the K_m changes with pH, with a consequential effect on the catalytic efficiency, k_{cat}/K_m . We believe that experimental artifacts with the rather sensitive chorismate mutase assay are the reasons for the observed complex regulation of *MtCM by the aromatic amino acids as reported by Prakash et al. (42). Consistent with our observation that *MtCM is not regulated by the aromatic amino acids, the X-ray structure of *MtCM does not have an allosteric regulatory site in the protein (see below). Lambert et al. (27), while working with the *AroQ homolog from the phytoparasitic nematode *Meloidogyne javanica*, also did not find any regulation by the aromatic amino acids. It

FIG. 5. Comparison of the structures of *MtCM, EcCM, and ScCM. In the far left diagram, the monomer of *MtCM is shown with the N-terminal part of the chain in blue and the C-terminal part of the same chain in salmon. In the next diagram, the EcCM dimer is shown with one monomer blue and the other monomer salmon. In the third diagram (from left), the monomer of ScCM is shown with the N-terminal part of the chain in blue and the C-terminal part of the same chain in salmon. Tryptophan (green) is shown in the regulatory site of ScCM; this is also the region of dimer contact in ScCM. In the far right diagram, the dimer of *MtCM is shown with one monomer colored according to the scheme used for the first diagram; the second monomer is shown in a lighter blue and pink. The main residues of the active site are shown as stick atomic models in green. The residues involved in the dimer interface are listed in Table 2; they are shown here as stick models in yellow.

therefore appears that *AroQ CM regulated by the aromatic amino acids is yet to be discovered.

Overall structure of *MtCM and comparison with other CM structures. CMs have evolved with different structural folds for catalyzing the pericyclic rearrangement of chorismate to prephenate. The genetically engineered 109-amino-acid EcCM (29) (AroQ_α subclass) and the 256-amino-acid ScCM (58) (AroQ_β subclass) function as dimers and are α-helical bundles. *MtCM belongs to the secreted subclass of CMs (*AroQ [4]), which was very recently renamed as the γ subclass of AroQ (AroQ_γ) CMs (40). The 127-residue *B. subtilis* CM (BsCM) (6, 26) (AroH class) and the 122-residue *T. thermophilus* CM (TtCM) (17) (AroH class) function as trimers with α-helices and β-sheets. Despite the differences in the structural folds among the CMs, these enzymes display remarkable similarity in their active sites. The ScCM (49) and *A. nidulans* CM (AnCM) (25) are a step further apart in that these enzymes display allosteric regulation by tryptophan and tyrosine whereas EcCM and BsCM are not regulated by the aromatic amino acids. However, the AroH class TtCM was shown to be inhibited by tyrosine (17). Due to the structural and functional diversity that exists among the CMs, we determined the three-dimensional crystal structure of *MtCM. As shown in Fig. 5, the 166-residue mature *MtCM protein forms an α-helical bundle with 10 helices, as defined by PROMOTIF (22). Eighty-six percent of the residues in *MtCM are included in helices, and one disulfide bond between Cys₁₆₀ and Cys₁₉₃ was observed. Sasso et al. (48) suggested an intramolecular disulfide bond between these two cysteine residues based on studies with reductants. *MtCM forms a dimer by the interactions defined in Table 2 which involve mainly residues in helix 4. The dimer is the same as that observed in the secreted CM from *Y. pestis* (unpublished data). We used the web service DALI (18) for automated superimposition of the refined structure on pro-

teins included in the PDB (2) to look for structural neighbors. The two top hits were PDB 1ECM (EcCM) and PDB 5SCM (ScCM).

Crystal structure analysis of *MtCM reveals the catalytic site. In the known structures of CMs such as EcCM (29) and BsCM (6, 26) and the modeled structure of TtCM (17), the active sites are formed from the contribution of residues from two chains whether these proteins function as dimers or trimers. The active site of *MtCM is made up of residues from a single chain (Fig. 6A) whereas in EcCM one critical residue, Arg₁₁, from the second chain of the dimer contributes to the active site (Fig. 6B). While this article was in preparation, we found a similar structural analysis of *MtCM by Ökvist et al. (40). Ökvist et al. also determined the structure with an endo-oxabicyclic transition state analog inhibitor, thus revealing the active-site residues in *MtCM. The analysis of our structure without the transition state analog but using the very close similarity to the EcCM structure with the transition state analog revealed the active-site residues identical to those identified by Ökvist et al. (40).

TABLE 2. Residue interactions involved in dimer interface in *MtCM

Interacting residues		Distance (A to B/B to A [Å])
A/B	B/A	
Hydrophobic		
Ala ₁₀₇	Leu ₁₁₈	
Tyr ₁₁₀	Phe ₁₁₃	
Hydrogen bonds		
Glu ₆₈ OE2	Lys ₁₁₇ NZ	2.8/2.8
Arg ₁₀₃ NH1	Lys ₁₁₇ O	2.8/2.9
Glu ₁₀₆ OE1	Lys ₁₁₇ NZ	2.8/2.8
Ala ₁₀₇ O	Ser ₁₁₄ OG	2.7/2.8

FIG. 6. The active site of *MtCM. The active-site residues of *MtCM are shown in stick form and labeled in panel A. The corresponding residues in the active site of EcCM are shown and labeled in panel B (29). The chains of the dimers of *MtCM and EcCM are colored differently. It is notable that in *MtCM all the active-site residues are from the same chain, whereas in EcCM residues from both chains contribute to the active site (29). In panel C, the dimer of EcCM is superimposed onto one monomer of the *MtCM dimer.

The structure analysis of *MtCM presented here and by Ökvist et al. (40) and Qamra et al. (43) show an interesting feature of its active site and a novel fold topology for the protein. The active site is formed from a single chain in the dimer similar to that observed for ScCM (57). The catalytic site of EcCM is formed from the residues Arg₂₈, Lys₃₉, Arg₅₁, Ser₈₄, and Gln₈₈ from one chain and Arg₁₁ from the second chain (Fig. 6B) (29). The residues that contribute to the active site in *MtCM are Arg₄₉, Lys₆₀, Arg₇₂, Thr₁₀₅, Glu₁₀₉, and Arg₁₃₄ (Fig. 6A); the numbering of the amino acids includes the signal sequence. MacBeath et al. (33) and Sasso et al. (48) already predicted all the active-site residues of *MtCM based on sequence alignments. Superimposition of the active site of one monomer of *MtCM onto one of the active sites of EcCM shows that a single chain of *MtCM covers most of the dimer of EcCM (Fig. 6C). This observation raises the question as to why *MtCM functions as a dimer. We have begun experiments to investigate this point. In this context it is pertinent to mention that MacBeath et al. (34) generated a monomeric form of *Methanococcus jannaschii* CM which is catalytically as efficient as the naturally occurring dimer.

Qamra et al. (43), in their structure analysis of *MtCM, showed a binding site for tryptophan at the small dimer interface despite overwhelming evidence that *MtCM is not regulated by any of the aromatic amino acids (Fig. 4) (48). Further, we clearly demonstrated that *MtCM is secreted out of *M. tuberculosis* cells. Since the shikimate pathway is absent in mammals, the *M. tuberculosis* host, *MtCM must not have any function in the nonexistent aromatic amino acid biosynthetic pathway. Hence, it is not scientifically tangible to have such an enzyme allosterically regulated by the aromatic amino acids as previously published (42, 43).

Mutagenesis of the conserved amino acid residues in *MtCM reinforces the catalytic site. We performed site-directed mutagenesis of the putative active-site residues based on the structure analysis presented here as well as based on the sequence alignment (48). We selected Arg₄₉, Lys₆₀, Asp₆₉,

Arg₇₂, Thr₁₀₅, Glu₁₀₉, and Arg₁₃₄ for mutagenesis. All the residues were mutated to alanine. Additionally, Glu₁₀₉ was mutated to Gln for two reasons: (i) the CM_{Ec} counterpart of this residue, Gln₈₈, is an active-site residue (30), and (ii) we wanted to explore the role of this residue in catalysis as a function of pH since the ScCM counterpart of this residue is Glu₂₄₆ which has been shown to be protonated for optimal activity (49).

Translocation of *MtCM into the periplasmic compartment of *E. coli* is shown in Fig. 1 (lane 4). The wild-type and the mutant proteins were purified under identical conditions from the periplasmic fluid of the respective clones [pET15b recombinant/BL21(DE3)] to homogeneity in a single chromatographic step as described in Materials and Methods (SDS-PAGE results not shown). We observed an identical elution behavior for all the mutant proteins as the wild-type enzyme on the molecular sieve Biosep SEC-3000 HPLC column. The mutant proteins were assayed for CM activity at a 12.5 nM concentration (5 pmol). The results of the mutagenesis and the activity analysis show that Arg₄₉, Lys₆₀, Arg₇₂, and Arg₁₃₄ are essential for catalysis, having only $\leq 1\%$ of the wild-type enzyme activity (Table 3). Thr₁₀₅Ala and Glu₁₀₉Ala mutations resulted in having 20% and 10% activity, respectively, and therefore are also catalytic site residues. Based on the mutagenesis data presented here and the complementary data from the crystal structure of *MtCM with the transition state analogue (TSA) determined by Ökvist et al. (40), we propose that the active-site residues in *MtCM correspond to the active-site residues in EcCM as follows: *MtCM Arg₄₉ is EcCM Arg₂₈; *MtCM Lys₆₀ is EcCM Lys₃₉; *MtCM Arg₇₂ is EcCM Arg₅₁; *MtCM Thr₁₀₅ is EcCM Ser₈₄; *MtCM Glu₁₀₉ is EcCM Gln₈₈, and *MtCM Arg₁₃₄ is EcCM Arg₁₁. An interaction between the main chain nitrogen of Asp₆₉ and the C₄ hydroxyl group of chorismate is proposed to be analogous to the interactions in the catalytic site of EcCM (Fig. 7). Mutation of Asp₆₉ to alanine has no effect on the enzyme activity (Table

TABLE 3. Characterization of *MtCM mutants by activity analysis

Mutation	CM activity (% of wild-type level) ^a
Arg ₄₉ Ala.....	1.0 ± 0.05
Lys ₆₀ Ala.....	0.5 ± 0.0
Asp ₆₉ Ala.....	110 ± 10
Arg ₇₂ Ala.....	0.8 ± 0.0
Thr ₁₀₅ Ala.....	20 ± 0.3
Glu ₁₀₉ Ala.....	10 ± 0.7
Glu ₁₀₉ Gln.....	40 ± 3
Arg ₁₃₄ Ala.....	0

^a Each protein was assayed for CM activity at 37°C using 5 pmol of the protein (purified from the periplasmic components of *E. coli*) and 1 mM chorismate. The pH of the reaction was 7.5. The reaction was allowed to proceed for 1 to 5 min, depending on the activity, and was terminated with HCl. The results are the average of two experiments. The 100% wild-type CM activity has a k_{cat} of 60 s⁻¹ per active site (at 37°C and pH 7.5).

3), as would be expected unless the specific side chain had an influence on the backbone conformation.

We focus our discussion on the variant glutamate/glutamine at the corresponding positions in various CMs. The activity of the wild-type *MtCM with Glu₁₀₉ was optimal under broad acidic conditions with only a slight increase in the activity from pH 4.0 ($k_{\text{cat}} = 55 \pm 3 \text{ s}^{-1}$) to pH 6.5 ($k_{\text{cat}} = 69.5 \pm 4.5 \text{ s}^{-1}$) (Fig. 8). While the activity of the wild-type enzyme was stable up to pH 7.5 ($k_{\text{cat}} = 66.5 \pm 4.5 \text{ s}^{-1}$), the activity was dramatically affected under alkaline conditions ($k_{\text{cat}} = 10.8 \pm 3 \text{ s}^{-1}$ at pH 9.5). A broad pH tolerance from pH 4.0 to 7.5 for optimal activity of *MtCM is in sharp contrast to the pH optimum observed for ScCM with Glu₂₄₆ (49) and AnCM (25) with Glu₂₅₃ in the corresponding positions. ScCM (49) and AnCM (25) exhibit a bell-shaped curve at pH 5.5 and pH 5.9, respectively, having only 5% to 10% of the activity at pH 4.0 and 25% to 35% of the activity at pH 7.5. Consistent with the fact that EcCM with a naturally occurring Gln₈₈ is an active enzyme (30), we observed that the Glu₁₀₉Gln mutation in *MtCM

FIG. 8. pH dependence of the wild-type (WT) *MtCM and the mutant Glu₁₀₉Gln. The wild-type and the mutant proteins were assayed using 11 pmol of *MtCM over a broad pH range in a universal buffer (33). The reaction was carried out at 37°C and terminated after 1 min when the wild-type enzyme was used or after 3 min when the mutant enzyme was used. Chorismate was used at a 1 mM concentration.

results in ca. 40% net activity at pH 7.5 (Fig. 8 and Table 3). The Glu₁₀₉Gln mutant of *MtCM exhibited ca. 27% of the wild-type activity at pH 4.0 ($k_{\text{cat}} = 15 \pm 0.8 \text{ s}^{-1}$). The activity of this mutant gradually increased to ca. 40% of the wild-type activity with an increase in pH to 7.5 ($k_{\text{cat}} = 26.6 \pm 2 \text{ s}^{-1}$) and remained constant up to pH 11.0, unlike the activity of the wild-type enzyme. While these results show that *MtCM is active at pH 4.0 and suggest a requirement for protonation of Glu₁₀₉ in *MtCM from the observed optimal activity in the acidic pH range, it is more difficult to explain why the enzyme is as active at pH 7.5 where protonation does not efficiently occur. It is conceivable that Glu₁₀₉ could have a perturbed elevated pK_a as was demonstrated for a helicase catalyzing ATP hydrolysis (11). Based on the structure of the wild-type protein and activity analysis of the mutant proteins, we propose that Arg₄₉, Lys₆₀, Arg₇₂, Thr₁₀₅, Glu₁₀₉, and Arg₁₃₄ constitute the active site of *MtCM through hydrogen bonds and electrostatic interactions with chorismate (Fig. 7). In addition to these residues, Ökvist et al. (40) proposed Gln₇₆ and Glu₁₀₆ as part of the catalytic site based on the structure with the TSA.

What is the role of secreted CM in *M. tuberculosis*? MacBeath et al. (33) and Calhoun et al. (4) have suggested and Sasso et al. (48) and Prakash et al. (42), along with us (Patterson et al., Abstr. Am. Soc. Microbiol. Biodefense Res. Meet.), have experimentally demonstrated in a heterologous *E. coli* expression system that *MtCM is secreted into the periplasmic compartment. First of all, there is no discrete periplasmic compartment in the gram-positive *M. tuberculosis* although there is a report of possible existence of a pseudoperiplasmic compartment in this organism (39). Hence, the cellular location/dislocation of the *AroQ is another emerging attribute of CMs. So, where is *MtCM expected to be localized in its native environment? To answer this question, we tested the *M. tuberculosis* culture filtrate for CM activity. We found CM activity in the culture filtrate of early exponential phase and late exponential phase cells (Fig. 2A) and thus established that *MtCM is secreted in *M. tuberculosis*. The biological role of the highly active and secreted *MtCM in *M. tuberculosis* is elusive at this time. The secreted *MtCM must have no function in the non-

FIG. 7. A schematic representation of the active site of *MtCM modeled with chorismate. The diagram depicts the hydrogen bonds and electrostatic contacts between the amino acid residues and chorismate in the center. Although we did not determine the structure of *MtCM with a TSA, we predict that the contacts will be identical to those of the EcCM structure with the TSA except that Arg₁₃₄ from the same chain contributes to the active site.

existent shikimate pathway in the macrophages of mammals, the target of *M. tuberculosis* infection. If *MtCM were to function in pathogenesis or virulence of *M. tuberculosis*, why has it evolved to be chorismate mutase? This needs further investigation.

The success of *M. tuberculosis* as a persistent pathogen lies in its interaction with the host macrophages and inhibition of the maturation of phagosomes to phagolysosomes (41). While there may be a number of mechanisms responsible for the arrest of phagosome maturation by *M. tuberculosis*, we would like to propose a possible role in this process for *MtCM. As we demonstrated here, *MtCM is active at pH 4.5 (Fig. 8) and therefore should maintain its structural integrity in the acidic pH range. A probable reason for *MtCM to have a broad acidic pH range may have to do with its destination in the *M. tuberculosis*-infected macrophage environment, which is acidic (53). *MtCM may have some unappreciated role in the host-pathogen interaction. Finally, as the shikimate pathway enzymes are not present in humans, these enzymes are often considered as therapeutic targets. Although the secreted CM is not as attractive as a drug target as an intracellular CM, knock-out for the Rv1885c-encoded secreted CM might prove to be useful as a vaccine candidate. What can be better to combat a debilitating disease caused by the pathogen *M. tuberculosis*, which claims about 8 million humans per year (50), than to use basic science to interrupt the shikimate pathway?

ACKNOWLEDGMENTS

We are grateful to Helena Boshoff and Clifton Barry III, National Institute of Allergy and Infectious Diseases, NIH, for their generous help in providing the culture filtrate and the cell extract of a laboratory strain of *M. tuberculosis* H₃₇R_v. We are also grateful to Sheldon Morris, Laboratory of Mycobacterial Diseases and Cellular Immunology, Center for Biological Evaluation and Research, Food and Drug Administration, for providing us with the culture filtrate of a clinical isolate of *M. tuberculosis* H₃₇R_v. We thank John Jakupciak for occasional help with the ABI 3100 DNA sequencer. We are grateful to Peter Kast for valuable discussions. We appreciate many valuable comments by the anonymous reviewers.

Certain commercial equipment or materials are identified in this paper in order to specify adequately the experimental procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

REFERENCES

- Andrews, P. R., E. N. Cain, E. Rizzardo, and G. D. Smith. 1977. Rearrangement of chorismate to prephenate. Use of chorismate mutase inhibitors to define the transition state structure. *Biochemistry* **22**:4848–4852.
- Berman, H. M., J. Westbrook, Z. Feng, G. L. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. 2000. The protein data bank. *Nucleic Acids Res.* **28**:235–242.
- Brennan, P. J., and H. Nikaido. 1995. The envelope of mycobacteria. *Annu. Rev. Biochem.* **64**:29–63.
- Calhoun, D. H., C. A. Bonner, W. Gu, G. Xie, and R. A. Jensen. 2001. The emerging periplasm-localized subclass of AroQ chorismate mutases, exemplified by those from *Salmonella typhimurium* and *Pseudomonas aeruginosa*. *Genome Biol.* **2**:30.01–30.16.
- Camus, J.-C., M. J. Pryor, C. Medigue, and S. T. Cole. 2002. Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H₃₇R_v. *Microbiology* **148**:2967–2973.
- Chook, Y. M., H. Ke, and W. N. Lipscomb. 1993. Crystal structure of the monofunctional chorismate mutase from *Bacillus subtilis* and its complex with a transition state analog. *Proc. Natl. Acad. Sci. USA* **90**:8600–8603.
- Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry III, F. Tekle, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M.-A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead, and B. G. Barrell. 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**:537–544.
- Cotton, R. G. H., and F. Gibson. 1965. The biosynthesis of phenylalanine and tyrosine; enzymes converting chorismic acid into prephenic acid and their relationship to prephenate dehydratase and prephenate dehydrogenase. *Biochim. Biophys. Acta* **100**:76–88.
- Davidson, B. E., and G. S. Hudson. 1987. Chorismate mutase—prephenate dehydrogenase from *Escherichia coli*. *Methods Enzymol.* **142**:440–450.
- Dopheide, T. A. A., P. Crewther, and B. E. Davidson. 1972. Chorismate mutase-prephenate dehydratase from *Escherichia coli* K-12. II. Kinetic properties. *J. Biol. Chem.* **247**:4447–4452.
- Frick, D. N., R. S. Rypma, A. M. I. Lam, and C. M. Frenz. 2004. Electrostatic analysis of the hepatitis C virus NS3 helicase reveals both active and allosteric site locations. *Nucleic Acids Res.* **32**:5519–5528.
- Goldman, D. S. 1956. Enzyme systems in the mycobacteria. II. The malic dehydrogenase. *J. Bacteriol.* **72**:401–405.
- Gray, J. V., B. Golinelli-Pimpaneau, and J. R. Knowles. 1990. Monofunctional chorismate mutase from *Bacillus subtilis*: purification of the protein, molecular cloning of the gene, and overexpression of the gene product in *Escherichia coli*. *Biochemistry* **29**:376–383.
- Grisostomi, C., P. Kast, R. Pulido, J. Huynh, and D. Hilvert. 1997. Efficient *in vivo* synthesis and rapid purification of chorismic acid using an engineered *Escherichia coli* strain. *Bioorg. Chem.* **25**:297–305.
- Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**:557–580.
- Haslam, E. 1993. Shikimic acid: metabolism and metabolites. Wiley, New York, N.Y.
- Helmstaedt, K., G. Heinrich, R. Merkl, and G. H. Braus. 2004. Chorismate mutase of *Thermus thermophilus* is a monofunctional AroH class enzyme inhibited by tyrosine. *Arch. Microbiol.* **181**:195–203.
- Holm, L., and C. Sander. 1996. Mapping the protein universe. *Science* **273**:595–603.
- Howard, A. April 1997, posting date. Data processing in macromolecular crystallography. In P. E. Bourne and K. Watenpugh, (ed.), *Crystallographic computing 7: proceedings from the Macromolecular Crystallography Computing School*. [Online] <http://www.sdsc.edu/Xtal/IUCr/CC/School96/IUCr.html>.
- Huang, L., W. M. Nakatsukasa, and E. W. Nester. 1974. Regulation of aromatic amino acid biosynthesis in *Bacillus subtilis* 168. Purification, characterization, and subunit structure of the bifunctional enzyme 3-deoxy-D-arabinoheptulosonate 7-phosphate synthetase-chorismate mutase. *J. Biol. Chem.* **249**:4467–4472.
- Hudson, G. S., V. Wong, and B. E. Davidson. 1984. Chorismate mutase/prephenate dehydrogenase from *Escherichia coli* K-12: purification, characterization, and identification of a reactive cysteine. *Biochemistry* **23**:6240–6249.
- Hutchinson, E. G., and J. M. Thornton. 1996. PROMOTIF—a program to identify and analyze structural motifs in proteins. *Protein Sci.* **5**:212–220.
- Jacobs, W. R. Jr., G. V. Kalpana, J. D. Cirillo, L. Pascopella, S. B. Snapper, R. A. Udani, W. Jones, R. G. Barletta, and B. R. Bloom. 1991. Genetic systems for mycobacteria. *Methods Enzymol.* **204**:537–555.
- Kast, P., M. Asif-Ullah, N. Jiang, and D. Hilvert. 1996. Exploring the active site of chorismate mutase by combinatorial mutagenesis and selection: the importance of electrostatic catalysis. *Proc. Natl. Acad. Sci. USA* **93**:5043–5048.
- Krappmann, S., K. Helmstaedt, T. Gerstberger, S. Eckert, B. Hoffmann, M. Hoppert, G. Schnappauf, and G. H. Braus. 1999. The *aroC* gene of *Aspergillus nidulans* codes for a monofunctional, allosterically regulated chorismate mutase. *J. Biol. Chem.* **274**:22275–22282.
- Ladner, J. E., P. Reddy, A. Davis, M. Tordova, A. J. Howard, and G. L. Gilliland. 2000. The 1.30 Å resolution structure of the *Bacillus subtilis* chorismate mutase catalytic homotrimer. *Acta Crystallogr. D* **56**:673–683.
- Lambert, K. S., K. D. Allen, and I. M. Sussex. 1999. Cloning and characterization of an esophageal-gland-specific chorismate mutase from the phytoparasitic nematode *Meloidogyne javanica*. *Mol. Plant Microbe Interact.* **12**:328–336.
- Laskowski, R. A., M. W. MacArthur, D. S. Moss, and J. M. Thornton. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**:283–291.
- Lee, A. Y., P. A. Karplus, B. Ganem, and J. Clardy. 1995. Atomic structure of the buried catalytic pocket of *Escherichia coli* chorismate mutase. *J. Am. Chem. Soc.* **117**:3627–3628.
- Liu, D. R., S. T. Cload, R. M. Pastor, and P. G. Schultz. 1996. Analysis of active site residues in *Escherichia coli* chorismate mutase by site-directed mutagenesis. *J. Am. Chem. Soc.* **118**:1789–1790.
- Lovell, S. C., I. W. Davis, W. B. Arendall III, P. I. W. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, and D. C. Richardson. 2003. Structure validation by Cα geometry: φ, ψ, and Cβ deviation. *Proteins* **50**:437–450.

32. Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall. 1951. Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**:265–275.
33. MacBeath, G., P. Kast, and D. Hilvert. 1998. A small, thermostable, and monofunctional chorismate mutase from the archaeon *Methanococcus jannaschii*. *Biochemistry* **37**:10062–10073.
34. MacBeath, G., P. Kast, and D. Hilvert. 1998. Redesigning enzyme topology by directed evolution. *Science* **279**:1958–1961.
35. McCoy, A. J., R. W. Grosse-Kunstleve, L. C. Storoni, and R. J. Read. 2005. Likelihood-enhanced fast translation functions. *Acta Crystallogr. D* **61**:458–464.
36. McRee, D. E. 1999. Practical protein crystallography, 2nd ed. Academic Press, San Diego, Calif.
37. Murshudov, G. N., A. A. Vagin, and E. J. Dodson. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**:240–255.
38. Neu, H. C., and L. A. Heppel. 1965. The release of enzymes from *Escherichia coli* by osmotic shock and during the formation of spheroplasts. *J. Biol. Chem.* **240**:3685–3692.
39. Niederweis, M. 2003. Mycobacterial porins—new channel proteins in unique outer membranes. *Mol. Microbiol.* **49**:1167–1177.
40. Ökvist, M., R. Dey, S. Sasso, S. Grahn, P. Kast, and U. Krengel. 2006. 1.6 Å Crystal structure of the secreted chorismate mutase from *Mycobacterium tuberculosis*: novel fold topology revealed. *J. Mol. Biol.* **357**:1483–1499.
41. Pethe, K., D. L. Swenson, S. Alonso, J. Anderson, C. Wang, and D. G. Russell. 2004. Isolation of *Mycobacterium tuberculosis* mutants defective in the arrest of phagosome maturation. *Proc. Natl. Acad. Sci. USA* **101**:13642–13647.
42. Prakash, P., B. Aruna, A. A. Sardesai, and S. E. Hasnain. 2005. Purified recombinant hypothetical protein coded by open reading frame Rv1885c of *Mycobacterium tuberculosis* exhibits a monofunctional AroQ class of periplasmic chorismate mutase activity. *J. Biol. Chem.* **280**:19641–19648.
43. Qamra, R., P. Prakash, B. Aruna, S. E. Hasnain, and S. C. Mande. 2006. The 2.15 Å crystal structure of *Mycobacterium tuberculosis* chorismate mutase reveals an unexpected gene duplication and suggests a role in host-pathogen interactions. *Biochemistry* **45**:6997–7005.
44. Reddy, P., P. S. Murthy, and T. A. Venkatasubramanian. 1975. Variations in the pathways of malate oxidation and phosphorylation in different species of mycobacteria. *Biochim. Biophys. Acta* **376**:210–218.
45. Reddy, P., A. Peterkofsky, and K. McKenney. 1989. Hyperexpression and purification of *Escherichia coli* adenylate cyclase using a vector designed for expression of lethal gene products. *Nucleic Acids Res.* **17**:10473–10488.
46. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
47. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
48. Sasso, S., C. Ramakrishnan, M. Gamper, D. Hilvert, and P. Kast. 2005. Characterization of the secreted chorismate mutase from the pathogen *Mycobacterium tuberculosis*. *FEBS J.* **272**:375–389.
49. Schnappauf, G., N. Sträter, W. N. Lipscomb, and G. H. Braus. 1997. A glutamate residue in the catalytic center of the yeast chorismate mutase restricts enzyme activity to acidic conditions. *Proc. Natl. Acad. Sci. USA* **94**:8491–8496.
50. Stewart, G. R., B. D. Robertson, and D. B. Young. 2003. Tuberculosis: a problem with persistence. *Nat. Rev. Microbiol.* **1**:97–105.
51. Sträter, N., G. Schnappauf, G. Braus, and W. N. Lipscomb. 1997. Mechanisms of catalysis and allosteric regulation of yeast chorismate mutase from crystal structures. *Structure* **5**:1437–1452.
52. Studier, F. W., and B. A. Moffatt. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**:113–130.
53. Sturgill-Koszycki, S., P. H. Schlesinger, P. Chakraborty, P. L. Haddix, H. L. Collins, A. K. Fok, R. D. Allen, S. L. Gluck, J. Heuser, and D. G. Russell. 1994. Lack of acidification in *Mycobacterium* phagosomes produced by exclusion of the vesicular proton-ATPase. *Science* **263**:678–681.
54. Terwilliger, T. C. 2003. Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr. D* **59**:38–44.
55. Xia, T., and R. A. Jensen. 1992. Monofunctional chorismate mutase from *Serratia rubidaea*: a paradigm system for the three-isozyme gene family of enteric bacteria. *Arch. Biochem. Biophys.* **294**:147–153.
56. Xia, T., J. Song, G. Zhao, H. Aldrich, and R. A. Jensen. 1993. The *aroQ*-encoded monofunctional chorismate mutase (CM-F) protein is a periplasmic enzyme in *Erwinia herbicola*. *J. Bacteriol.* **175**:4729–4737.
57. Xue, Y., and W. N. Lipscomb. 1995. Location of the active site of allosteric chorismate mutase from *Saccharomyces cerevisiae*, and comments on the catalytic and regulatory mechanisms. *Proc. Natl. Acad. Sci. USA* **92**:10595–10598.
58. Xue, Y., W. N. Lipscomb, R. Graf, G. Schnappauf, and G. Braus. 1994. The crystal structure of allosteric chorismate mutase at 2.2-Å resolution. *Proc. Natl. Acad. Sci. USA* **91**:10814–10818.
59. Zuber, M., T. A. Patterson, and D. L. Court. 1987. Analysis of *nutR*, a site required for transcription antitermination in phage lambda. *Proc. Natl. Acad. Sci. USA* **84**:4514–4518.