Further Insight into S-Adenosylmethionine-dependent Methyltransferases: STRUCTURAL CHARACTERIZATION OF Hma, AN ENZYME ESSENTIAL FOR THE BIOSYNTHESIS OF OXYGENATED MYCOLIC ACIDS IN MYCOBACTERIUM TUBERCULOSIS

Fanny Boissier, Fabienne Bardou, Valérie Guillet, Sandrine Uttenweiler-Joseph, Mamadou Daffé, Annaïk Quémard, Lionel Mourey

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 7, pp. 4434–4445, February 17, 2006

Further Insight into *S*-Adenosylmethionine-dependent Methyltransferases

*STRUCTURAL CHARACTERIZATION OF Hma, AN ENZYME ESSENTIAL FOR THE BIOSYNTHESIS OF OXYGENATED MYCOLIC ACIDS IN MYCOBACTERIUM TUBERCULOSIS**

Received for publication, September 19, 2005, and in revised form, November 18, 2005 Published, JBC Papers in Press, December 15, 2005, DOI 10.1074/jbc.M510250200

Fanny Boissier[‡], Fabienne Bardou[‡], Valérie Guillet[‡], Sandrine Uttenweiler-Joseph[§¶], Mamadou Daffé[‡], Annaïk Quémard[‡], and Lionel Mourey[‡1]

From the Départements [‡]Mécanismes Moléculaires des Infections Mycobactériennes and [¶]Protéome et Cibles Thérapeutiques and the [§]Plate-Forme Protéomique, Institut de Pharmacologie et de Biologie Structurale du CNRS et de l'Université Paul Sabatier, 31077 Toulouse Cedex 04, France

Mycolic acids are major and specific components of the cell envelope of Mycobacteria that include *Mycobacterium tuberculosis*, the causative agent of tuberculosis. Their metabolism is the target of the most efficient antitubercular drug currently used in therapy, and the enzymes that are involved in the production of mycolic acids represent important targets for the development of new drugs effective against multidrug-resistant strains. Among these are the *S*-adenosylmethionine-dependent methyltransferases (SAM-MTs) that catalyze the introduction of key chemical modifications in defined positions of mycolic acids. Some of these subtle structural variations are known to be crucial for both the virulence of the tubercle bacillus and the permeability of the mycobacterial cell envelope. We report here the structural characterization of the enzyme Hma (MmaA4), a SAM-MT that is unique in catalyzing the introduction of a methyl branch together with an adjacent hydroxyl group essential for the formation of both keto- and methoxymycolates in *M. tuberculosis*. Despite the high propensity of Hma to proteolytic degradation, the enzyme was produced and crystallized, and its three-dimensional structure in the apo-form and in complex with *S*-adenosylmethionine was solved to about 2 Å. The structures show the important role played by the modifications found within mycolic acid SAM-MTs, especially the α2-α3 motif and the chemical environment of the active site. Essential information with respect to cofactor and substrate binding, selectivity and specificity, and about the mechanism of catalytic reaction were derived.

Mycolic acids, α-branched β-hydroxylated long chain fatty acids, are the hallmark of the *Mycobacterium* genus that comprises the causative agents of human diseases such as tuberculosis and leprosy, *Mycobacterium tuberculosis* and *Mycobacterium leprae*, respectively. These major cell envelope components play an important role in the structure and function of the mycobacterial cell envelope (1, 2). For instance, mycolic acids attached to the cell wall arabinogalactan are organized with other lipids to form an outer permeability barrier with an extremely low fluidity that confers an exceptional low permeability to mycobacteria and may explain their intrinsic resistance to many antibiotics (3). Similarly, trehalose mycolates have been implicated in numerous biological functions related both to the physiology and virulence of *M. tuberculosis* (1).

Numerous studies have been and are currently devoted to understanding the structures and biosynthesis of mycolic acids, primarily because they are specific to the *Mycobacterium* genus, and their metabolism is the only clearly identified target inhibited by the major anti-tubercular drug isoniazid (4–8). With the resurgence of tuberculosis infections caused by multidrug-resistant strains and the need for the development of new anti-tubercular drugs, deciphering the biosynthesis pathway leading to mycolates still represents a major objective. Although much work remains to be done to complete their biosynthetic scheme, it is known that two mycobacterial fatty acid synthases participate in the formation of all types of mycolates. Fatty acid synthase I is necessary to produce $C_{16,18}$ and C_{22-26} saturated fatty acids, which may be either directly incorporated into mycolates as the α-branch chain or used as substrates of the acyl carrier protein-dependent elongation system, fatty acid synthase-II, to produce the long meromycolic chain (for a recent review, see Ref. 9).

Mycolic acids usually occur in mycobacterial species as a mixture of various related molecules with different chemical groups at the so-called "proximal" and "distal" positions of their meromycolic chain (Fig. 1*A*). In members of the *M. tuberculosis* complex (*M. tuberculosis*, *Mycobacterium africanum*, *Mycobacterium bovis*, *Mycobacterium microti*, and *Mycobacterium canetti*), three types of mycolates are commonly encountered (10, 11). The least polar type I α-mycolates are composed of C_{76-82} fatty acids (12) and contain two *cis* cyclopropyl groups. The more polar type III and IV mycolates consist of C_{82-89} (12) and contain a *cis* or *trans* (with a methyl group on the vicinal carbon atom) cyclopropyl group at the proximal position and a keto or methoxy group (with a methyl group on the vicinal carbon atom) at the distal position (Fig. 1*A*). These discrete structural variations in mycolates may be of crucial biological importance since it has been shown that mutations resulting in the loss of these chemical functions, and particularly the keto and methoxy groups, profoundly modify the permeability of the cell envelope to solutes and severely affect the virulence and pathogenicity of the mutant strains in experimental infections (13–16). Accordingly, the enzymic systems that introduce the chemical modifications in the mycolic acid chain merit special attention. One of the eight genes that encode putative mycolic acid *S*-adenosylmethionine (SAM)[2]-dependent methyltransferases (MTs) in *M. tuberculosis* (17), the *mma4* (18) or *hma* (15) gene, has been shown to be necessary and sufficient for the synthesis of both keto- and methoxymycolic acids. Indeed, inactiva-

* This work was supported by the CNRS and the Ministère de l'Education Nationale de l'Enseignement Supérieur et de la Recherche (Program Action Concertée Incitative: Molécules et Cibles Thérapeutiques). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] To whom correspondence should be addressed: IPBS-CNRS (UMR 5089), 205 route de Narbonne, 31077 Toulouse Cedex 04, France. Tel.: 33-561-175-436; Fax: 33-561-175-994; E-mail: lionel.mourey@ipbs.fr.

[2] The abbreviations used are: SAM, *S*-adenosylmethionine; SAM-MT, *S*-adenosylmethionine-dependent methyltransferase; MT, methyltransferase; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; SAH, *S*-adenosylhomocystein; CTAB, cetyltrimethylammonium bromide; DDDMAB, didecyldimethylammonium bromide; AcpM, acyl carrier protein from *M. tuberculosis*; MES, 4-morpholineethanesulfonic acid; r.m.s.d., root mean square deviation.

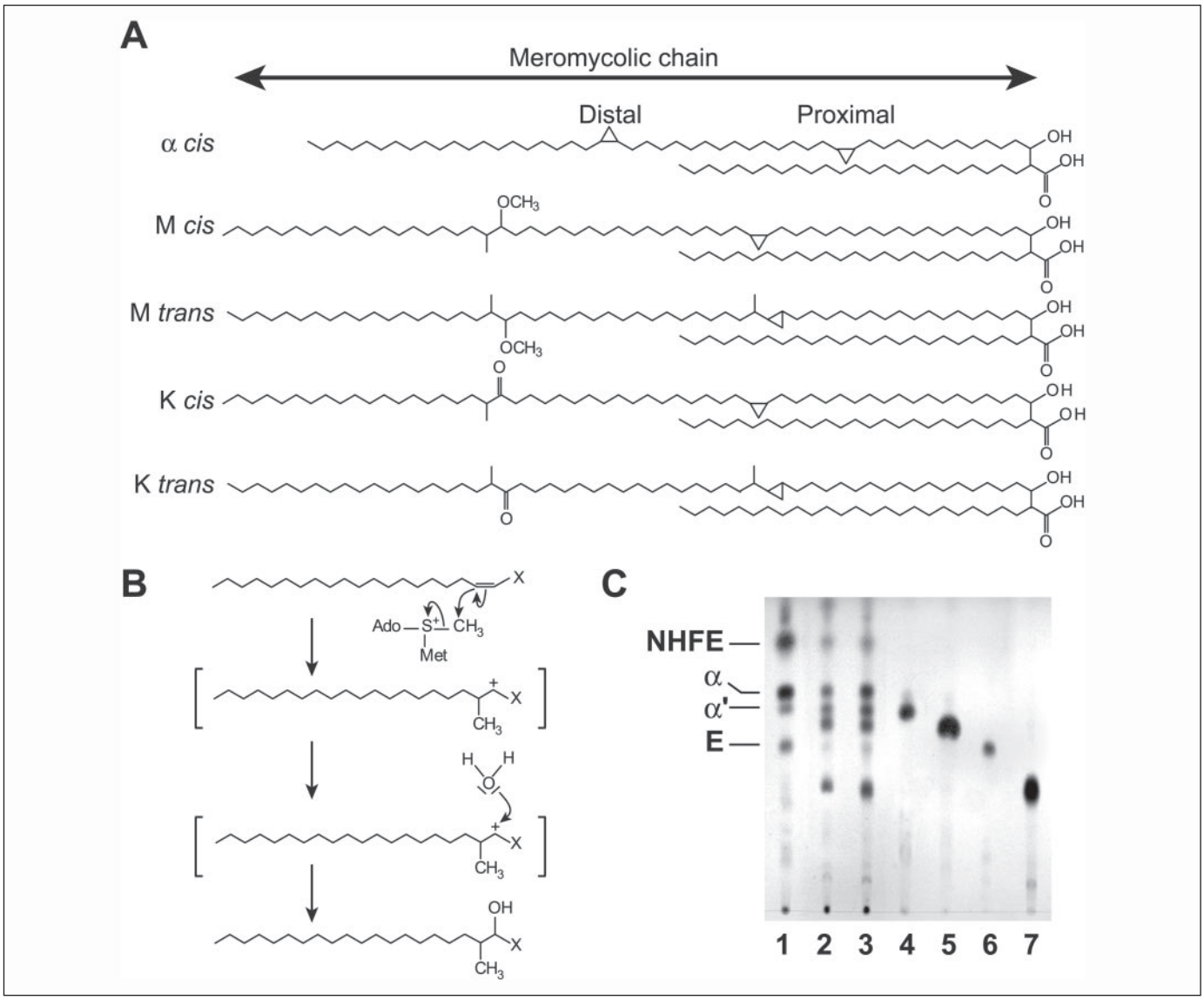

FIGURE 1. *A*, structures of α-methoxy- (*M*), and keto- (*K*) mycolic acids from *M. tuberculosis*. The *cis* or *trans* geometry of the proximal cyclopropyl is indicated. The major sizes of mycolic acids in *M. tuberculosis* H37Rv are C_{78-80} for α and C_{84-87} for oxygenated mycolic acids. *B*, hypothetical mechanism for the reaction catalyzed by the *S*-adenosylmethionine-dependent methyltransferase Hma. *Ado*, adenosyl; $X = (CH_2)_n$-CO carrier, where $n \geq 0$. *C*, TLC profile of the mycolic acid methyl esters from *M. smegmatis* recombinant strains. *Lane 1*, *M. smegmatis*/pMV261; *lane 2*, *M. smegmatis*/pMV261::*hma*; *lane 3*, *M. smegmatis*/pMV261::*h-hma*; *lane 4*, purified α′-mycolate; *lane 5*, ketomycolate; *lane 6*, epoxymycolate (*E*); *lane 7*, hydroxymycolate. *NHFE*, non-hydroxylated fatty esters.

tion of the gene resulted in the suppression of the synthesis of both types of mycolic acid (12, 15, 19). Consistently, transformation of *Mycobacterium smegmatis* with the *hma* gene resulted in the production of large amounts of hydroxymycolic acids with an adjacent methyl branch, structurally related to the ketomycolic acids of *M. tuberculosis* (20). Trace amounts of these hydroxymycolic acids were also detected in mycobacterial species producing keto- and/or methoxymycolates, further supporting the hypothesis that hydroxymycolic acids can be the precursors of both keto- and methoxymycolic acids (21). Analysis of the mycolic acids elaborated by the *hma* mutant showed that the strain accumulates a new type of α-mycolate in which a distal double bond replaces a cyclopropyl ring (19). Thus, the Hma protein would transfer a methyl group from the SAM cofactor and, subsequently or simultaneously, a water molecule onto the double bound of ethylene substrates, leading to the formation of an hydroxylated product (15, 18, 20, 21) (Fig. 1*B*). Other evidence also suggests that SAM-dependent methyltransferases would operate on full-length meromycolic derivatives (13, 22). To understand the mechanism of catalysis of this reaction and as a prerequisite for drug design, we have studied the Hma protein and solved its three-dimensional structure in the apo-form and in complex with SAM to 2.1 and 2.0 Å of resolution, respectively.

MATERIALS AND METHODS

Analysis of the Mycolic Acid Profile of Recombinant M. smegmatis Strains—The *hma* (*mmaA4*, *Rv0642c*) gene from *M. tuberculosis* H37Rv was amplified by PCR from genomic DNA and first cloned into the BamHI restriction site of the pQE30 plasmid (Qiagen), downstream the poly-His-coding region. This construction removed the first methionine residue from the sequence deduced from the gene and added 12 residues, including a non-cleavable His_6-tag at the N terminus, leading to a protein that contains 312 amino acid residues. Using this construct as a template, the *hma* gene alone or together with the poly-His tag (*h-hma*) was amplified by PCR and cloned into the EcoRI and HindIII sites of the mycobacterial expression vector pMV261. The strain

M. smegmatis mc^{2}155 was transformed by electroporation (23) with the plasmid pMV261, pMV261::*hma*, or pMV261::*h-hma*. The recombinant strains were selected on 7H10 medium (Middlebrook) supplemented with 0.2% glycerol and 10 μg/ml kanamycin. Isolated colonies were used to inoculate cultures grown on the same medium. Bacteria were submitted to saponification as previously described (12). After methylation using diazomethane, the mycolate patterns were analyzed by TLC on silica gel 60-coated plates (0.25 mm thickness, Macherey-Nagel) with elution in dichloromethane. Purified mycolic acid methyl esters were used as standards (10). Fatty acid methyl esters were revealed by spraying molybdophosphoric acid (10% in ethanol) and charring.

Expression and Purification of Hma for Structural Studies—The construction used for protein expression and purification was obtained by cloning the *hma* gene into the NdeI and BamHI sites of the expression vector pET-15b (Novagen). This construction removed the first three residues (Met-Thr-Arg) from the sequence deduced from the gene and added 20 residues, including a 17-residue-long cleavable His_6-tag at the N terminus, leading to a fusion protein that contains 318 amino acid residues. The overexpression of the recombinant Hma fusion protein was carried out in *Escherichia coli* BL21(DE3)pLysS. Cultures were grown at 37 °C in Luria broth supplemented with 50 μg/ml ampicillin. One millimolar isopropyl-1-thio-β-D-galactopyranoside was added for induction when cell density reached an A_{600} of ~0.5–0.9, and cell cultures were incubated for another 3 h at 37 °C. Cells were harvested by centrifugation at 3000 × *g* for 10 min at 4 °C, and the pellets were washed with 50 mM MES-NaOH, 0.3 M NaCl, pH 6.5, and placed overnight at −20 °C. Bacteria were then lysed by sonication and centrifuged at 5000 × *g* for 1 h at 4 °C.

Recombinant Hma was purified by using fast protein liquid chromatography on an ÄKTA Purifier system (Amersham Biosciences). First, the supernatant was applied to a chelating-Sepharose nickel affinity column (Amersham Biosciences). After extensive washes with 5 mM imidazole in 50 mM MES-NaOH, 0.3 M NaCl, pH 6.5, the protein was eluted with 0.15 M imidazole in the same buffer. Fractions corresponding to the protein were pooled and concentrated, and the engineered His_6 tag was removed by thrombin (Novagen) (2 units/mg of protein, 4 h at 20 °C) if necessary. Concentrated tagged and untagged proteins were further purified by gel filtration with a HiLoad 16/60 Superdex-75 prepgrade chromatography column (Amersham Biosciences) and both eluted with 50 mM MES-NaOH, 0.15 M NaCl, 2 mM EDTA, 0.2 mM phenylmethylsulfonyl fluoride, pH 6.5. Fractions containing pure recombinant proteins were pooled, concentrated to 3–10 mg/ml using Centricon YM-10 (Millipore) devices, and stored at 4 °C. Protein concentration was determined by measuring UV spectra ($A_{280\ =}$ 0.93 at 1 mg/ml, 1 cm).

Electrophoretic Analysis and Mass Spectrometry—Protein purity was checked throughout purification using SDS-PAGE with a 15% acrylamide concentration. For mass spectrometry, SDS-PAGE separations were conducted using 12% polyacrylamide gels stained with Coomassie Brilliant Blue. Passive elution of proteins from polyacrylamide gels was achieved as described in Kurth and Stoffel (24) and Claverol *et al.* (25). Briefly, the gel pieces were excised and subsequently washed with H_2O, and the proteins were allowed to diffuse out of the gel overnight at 37 °C by incubation in 20 μl of 0.1 M sodium acetate, 0.1% SDS, pH 8.2. Coomassie Brilliant Blue, SDS and salts were removed from the protein sample after passive elution by hydrophilic interaction chromatography using a ZipTipHPL according to the manufacturer's instructions (Millipore). Briefly, the ZipTipHPL was rehydrated in buffer A (H_2O/CH_3CN/CH_3COOH 50/50/0.1, pH 5) and equilibrated with buffer B (H_2O/CH_3CN/CH_3COOH 10/90/0.1, pH 5.5). Protein eluates were diluted in 200 μl of buffer B and loaded onto the ZipTipHPL. Salts were removed by washing with buffer B, and proteins were eluted with 4 μl of H_2O/CH_3CN/HCOOH (49/50/1).

Electrospray ionization analysis was performed using an electrospray ionization quadrupole-time of flight mass spectrometer (QSTAR Pulsar, Applied Biosystems, Foster City, CA) operating in positive mode. A potential of 1–2 kV was applied to the precoated nanoelectrospray needles (New Objective, Picotips, Econotips) in the ion source. Instrument operation, data acquisition, and analysis were performed using Analyst® QS 1.0 software and Bioanalyst TM extensions.

After extraction of the protein from the gel spot and desalting, acetonitrile was evaporated from the eluate at room temperature. Five microliters of trypsin solution (Promega) at 12.5 ng/μl in 12.5 mM NH_4HCO_3 were added, and the sample was incubated overnight at 37 °C. MALDI-TOF mass spectroscopy analyses were performed on a MALDI-TOF/TOF instrument (4700 Proteomics Analyzer; Applied Biosystems). A 0.5-μl volume of trypsin digest was applied on the MALDI target plate with 0.3 μl of matrix solution (α-cyano-4-hydroxycinnamic; 5 mg/ml in H_2O/acetonitrile/TFA, 50/50/0.1). Mass spectra were acquired in automated positive reflector mode, from *m*/*z* 700 to 3500 and calibrated with external calibration. Peak lists from peptide mass mapping spectra were compared manually with the theoretical molecular masses of the trypsin peptides of Hma.

Crystallization and X-ray Data Collection—Crystallization was performed at 12 °C by vapor equilibration using the hanging-drop method. Protein samples were concentrated to 3–10 mg/ml in the appropriate buffer (50 mM MES, 50 mM NaCl, pH 6.5). Drops were prepared by mixing equal volumes of protein and reservoir solutions; reservoir volumes of 500 μl were used. Basic, extension, and low ionic screens from Sigma were systematically used for initial screenings. X-ray diffraction quality crystals of Hma were obtained for the His_6-tagged protein in the presence of 4–28% polyethylene glycol 3350, pH 5–9.

All crystals were cryocooled in a stream of nitrogen gas at 100 K after a 3-min immersion in the crystallization solution supplemented with 20% (v/v) glycerol and stored in liquid nitrogen if necessary. For preparation of the binary complex with the cofactor, crystals were soaked in a solution containing both the cryoprotectant and 50 mM *S*-adenosylmethionine for 2–3 min. The various crystal forms were evaluated in-house at 285 and 100 K using a Rigaku RU300 rotating-anode source operating at 50 kV and 90 mA and a MarResearch Mar345dtb image-plate area detector. Diffraction data used for structure determination and refinement were collected at multistation beam-line ID14 of the European Synchrotron Radiation Facility (Grenoble, France).

Data Processing and Phasing—All crystallographic calculations were performed using the CCP4 suite (26) as implemented in the graphical user interface (27). X-ray diffraction data were processed using MOSFLM (28) and scaled with SCALA (29). The structure of Hma in its apo-form was solved using molecular replacement with the program Phaser (30) and the structure of apoCmaA1 (31) (PDB entry code 1KP9). The search model was truncated as a polyalanine except for strictly conserved residues among the family of mycolic acid SAM-MTs from *M. tuberculosis*. The structure of Hma in complex with its cofactor was solved using the refined model of apoHma and molecular replacement with the program Phaser to compensate for variation along the unit cell *c* axis.

Model Building and Crystallographic Refinement—Model building of apoHma was first carried out with ARP/wARP using the "warpNtrace" automated procedure (32). The initial map used for the calculation corresponded to the molecular replacement solution where several regions

of the protein with different conformation were removed. All structures were then constructed manually in sigmaA-weighted electron density maps (33) using TURBO-FRODO. Restrained refinements of the structures were performed with the program REFMAC (34) using a bulk solvent correction based on the Babinet principle and minimizing a maximum likelihood target function. Solvent molecules were automatically added as neutral oxygen atoms using wARP (35). In the last stages of refinement, TLS parameters were refined using a single group for the whole molecule, which resulted in a similar improvement of the *R* and R_{free} values.

Production of the Figures—Fig. 3*A* was produced using TopDraw from the CCP4 graphical user interface (27). Figs. 3*B*, 5, and 6 were produced using BobScript (36). Fig. 3*B* was rendered using RASTER3D (37). Fig. 4 was produced using ESPript (38) with a sequence alignment edited manually and secondary structure assignment performed with STRIDE (39). Figs. 7–9 were produced using PyMol (58).

RESULTS

Production and Characterization of Hma—The *hma* gene was cloned both into pQE30 and pET-15b expression vectors downstream of a His_6 tag-encoding sequence, and these constructs were used to transform *E. coli* JM109 and BL21(DE3)pLysS strains, respectively. Recombinant Hma proteins were overproduced from these strains and purified to homogeneity by metal-chelating affinity and size exclusion chromatography. As observed by SDS-PAGE, the tagged protein purified from the *E. coli* JM109/pQE30::*hma* strain and stored at 4 °C underwent progressive proteolysis ultimately leading to total degradation (data not shown). This phenomenon was less marked for the tagged and untagged proteins purified from *E. coli* BL21(DE3)pLysS/pET-15b::*hma*, which were, thus, subsequently used throughout this work. Nevertheless, SDS-PAGE of the latter proteins, freshly prepared and stored at 4 °C, led to the appearance of two fragments. Moreover, we observed that these two fragments underwent further processing until they reached a stable size within a few weeks (Fig. 2*A*). The different molecular species, *i.e.* tagged, untagged, and cleaved Hma, were characterized in native conditions. Analytical size exclusion chromatography experiments showed no significant change in the elution volume of Hma, and cleaved Hma and native PAGE of the cleaved protein revealed a single band (data not shown). These results suggest that the two fragments of cleaved Hma remain associated.

Mass spectrometry analyses were then performed to better characterize these fragments. The bands were excised from the SDS gel (Fig. 2*A*), and the corresponding proteins were eluted, desalted, and analyzed by electrospray ionization-mass spectroscopy. For the untagged protein (*bands 1* and *3*), the experimental molecular mass (34,561 Da) is in good agreement with the theoretical mass of Hma (34,563 Da). On the other hand, the mass shift observed between the measured (*band 2*) and expected molecular masses of His_6-tagged Hma (36,313 ± 1 Da and 36,445 Da, respectively) suggests the loss of the first methionine residue (131 Da). Molecular masses of proteins corresponding to bands 4 and 5 were 21,738 ± 3 and 12,838 ± 2 Da, respectively. The discrepancy between the sum of these two molecular masses (34,576 ± 5 Da) and the theoretical value for Hma (34,563 Da) may correspond to the formation of the peptide bond (−18 Da), suggesting that fragments 4 and 5 arose from a single cleavage of the untagged protein. In addition, MALDI-TOF analyses were conducted on trypsin digests of proteins in bands 1, 2, 6, and 7. Peptide mass fingerprints of bands 1 and 2 confirmed the presence of untagged and tagged Hma, respectively. The difference between mass spectra of bands 1 and 2 relies on the specific presence of a unique peptide at *m*/*z* 1768.83. Data base searches allowed identifying

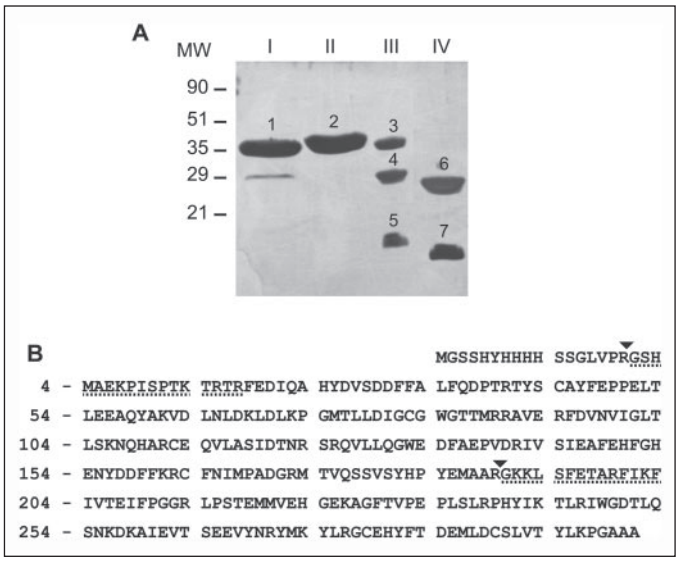

FIGURE 2. **Analysis of the purified Hma protein.** *A,* SDS-PAGE of untagged Hma (*lane I*), tagged Hma (*lane II*), untagged Hma stored at 4 °C during 2 and 4 weeks (*lanes III* and *IV*, respectively). Proteins corresponding to *bands 1–7* were further analyzed by mass spectrometry. *B,* sequence coverage of the different Hma molecular entities. Electrospray ionization quadrupole-time of flight and MALDI-TOF experiments (i) allowed precise characterization of bands 1 and 3 as an untagged protein and band 2 as a tagged protein in which the first methionine was lost, (ii) showed that bands 4 and 5 correspond to the N- and C-terminal fragments, respectively, of untagged Hma cleaved between Arg-189 and Gly-190, and (iii) provided evidence of further proteolysis leading to fragments 6 and 7 due to the N-terminal degradation (*underlined regions*) of fragments 4 and 5, respectively. *Arrow heads* indicate cleavage sites.

this peptide as the fusion peptide of the tagged protein and confirmed the methionine loss. Peptide mass fingerprints of proteins corresponding to bands 6 and 7 covered the N- and C-terminal portions of the Hma protein sequence (Fig. 2*B*). Molecular weight determination of fragments corresponding to bands 4 and 5 and mass mapping of fragments corresponding to bands 6 and 7 suggest the existence of a specific proteolytic cleavage between arginine 189 and glycine 190 followed by additional processing at the N termini of primary fragments and more specifically between amino acids 189 and 204 (Fig. 2*B*).

As a control to assess the susceptibility of Hma to proteolysis, tagged and untagged proteins were both submitted to the specific endopeptidase activity of members of the chymotrypsin family (40). Interestingly, SDS-PAGE analyses of digestion products revealed patterns similar to those mentioned above (data not shown), thus confirming the existence of a proteolytic susceptibility domain. Accordingly, the region covering residues 189–204 contains several residues compatible with putative P_1 sites and corresponds to helix α3, which is exposed to solvent. To limit cleavage, purification of the tagged protein was performed in 50 mM MES-NaOH, 0.15 M NaCl, pH 6.5, supplemented with 2 mM EDTA and 0.2 mM phenylmethylsulfonyl fluoride, which consequently insured a much longer life time of the protein.

Finally, we checked that a polyhistidine tag does not impair the function of Hma *in vivo* within *Mycobacterium*. For this purpose, *M. smegmatis* was transformed with vectors expressing either the *hma* gene alone or together with a sequence encoding a His_6-tag. TLC analysis revealed that the mycolic acid profile of both recombinant strains was similar, thus showing that the presence of the His_6-tag at the N terminus of the Hma protein does not affect its function (Fig. 1*C*).

Crystallization and Structure Determination—The crystallization condition that initially led to crystalline material was observed with the untagged and cleaved proteins, where clusters of needles and plates were obtained, respectively. However, none of these two crystal forms diffracted x-rays. In contrast, crystals suitable for diffraction studies

were obtained with freshly prepared His_6-tagged Hma in 50 mM MES, 50 mM NaCl, pH 6.5, in the presence of EDTA and phenylmethylsulfonyl fluoride. Diamonds, cubes, and sword-shaped rods were grown using 4–28% polyethylene glycol 3350 at pH 5–9 in the appropriate buffer. In fact, these three forms only differ by their crystal habit since preliminary crystallographic analysis revealed the same lattice type and similar cell parameters. Two complete data sets were collected from single crystals corresponding to Hma in its apo-form (apoHma) and in complex with S-adenosylmethionine (Hma-SAM) to 2.1 and 2.0 Å resolution, respectively. Statistics of the x-ray data processing are given in Table 1. Crystals of Hma were unambiguously assigned to the trigonal $P3_121$ space group based on data scaling and molecular replacement calculations. There is one molecule per asymmetric unit giving a V_M of protein (Matthews coefficient) and a solvent content of 2.60 $Å^3$ Da^{-1} and 53%, respectively.

The structure of apoHma was solved using molecular replacement and the three-dimensional coordinates of apoCmaA1 (31) as a search model. A unique solution was obtained with Phaser, giving Z-scores (number of S.D. above the mean value) for the rotation and translation functions of 7.3 and 14.3, respectively, and a log (likelihood gain) of 164.2. The structure of apoHma was then used as a search model for the molecular replacement of Hma-SAM to compensate for variation along the unit cell *c* axis. The final structure of apoHma comprises 277 of 318 amino acids found in the tagged protein. The final *R* and R_{free} factors are 0.195 and 0.232, respectively, for all data between 30.0 and 2.1 Å (Table 2). The refined model of Hma-SAM comprises 281 amino acids, and the final *R* and R_{free} factors are 0.179 and 0.239, respectively, for all data between 30.0 and 2.0 Å. Both structures have >94% of the residues in the most favored region of the Ramachandran plot and none in the disallowed region, as defined by PROCHECK (41). The missing residues had poorly defined electron density and belong to the N termini in both cases and to a loop only disordered in the structure of the apo-form. However, SDS-PAGE of dissolved crystals revealed a single protein band corresponding to the tagged protein.

TABLE 1
Data collection statistics

Crystal	apoHma	Hma-SAM
Beam line	ID14–3	ID14–4
Space group	$P3_121$	$P3_121$
Unit cell dimensions, Å	$a = b = 56.7$ $c = 206.2$	$a = b = 56.8$ $c = 199.0$
Molecules/asymmetric unit	1	1
Maximum resolution[a], Å	2.1 (2.21-2.10)	2.0 (2.11-2.00)
No. measured reflections	220,083 (33,198)	98,707 (10,537)
No. unique reflections	22,776 (3,315)	25,549 (3,284)
Completeness, %	97.5 (100)	97.7 (88.1)
R_{sym}, %	10.0 (18.6)	8.5 (22.3)
$I/\sigma I$	3.8 (3.7)	4.8 (2.9)
B_{Wilson}, $Å^2$	42.3	30.6

[a] The numbers in parentheses are for the highest resolution shell.

Overall Structure of Hma—The tertiary structure of Hma consists of the so-called core SAM-MT fold (for a review, see Ref. 42) with an embellishment pattern characteristic of small molecule and lipid SAM-MTs (Fig. 3*A*). The protein core contains seven strands arranged in a mix β-sheet, in the order β3-β2-β1-β4-β5-β7-β6 where all strands except strand β7 are parallel, flanked on each side by three helices (Fig. 3*B*). Strand β6, which is six residues long and contains an antiparallel classic β-bulge in other known structures of mycolic acid SAM-MTs, is extremely short in Hma, with only two contributing residues, Leu-235 and Ser-236. This is due to a distortion of the polypeptide backbone at positions Pro-232—Glu-233—Pro-234, which provides a local twist and outward protrusion, thereby creating a small cavity that is filled with 3 water molecules. These water molecules are hydrogen-bond partners for the main chain oxygen atoms of residues Pro-232 and Glu-233 and the main chain nitrogen atom of residue Thr-293 on strand β7 and, thus, ensure stabilization. The six helices (αZ, αA to αE) run in the same N to C direction on both sides of the central β-sheet. They are α-helical in nature and about the same length (10–15 residues) except for the short helix C inserted between strands β3 and β4. Helix C, which is not always conserved in the core fold of the SAM-MTs (42), comprises three residues that form one turn of the 3_{10} helix in all mycolic acid SAM-MTs, including Hma. Like other SAM-MTs, Hma displays individual variations (the embellishment pattern) to the core fold. As in other mycolic acid SAM-MTs (42), these variations correspond to 3_{10}-

TABLE 2
Refinement statistics

Crystal	ApoHma	Hma-SAM
Resolution range[a], Å	30-2.1 (2.15-2.10)	30-2.0 (2.05-2.00)
No of residues/total expected[b]	277/318	281/318
Missing residues	tag, 4–21, 151–153	tag, 4–20
Missing side chains	Lys-106, Glu-149, Glu-154, Lys-258, Glu-261	Ile-21, Lys-106, His-182, Tyr-184, Lys-191, Lys-202, Glu-261, Glu-265
No of protein atoms	2218	2,236
No of water molecules	124	153
No of cofactor atoms		27
No of heteroatoms	2342	2416
No of reflections work/test[a]	20,362/1,173 (1,570/ 86)	22,882/1302 (1,482/93)
Crystallographic *R* factor/R_{free}[a]	0.195/0.232 (0.207/0.260)	0.179/0.239 (0.201/0.304)
r.m.s.d. bond lengths, Å	0.021	0.016
r.m.s.d. bond angles, °	1.736	1.539
r.m.s.d. planarity, Å	0.008	0.006
Mean temperature factor, $Å^2$		
Main chain[c]	53.3	34.2
Side chain[c]	56.1	36.9
Solvent	55.0	42.0
Cofactor		35.6

[a] The numbers in parentheses are for the highest resolution shell.
[b] Taking into account the changes brought to the sequence.
[c] Full *B* factors that include the contribution from the TLS parameters.

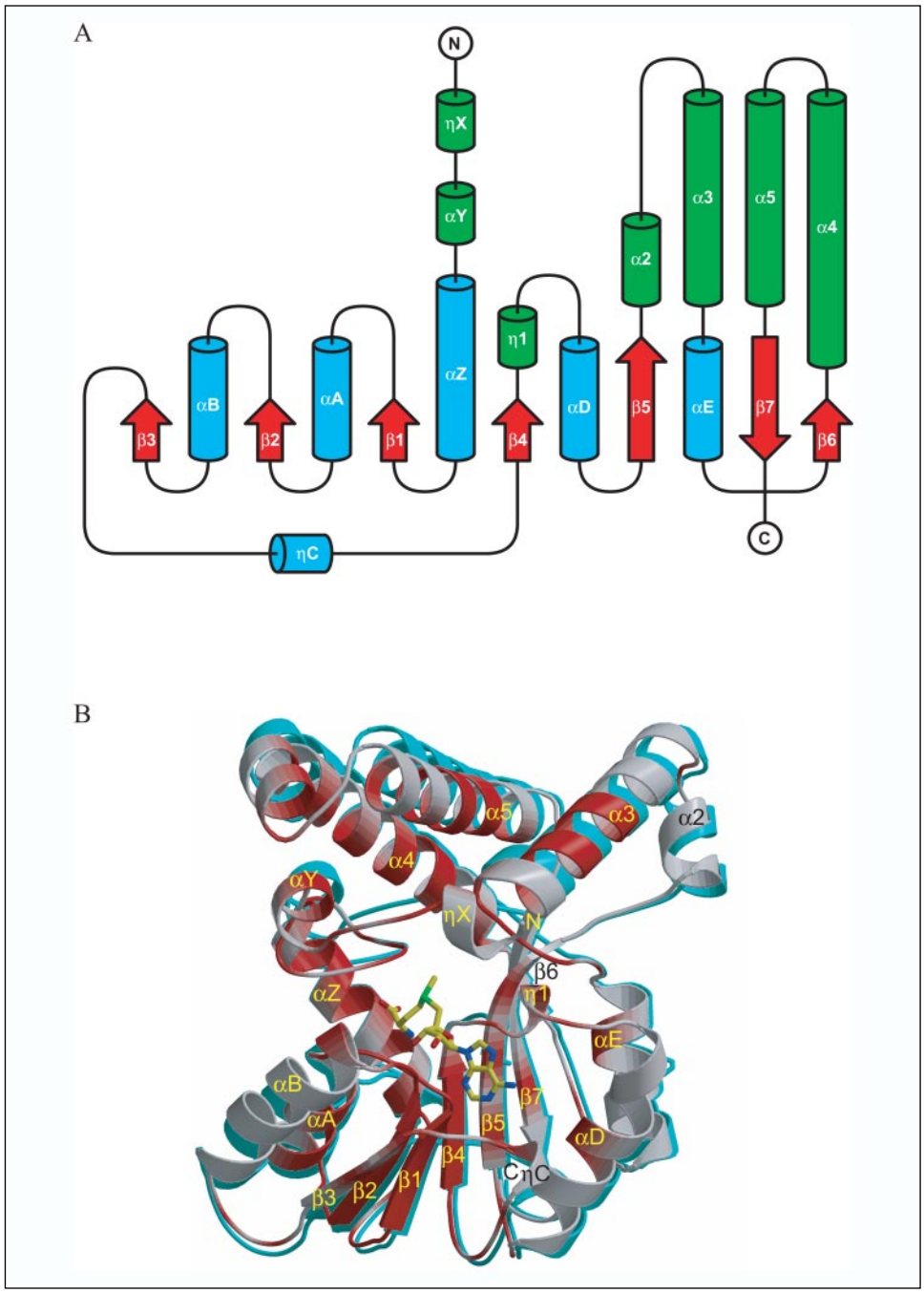

FIGURE 3. **Three-dimensional structures of Hma.** *A*, topology diagram showing helices and *β*-strands of the conserved SAM-MTs core in *cyan* and *red*, respectively. Helices of the embellishment pattern are shown in *green*. *B*, ribbon representation of the superimposed structures of apo-Hma (*cyan*) and Hma-SAM colored according to sequence similarity as in Fig. 4. The N and C termini and the secondary structure elements are labeled. The cofactor is shown as *sticks* with carbon, nitrogen, oxygen, and sulfur atoms in *yellow*, *blue*, *red*, and *green*, respectively.

and α-helical N-terminal additions (ηX and αY, respectively), to the insertion of a short 3_{10} helix (η1) between strand β4 and helix αD, and to insertions of two long antiparallel alpha helices both between strand β5 and helix αE (α2-α3) and between strands β6 and β7 (α4-α5) (Fig. 3*A*). It should be noted that helices ηX and η1 are only formed upon binding of the cofactor and are, thus, absent in the structure of apoHma. All helical inserts form an active site cover lying on top of the core, on the C-terminal side of the central β-sheet (Fig. 3*B*). The amino acid sequence of Hma comprises six cysteine residues at positions 44, 82, 112, 163, 278, and 289, three of which (Cys-44, -82, and -278) are conserved among mycolic acid SAM-MTs (Fig. 4). These cysteine residues are neither exposed at the surface of the protein nor engaged in disulfide bridges despite the close proximity of Cys-44 and Cys-289 and of Cys-82 and Cys-112, whose SG atoms are 4.7 and 4.0 Å apart, respectively (the typical distance observed in a disulfide bridge is 2.05 Å). Rather, the cysteine residues of Hma participate in the set of hydrophobic interactions that help stabilize the tertiary structure of the protein. Furthermore, there is one *cis* peptide bond formed between residues Glu-48 and Pro-49. Glu-48 is strictly conserved among mycolic acid SAM-MTs, whereas Pro-49 is only found in the sequence of Hma, with an arginine residue in all other sequences (Fig. 4).

The asymmetric unit of Hma crystals consists of a single molecule. Inspection of the crystal packing did not allow the identification of a specific interface that could mediate oligomerization through crystallographic symmetry. This corroborates results obtained using analytical gel filtration and small angle x-ray scattering, which both indicate that the protein behaves as a monomer in solution (data not shown).

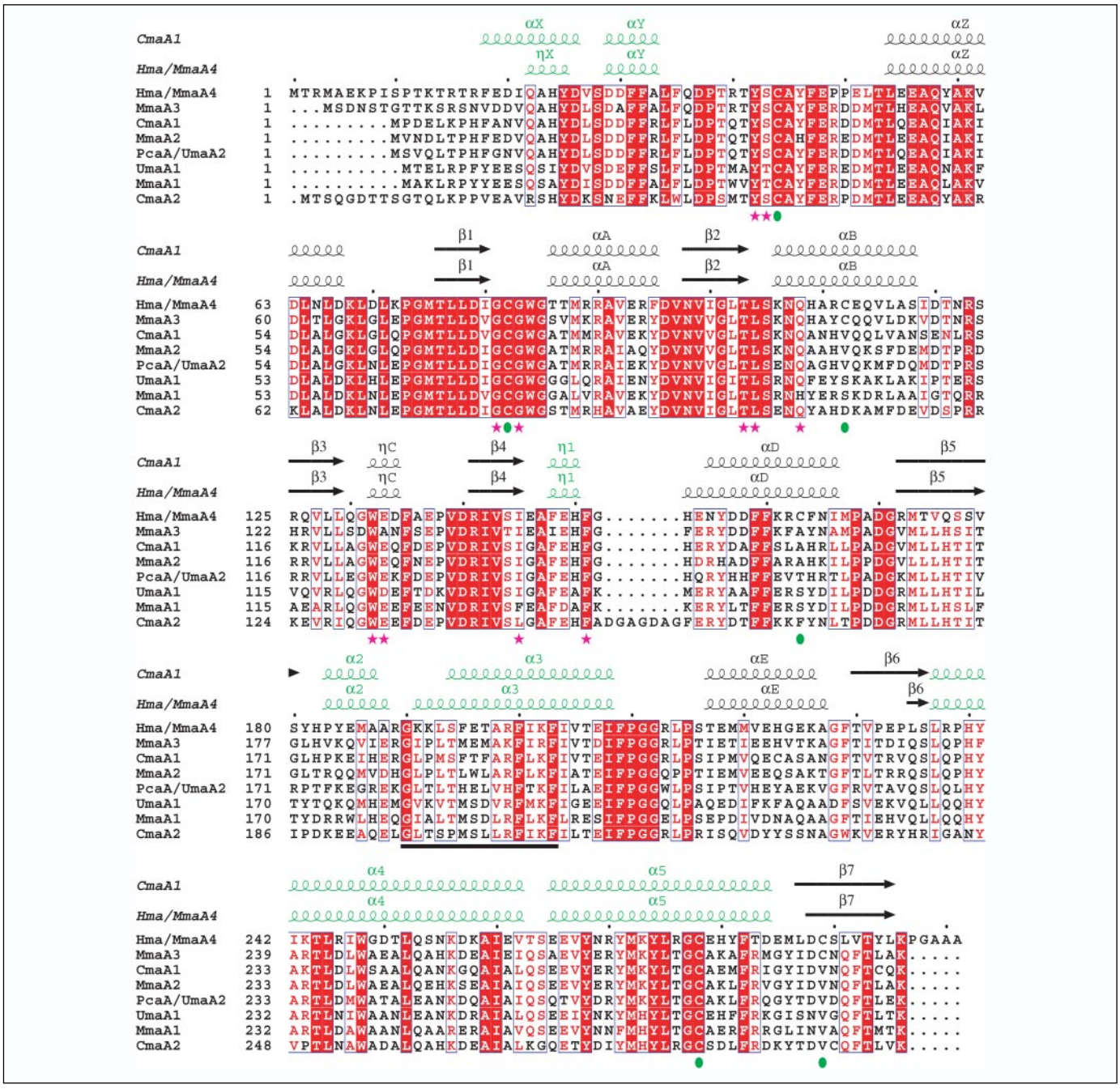

FIGURE 4. **Sequence alignment for *M. tuberculosis* mycolic acid SAM-MTs.** Sequences were displayed from *top* to *bottom* by decreasing order of homology with respect to Hma (percentages of identity/similarity: MmaA3, 59/74; CmaA1, 56/68; MmaA2, 53/65; PcaA, 51/65; UmaA1, 49/66; MmaA1, 48/64; CmaA2, 47/62). Sequence similarities are *highlighted* in *red*, whereas sequence identities are shown as *white letters* on a *red background*. Secondary structure elements (*arrows* for β-strands and *coils* for helices) of Hma in complex with SAM and of CmaA1 in complex with SAH and CTAB are indicated at the *top*. Secondary structure elements that participate to the embellishment pattern are in *green*. *Green circles*, cysteine residues of Hma; *purple stars*, residues that make contact with the cofactor. Residues of Hma that might ultimately be eliminated upon proteolytic cleavage are *underlined*.

Comparison of the Apo-form and Binary Complex of Hma—Superimposition of the apoHma and Hma-SAM structures led to a root mean square deviations (r.m.s.d.) value of 0.9 Å for 277 Cα atoms. In fact, the largest deviations (r.m.s.d. > 1 Å) occur in five peptide segments corresponding to residues 22–29, 105–120, 147–153, 187–192, and 256–267 (Fig. 3*B*). Excluding these segments, the superposition gives an r.m.s.d. value of 0.5 Å for 231 Cα atoms. Residues 22–29 at the N terminus of Hma adopt a slightly different conformation with residues 22–25, forming helix ηX in the structure of the binary complex. Residues 105–120 encompass helix αB, which is translated along its axis toward the cofactor binding cleft upon SAM binding. Residues 187–192 are opposite to the SAM-binding site and define a loop that connects the two helices (α2-α3) inserted between strand β5 and helix αE. Residues 256–267 connect and are part of the two helices (α4-α5) inserted between the C-terminal strands β6 and β7. The segment 147–153 comprises 3 residues (151–153) that could not be traced in the electron density of the apo-form. Furthermore, residues 148–150, whose electron density was clearly interpretable in both structures, undergo a major structural rearrangement since these residues move apart and form the 3_{10} helix η1 upon SAM binding. Among the five peptide stretches that differ in the two structures of Hma, only one is displaced to create space for SAM binding, whereas the four others rather transmit a closing of the structure.

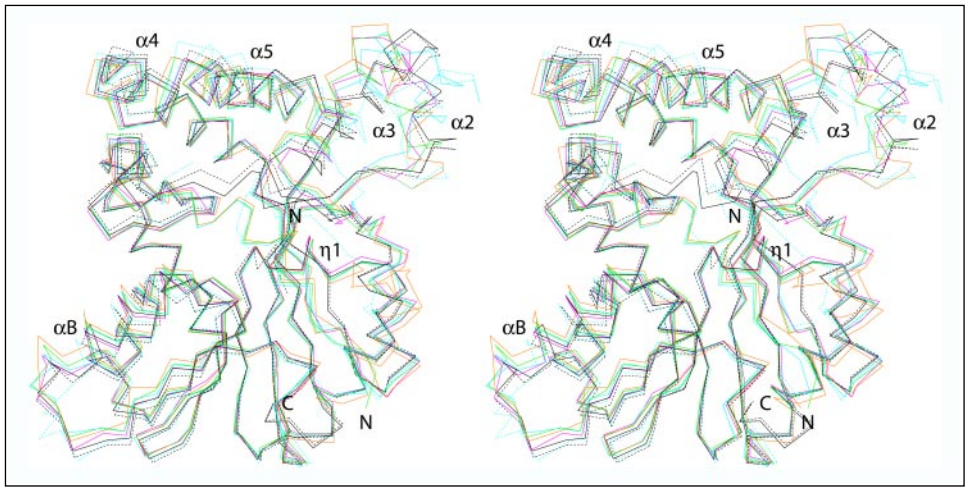

FIGURE 5. **Structural variation among mycolic acid SAM-MTs.** Stereo view of the superimposed α-carbon traces of Hma (*black*), CmaA1 (*cyan*), MmaA2 (*green*), PcaA (*magenta*), and CmaA2 (*orange*). Apo structures are represented by *dotted lines*.

Comparison of all Known Mycolic Acid SAM-MTs Structures—The crystal structures of several SAM-MTs potentially involved in mycolic acid biosynthesis have been reported (31); CmaA1 in its apo-form (PDB code 1KP9) and in complex with the cofactor product SAH and either cetyltrimethylammonium bromide (CTAB) (PDB code 1KPG) or didecyldimethylammonium bromide (DDDMAB) (PDB code 1KPH), the ternary complex of CmaA2 with SAH and DDDMAB (PDB code 1KPI), and the binary complex of PcaA with SAH (PDB code 1L1E). The three-dimensional coordinates of MmaA2 in complex with SAH and CTAB were also deposited in the Protein Data Bank (PDB code 1TPY). These six structures and the two of Hma (this work) were superimposed on one another. Strikingly, a rather complex picture emerged from this overall comparison (Fig. 5) that we have tried to rationalize below. First, it has to be mentioned that three groups of close structures can be distinguished; the structures of both ternary complexes of CmaA1 (r.m.s.d. of 0.5 Å for 283 Cα atoms), the structures of the ternary complexes of CmaA2 and MmaA2 (0.8 Å/276 Cα atoms), and the structures of apoHma and Hma-SAM (0.9 Å/277 Cα atoms). Otherwise, the process of overall comparison lead to r.m.s.d. values between 1.4 and 3.4 Å, the most deviating structure being that of apoCmaA1. We then checked for specific differences and noted that they were confined to a limited number of polypeptide stretches, some of which have been described above when comparing the two structures of Hma. The first set of differences occurs at the N termini, which were systematically found to be shorter in the structures of apo-forms (Hma, CmaA1) or of binary complexes with either the cofactor substrate (Hma-SAM) or the cofactor product (PcaA-SAH) due to disorder, as reflected by the poor electron density for those regions. The second affected region corresponds to helix B and its connecting loops to the preceding β2 and following β3 strands. The largest deviations in this region are typical of Hma, bound to SAM or not, and of apoCmaA1. The third region corresponds to the connecting loop between strand β4 and helix αE, which adopts a variable extended conformation in the apo structures of Hma and CmaA1, whereas it forms a helical turn (η1) conserved in position in all binary and ternary complexes. The fourth region is the longest and most affected region in terms of structural variation. It includes the embellishment pattern corresponding to helix α2, the connecting loop between helices α2 and α3, and to a lesser extent helix α3. The observed discrepancy seems to be rather protein-dependent than related to the different states of a given enzyme. Indeed, the two structures of Hma look nearly similar when compared to such a wealth of conformations observed in the current set of known structures (Fig. 5). However, the large differences observed for CmaA1 when comparing the structures of the apo-form and of the ternary complexes may be an indication that more complicated relationships could exist. The fifth and last deviation is specific to the apo structures, which differ from structures of all complexes in the embellishment at the C-terminal part of α4, the loop connecting α4 and α5, and α5, which is slightly affected.

Cofactor Binding to Mycolic acid SAM-MTs—Soaking of crystals of apoHma in a solution containing both the cryoprotectant and 50 mM SAM for 2–3 min was sufficient to allow binding of the cofactor. Binding of the cofactor substrate to Hma (Fig. 6) occurs in the same position as the one observed for the cofactor product in the structures of the different complexes of CmaA1, CmaA2, PcaA (31), and MmaA2, which were prepared by cocrystallization. The SAM-binding site forms a crevice on the C-terminal side of the central β-sheet and apical to strand β1 (Fig. 3*B*). It is delineated by several protein segments, five of which are directly involved in polar and/or van der Waals interactions with the cofactor. As described above, the segment encompassing residues 147–153 undergoes a structural rearrangement upon SAM binding, and residues 148–150 move apart and refold as helix η1. Contacting residues belong to the connecting loops αY-αZ (Tyr-42 and Ser-43), β1-αA (Gly-81 and Gly-83), β2-αB (Leu-104), and η1-αD (Phe-151), to the C-terminal tips of strands β2 (Thr-103) and β4 (Ile-145), and to helices αB (Gln-108) and ηC (Trp-132 and Glu-133). These residues are conserved in the sequences of other mycolic acid SAM-MTs, except for Glu-133, which is replaced by an alanine in MmaA3 (Fig. 4), and some of them take part in previously identified sequence motifs (31, 42). Furthermore, superimposition of SAM and SAH from all mycolic acid SAM-MT complexes showed that both molecules bind with the same conformation (Fig. 6), leading to r.m.s.d. values of about 0.2 Å for 26 common atoms, *i.e.* excluding the methyl carbon atom of SAM. The surroundings of the cofactor molecules are similar in all structures with one exception. Indeed, formation of ternary complexes, at least using cationic detergents, seems to further sequester the cofactor due to the closer proximity of the 20 first amino acids that wrap around the binding pocket. As a consequence, the side chains of a strictly conserved tyrosine (Tyr-16 in CmaA1) and of a valine (Val-12 in CmaA1, conserved in CmaA2, MmaA2, PcaA; Ile-21 in Hma) are at van der Waals contact of the amino acid portion and adenine ring of the cofactor, respectively. Because in the superimposed structures the hydroxyl group of Tyr-16 of CmaA1-SAH-CTAB is at 2.7 Å of the cofactor methyl group in Hma-SAM (Fig. 6), it is unlikely that the tyrosine residue occupies the same position in the presence of the lipid and SAM substrates because of steric hindrance.

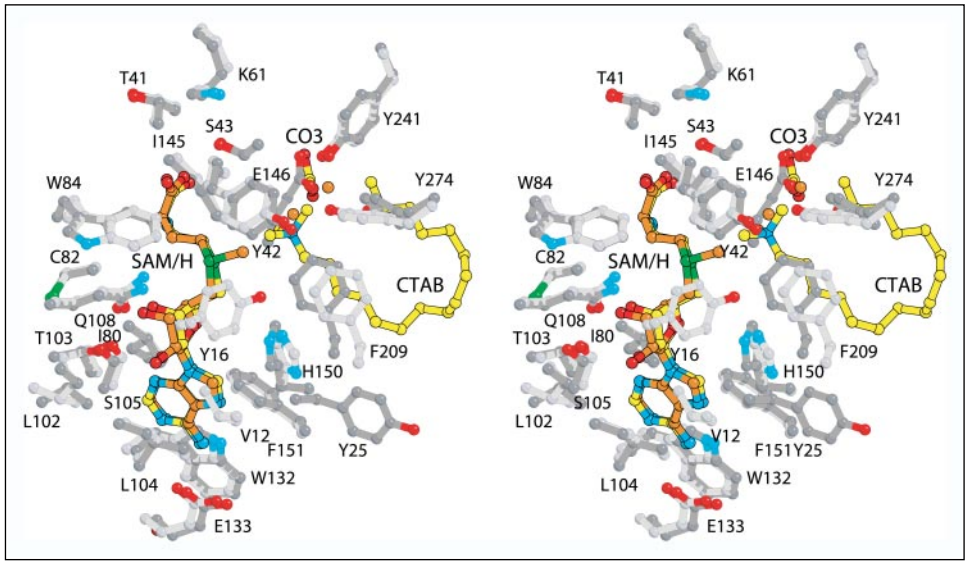

FIGURE 6. **Cofactor binding site and active site architecture of mycolic acid SAM-MTs.** Stereo image of the chemical environment of the cofactor and of a cationic lipid as observed in the structures of Hma-SAM (protein, *dark gray carbon atoms*; SAM, *orange carbon atoms*) and CmaA1-SAH-CTAB (protein, *light gray carbon atoms*; SAH, CTAB, and bicarbonate, *yellow carbon atoms*). Except Val-12 and Tyr-16 of CmaA1, all numbers are for residues of Hma. Water molecules of Hma are represented in *orange*.

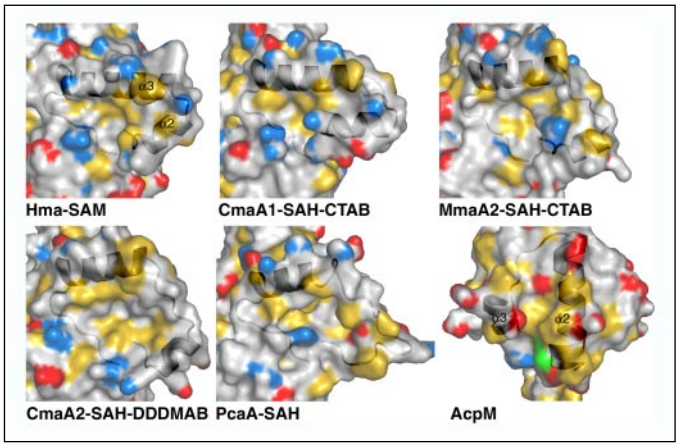

FIGURE 7. **Molecular surface topography of the α2-α3 motif of mycolic acid SAM-MTs and of the recognition motif of AcpM.** All molecular surfaces were displayed using the same scale. Side chain atoms of hydrophobic residues are in *gold*. Side chain nitrogen atoms of basic or acidic residues are in *blue* and *red*, respectively. The serine residue of AcpM that bears the 4′-phosphopantetheine prosthetic group is in *green*. The α2-α3 motif of SAM-MTs and the second and third helices of the AcpM bundle are displayed as *ribbons*.

Interaction with AcpM and Substrate Binding—Mycolic acid SAM-MTs probably exert their action on either full-length meromycolate chains or on shorter intermediates (13, 22). These substrates are likely fueled by the fatty acid synthase II system (43) in the form of acyl-ACP, where the acyl chains are covalently linked to the 4′-phosphopantetheine prosthetic group of the acyl carrier protein from *M. tuberculosis* (AcpM). Thus, in addition to binding and exchanging the cofactor, mycolic acid SAM-MTs have the inherent capability to recognize/interact with AcpM and would host a long acyl chain ($\geq C_{20}$) to perform the reaction. Indirect clues about these specific features have been brought by the structures of the ternary complexes of CmaA1 and CmaA2 with SAH and CTAB/DDDMAB (31). In these structures there is a hydrophobic tunnel extending from the surface to the cofactor binding site. The entrance of the tunnel is delineated by the long segment that comprises residues 180–216 of Hma, encompassing helices α2 and α3, and the extended peptide that connects helices α3 and αE (Fig. 3). In the following, this segment will be referred to as the α2-α3 motif.

The α2-α3 motif of Hma displays basic/hydrophobic patches that would be compatible with its interaction with AcpM, as already observed and discussed for CmaA1/2 and PcaA (31). Indeed, residues 180–216 contain seven basic residues, 5 of which define two regions of solvent-exposed positive charges aligned along α3 and at the bottom of the motif, respectively. These two regions are separated by a shallow hydrophobic cleft with a deeper cavity, the latter providing some kind of vestibule to the hydrophobic tunnel (Fig. 7). Although not identical, these electrostatic features seem to be conserved among all the mycolic acid SAM-MTs of known structures (Fig. 7). On the other hand, the solution structure of AcpM has been solved by NMR spectroscopy (44). It shares the common four-α-helix bundle fold found in other bacterial ACP structures. The second helix of the bundle bears at its N terminus the serine residue to which the prosthetic group is attached and is responsible for the interaction of ACP with its protein partners. This helix as well as the third helix of the bundle define an acidic/hydrophobic patch that could provide the shape complementarity and polar/electrostatic determinants for mycolic acid SAM-MTs molecular recognition and binding (Fig. 7).

In the ternary complexes of CmaA1/2 and MmaA2, the hydrophobic tunnel is filled by the cationic detergents, which adopt a U-shape conformation, with the quaternary ammonium ion rejoining the peptide part of the cofactor. The entrance of the tunnel is closed in both structures of Hma due to the steric obstruction by three residues, Ile-201,Val-205, and Leu-214, whose counterparts in CmaA1 (Leu-192, Val-196, and Leu-205, respectively) establish hydrophobic contacts with the alkyl chain of the detergents (Fig. 8). In the structures of Hma, Ile-201 and Val-205 of helix α3, point in the same direction toward the entrance. Because of the downward position of α3, their side chains are in close vicinity to the loop between α3 and αE and especially to the side chain of Leu-214. Furthermore, the side chains of Ile-201, Val-205, and Leu-214 adopt conformers that maximize hydrophobic contacts, thus minimizing the size of the tunnel aperture. A similar situation is found with the structure of PcaA in complex with the cofactor product, whereas it is the global positioning of α3 and the conformation of residues 137–144 that completely obstruct the tunnel in apoCmaA1.

Mycolic Acid SAM-MTs Specificity—Further comparison of mycolic acid SAM-MTs structures in the α2-α3 specific area revealed that they can be partitioned into three classes. Indeed, looking perpendicular both to the hydrophobic tunnel and to the helix α3 axis leaves α2 nearly in the same plane as α3 in the case of CmaA1 and Hma, below this plane in the case of CmaA2 and MmaA2, and further down in the case of PcaA

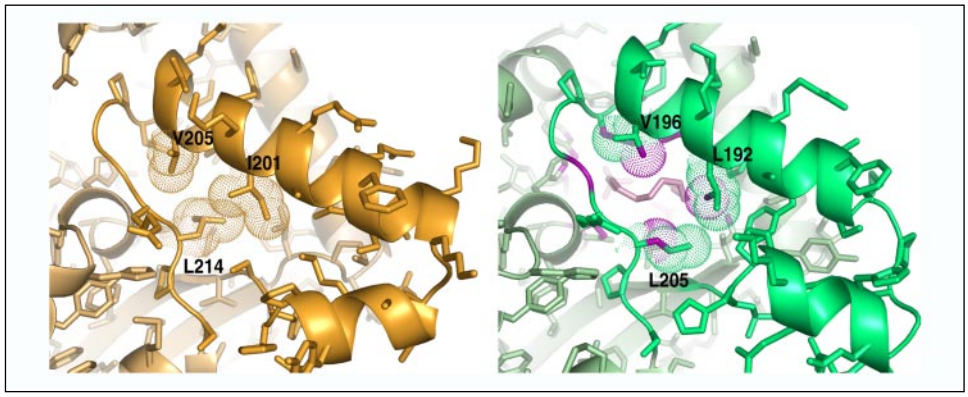

FIGURE 8. **Closed and open state of the mycolic acid SAM-MTs hydrophobic tunnel.** Shown is a Close view inside the α2-α3 motif (*vivid colors*) toward the hydrophobic tunnel comparing the aperture size in Hma-SAM (*left*) and CmaA1-SAH-CTAB (*right*). CTAB is in *pink*, and CmaA1 atoms within 4 Å of the cationic detergent are in *purple*. The hydrophobic residues directly restricting the aperture were labeled. *Dots* represent van der Waals surfaces.

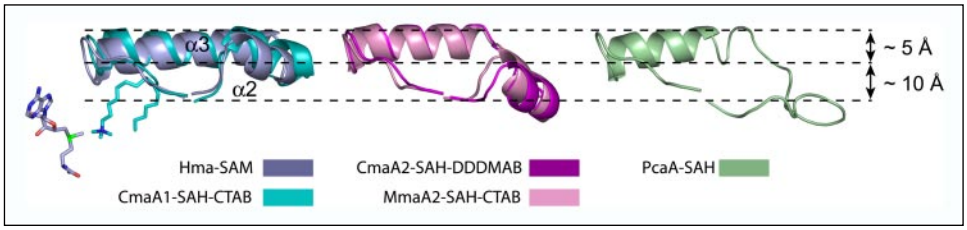

FIGURE 9. **Structural variation in the α2-α3 motif of mycolic acid SAM-MTs.** SAM and CTAB from the structures of Hma-SAM and CmaA1-SAH-CTAB, respectively, have been represented as *sticks* on the *left*. Selected dimensions are indicated.

(Fig. 9). It is worthwhile mentioning that the geometry and the position of the loop β5-α2 and helix α2 are conserved within a class. In addition, the N terminus of helix α3 of Hma is longer by one turn when compared with those of CmaA1, CmaA2, and MmaA2 and by two turns with respect to PcaA (Fig. 9). Therefore, structural variation of helix α3 also participates to the definition of the molecular surface in this area. Because helix α3 of Hma (residues 191–208) contains a proteolytic susceptibility domain where residues 190–203 could be removed upon processing, this might ultimately lead to a severely truncated hydrophobic channel with a substrate binding site open for ready access.

It has been hypothesized that the conformation of mycolic acid SAM-MTs between strand β5 and helix αE, *i.e.* the α2-α3 motif, could be a structural discriminant of proximal *versus* distal specificity by allowing the acyl-AcpM to sit closer/farther from the active site, thereby favoring the reaction at the proximal/distal position (31). Hma is a distal kind of enzyme responsible for the production of methyl-branched hydroxy-mycolic acids (15, 18, 20), whereas PcaA and CmaA2 are required for the introduction of proximal cyclopropane rings in α-(14) and oxygenated mycolates (45, 46), respectively. Furthermore, the function of MmaA2 is dual since it has been shown to be required for introduction of the distal cyclopropane ring in α-mycolates (47) and of the proximal cyclopropane ring in oxygenated mycolates (18, 20). Thus, the structural variation observed in the α2-α3 motif might not only be related to the proximal *versus* distal specificity of mycolic acid SAM-MTs but could also play an important role with respect to their biochemical functions.

DISCUSSION

Mycolic acids are very important components of Mycobacteria, including *M. tuberculosis*. Their structures, which are modulated by chain length and chemical modifications, determine in part the degree of protection of bacilli against the hostile environment of the host. For instance, the deletion of the proximal cyclopropane ring of α-mycolates affects long term persistence in infected mice (14). Furthermore, deletion of keto- and methoxymycolates leads to restricted growth of the corresponding *M. tuberculosis* mutant strain in the mouse model of infection (15). In addition to the severe effects on virulence and pathogenicity, the structural and also the quantitative variations in mycolates may be of crucial biological importance with respect to the permeability of the cell envelope to solutes (15, 48–50). Thus, studying the structure-function relationships of the enzymes involved in the chemical modifications of mycolates deserves special attention. This would help in finding new antitubercular drugs more effective against *M. tuberculosis*, including multiple-drug-resistant strains.

Except CmaA1, which has no discernable role in mycolic acid modification (9, 47), all mycolic acid SAM-MTs three-dimensional structures known so far are those of the cyclopropane synthases CmaA2, MmaA2, and PcaA (or UmaA2). Hma, also known as MmaA4, is unique in that it is the enzyme responsible for the production of the precursors for all oxygenated mycolates in *M. tuberculosis* (15, 20). The methyl-branched hydroxy(mero)mycolates produced by Hma would be transformed either into methoxy(mero)mycolates by the *O*-methyltransferase MmaA3 (51, 52) or into keto(mero)mycolates by an enzyme, probably a dehydrogenase, which has not yet been identified. The precise mechanism by which Hma catalyzes the introduction of the methyl branch and adjacent hydroxyl is also not known. As for the mycolic acid cyclopropane synthases, it has been suggested to proceed through a high energy carbocation intermediate formed upon methyl group addition to olefin precursors (18), and long chain ethylene compounds presumably are the substrates of Hma (19). It has also been proposed that mycolic acid SAM-MTs could catalyze the methylations on a strongly nucleophilic site such as a β-keto ester that might be obtained by the Claisen-type condensation of two esters in an alternative mycolate biosynthesis pathway (53).

With respect to cofactor binding, the structures of Hma reveal that the transition from the apo-form to the binary complex with the SAM substrate involves a single major conformational change, *i.e.* the refolding of residues 147–153 leading to the formation of helix η1, and subtle structural variations of smaller amplitude. The structures of Hma show that helical stabilization of the flexible N-terminal extremities of mycolic acid SAM-MTs, corresponding to the first 15–20 amino acid residues, does not seem to depend on the presence of the cofactor. On the other hand, the available structural data strengthen the hypothesis that the N-terminal extremities of mycolic acid SAM-MTs need to go

through a conformational change for the enzymes to turn over and, more specifically, for the exchange of the product SAH and the substrate SAM as previously suggested (31). Concerning substrate binding and the hydrophobic tunnel, the available structural data indicate that mycolic acid SAM-MTs in their apo-form or in binary complex with SAM/SAH are not prone for direct abstraction of their lipid ligands. Indeed, the hydrophobic tunnel may exist in a closed and an open state. Widening of the tunnel aperture would not depend upon interaction with AcpM, as evidenced from the structures of the ternary complexes of cyclopropane synthases with SAH and CTAB or DDDMAB (31). The entrance of the hydrophobic tunnel is delineated by the α2-α3 motif that contributes to the specific embellishment pattern characteristic of mycolic acid SAM-MTs. This motif displays the largest deviation among mycolic acid SAM-MTs and might play a pivotal role regarding the specific function of each of these enzymes. Indeed, we showed that the precise structure of the α2-α3 motif can be partitioned into three classes that correspond to the following known modifications: (i) in the distal position of oxygenated mycolates, (ii) in the proximal position of oxygenated mycolates and in the distal position of α-mycolates, and (iii) in the proximal position of α-mycolates. The exact functional role played by the α2-α3 motif for such selectivity and specificity might be related to differences in the length of the three polymethylene parts that compose the meromycolic chain according to substituents (53). In addition, it may explain the reported partial redundancies between the activity of the different mycolic acid SAM-MTs (14, 45, 47).

It has been suggested that starting from the high energy cation intermediate, the concomitant displacement of an hydrogen from the incoming methyl group would lead to the formation of a cyclopropane ring, whereas the concerted addition of an hydroxyl group would lead to the formation of oxygenated compounds (18, 54). This would require the presence of a general base that will abstract a proton in the case of cyclopropane synthase and the presence of a residue or water molecule in the active site of Hma, which may facilitate hydroxylation. In the latter case the general base might, for instance, activate a water molecule for in-line nucleophilic attack. Interestingly, a carbonate ion has been found in the active site of CmaA1, CmaA2, MmaA2, and PcaA that could, for example, serve as the general base necessary to complete cyclopropanation reaction (31). The preponderant role played by carbonate in the catalysis of the formation of cyclopropane rings has been demonstrated for the closely related *E. coli* cyclopropane fatty acid synthase (55, 56). Because the closest oxygen atom of the carbonate ion is more than 5 Å away from the methyl carbon atom of SAM, as calculated from the superimposed structures, this would mean that abstraction of the proton from the incoming methyl group is not transient and implies local adjustments in the active site. No carbonate ion was found in Hma, but the carboxylate group of residue Glu-146, located between β4 and η1, occupies exactly the same position (Fig. 6). Thus, Glu-146, which is conserved in MmaA3 (a glycine residue is found in all cyclopropane synthases), may function as a general base (31). Moreover, scrutinizing the active site of Hma allowed the identification of a network of polar interactions involving the side chains of Glu-146 and Tyr-241 on α4 and Tyr-274 on α5 and two water molecules. One water molecule is hydrogen-bonded to the carboxylate of Glu-146, which in turn interacts with the OH group of Tyr-241. The other water molecule is hydrogen-bonded to the OH groups of both Tyr-241 and Tyr-274 (Fig. 6). These two water molecules are located 1.0 and 3.5 Å away, respectively, from the positively charged quaternary nitrogen atom of DDDMAB/CTAB as observed in the ternary complexes. Tyr-241 is conserved among mycolic acid SAM-MTs, except in MmaA3, where a phenylalanine residue is found, whereas Tyr-274 is strictly conserved. Hence, assignment of the general base and water molecule that would be involved in catalysis is not straightforward and will require further enzymatic and structural studies.

Finally, the propensity of Hma to undergo proteolytic degradation *in vitro* is intriguing. This is to our knowledge the first report of such an instability for a mycolic acid SAM-MT. There is no evidence that such a processing of the enzyme would also occur *in vivo* and if it might be ultimately related to a regulatory or a metabolic role. However, it might be reminiscent of what has been reported for *E. coli* cyclopropane fatty acid synthase whose metabolic instability is responsible for the loss of activity (57).

Acknowledgments—We are grateful to Drs. Eugenie Dubnau (PHRI, Newark, NJ) and Marie-Antoinette Lanéelle (Institut de Pharmacologie et de Biologie Structurale du CNRS (IPBS)) for help and for fruitful discussions. We thank Dr. Stéphanie Ducasse Cabanot (IPBS) for precious help during the early stage of this work and Dr. Isabelle Saves (IPBS) for help in cloning the hma gene. We thank the scientific staff of European Synchrotron Radiation Facility (Grenoble, France) for excellent data collection facilities.

REFERENCES

1. Daffé, M., and Draper, P. (1998) *Adv. Microb. Physiol.* **39,** 131–203
2. Goren, M. B., and Brennan, P. J. (1979) in *Tuberculosis* (Youmans, G. P., ed) pp. 63–193, W. B. Saunders Co., Philadelphia, PA
3. Brennan, P. J., and Nikaido, H. (1995) *Annu. Rev. Biochem.* **64,** 29–63
4. Winder, F. G., and Collins, P. B. (1970) *J. Gen. Microbiol.* **63,** 41–48
5. Takayama, K., Wang, L., and David, H. L. (1972) *Antimicrob. Agents Chemother.* **2,** 29–35
6. Davidson, L. A., and Takayama, K. (1979) *Antimicrob. Agents Chemother.* **16,** 104–105
7. Quémard, A., Lacave, C., and Lanéelle, G. (1991) *Antimicrob. Agents Chemother.* **35,** 1035–1039
8. Rozwarski, D. A., Grant, G. A., Barton, D. H., Jacobs, W. R., Jr., and Sacchettini, J. C. (1998) *Science* **279,** 98–102
9. Takayama, K., Wang, C., and Besra, G. S. (2005) *Clin. Microbiol. Rev.* **18,** 81–101
10. Daffé, M., Lanéelle, M. A., Asselineau, C., Levy-Frebault, V., and David, H. (1983) *Ann. Microbiol.* (*Paris*) **134,** 241–256
11. Minnikin, D. E., Minnikin, S. M., Parlett, J. H., Goodfellow, M., and Magnusson, M. (1984) *Arch. Microbiol.* **139,** 225–231
12. Laval, F., Lanéelle, M. A., Deon, C., Monsarrat, B., and Daffé, M. (2001) *Anal. Chem.* **73,** 4537–4544
13. Yuan, Y., Zhu, Y., Crane, D. D., and Barry, C. E., III (1998) *Mol. Microbiol.* **29,** 1449–1458
14. Glickman, M. S., Cox, J. S., and Jacobs, W. R., Jr. (2000) *Mol. Cell* **5,** 717–727
15. Dubnau, E., Chan, J., Raynaud, C., Mohan, V. P., Lanéelle, M. A., Yu, K., Quémard, A., Smith, I., and Daffé, M. (2000) *Mol. Microbiol.* **36,** 630–637
16. Rao, V., Fujiwara, N., Porcelli, S. A., and Glickman, M. S. (2005) *J. Exp. Med.* **201,** 535–543
17. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., III, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., and Barrell, B. G. (1998) *Nature* **393,** 537–544
18. Yuan, Y., and Barry, C. E., III (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 12828–12833
19. Dinadayala, P., Laval, F., Raynaud, C., Lemassu, A., Lanéelle, M. A., Lanéelle, G., and Daffé, M. (2003) *J. Biol. Chem.* **278,** 7310–7319
20. Dubnau, E., Lanéelle, M. A., Soares, S., Benichou, A., Vaz, T., Prome, D., Prome, J. C., Daffé, M., and Quémard, A. (1997) *Mol. Microbiol.* **23,** 313–322
21. Quémard, A., Lanéelle, M. A., Marrakchi, H., Prome, D., Dubnau, E., and Daffé, M. (1997) *Eur. J. Biochem.* **250,** 758–763
22. Qureshi, N., Sathyamoorthy, N., and Takayama, K. (1984) *J. Bacteriol.* **157,** 46–52
23. Stover, C. K., de la Cruz, V. F., Fuerst, T. R., Burlein, J. E., Benson, L. A., Bennett, L. T., Bansal, G. P., Young, J. F., Lee, M. H., Hatfull, G. F., Snapper, S. B., Barletta, R. G., Jacobs, W. R., Jr., and Bloom, B. R. (1991) *Nature* **351,** 456–460
24. Kurth, J., and Stoffel, W. (1990) *Biol. Chem. Hoppe-Seyler* **371,** 675–685
25. Claverol, S., Burlet-Schiltz, O., Gairin, J. E., and Monsarrat, B. (2003) *Mol. Cell. Proteomics* **2,** 483–493
26. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. D Biol. Crystallogr.* **50,** 760–763
27. Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) *Acta Crystallogr. D Biol. Crystallogr.* **59,** 1131–1137

28. Leslie, A. G. W. (1987) in *Proceedings of the Daresbury Study Weekend: Computational Aspects of Protein Crystal Data Analysis* (Helliwell, J. R., Machin, P. A., and Papiz, M. Z., eds) pp. 39–50, Science and Engineering Research Council, Daresbury Laboratory, Warrington, United Kingdom
29. Evans, P. R. (1993) in *Proceedings of the CCP4 Study Weekend: Data Collection and Processing*, pp. 114–122, Science and Engineering Research Council, Daresbury Laboratory, Warrington, United Kingdom
30. Storoni, L. C., McCoy, A. J., and Read, R. J. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60,** 432–438
31. Huang, C. C., Smith, C. V., Glickman, M. S., Jacobs, W. R., Jr., and Sacchettini, J. C. (2002) *J. Biol. Chem.* **277,** 11559–11569
32. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6,** 458–463
33. Read, R. J. (1986) *Acta Crystallogr. A* **42,** 140–149
34. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D. Biol. Crystallogr.* **53,** 240–255
35. Perrakis, A., Sixma, T. K., Wilson, K. S., and Lamzin, V. S. (1997) *Acta Crystallogr. D Biol. Crystallogr.* **53,** 448–455
36. Esnouf, R. M. (1997) *J. Mol. Graph.* **15,** 133–138
37. Merritt, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. D Biol. Crystallogr.* **50,** 869–873
38. Gouet, P., Courcelle, E., Stuart, D., and Metoz, F. (1998) *Bioinformatics* **15,** 305–308
39. Frishman, D., and Argos, P. (1995) *Proteins* **23,** 566–579
40. Rawlings, N. D., Tolle, D. P., and Barrett, A. J. (2004) *Nucleic Acids Res.* **32,** 160–164
41. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
42. Martin, J. L., and McMillan, F. M. (2002) *Curr. Opin. Struct. Biol.* **12,** 783–793
43. Bloch, K. (1977) *Adv. Enzymol. Relat. Areas Mol. Biol.* **45,** 1–84
44. Wong, H. C., Liu, G., Zhang, Y. M., Rock, C. O., and Zheng, J. (2002) *J. Biol. Chem.* **277,** 15874–15880
45. Glickman, M. S., Cahill, S. M., and Jacobs, W. R., Jr. (2001) *J. Biol. Chem.* **276,** 2228–2233
46. George, K. M., Yuan, Y., Sherman, D. R., and Barry, C. E., III (1995) *J. Biol. Chem.* **270,** 27292–27298
47. Glickman, M. S. (2003) *J. Biol. Chem.* **278,** 7844–7849
48. Gao, L. Y., Laval, F., Lawson, E. H., Groger, R. K., Woodruff, A., Morisaki, J. H., Cox, J. S., Daffé, M., and Brown, E. J. (2003) *Mol. Microbiol.* **49,** 1547–1563
49. Liu, J., Barry, C. E., III, Besra, G. S., and Nikaido, H. (1996) *J. Biol. Chem.* **271,** 29545–29551
50. Jackson, M., Raynaud, C., Lanéelle, M. A., Guilhot, C., Laurent-Winter, C., Ensergueix, D., Gicquel, B., and Daffé, M. (1999) *Mol. Microbiol.* **31,** 1573–1587
51. Behr, M. A., Schroeder, B. G., Brinkman, J. N., Slayden, R. A., and Barry, C. E., III (2000) *J. Bacteriol.* **182,** 3394–3399
52. Dubnau, E., Marrakchi, H., Smith, I., Daffé, M., and Quémard, A. (1998) *Mol. Microbiol.* **29,** 1526–1528
53. Asselineau, C., Asselineau, J., Lanéelle, G., and Lanéelle, M. A. (2002) *Prog. Lipid Res.* **41,** 501–523
54. Grogan, D. W., and Cronan, J. E., Jr. (1997) *Microbiol. Mol. Biol. Rev.* **61,** 429–441
55. Iwig, D. F., Uchida, A., Stromberg, J. A., and Booker, S. J. (2005) *J. Am. Chem. Soc.* **127,** 11612–11613
56. Courtois, F., and Ploux, O. (2005) *Biochemistry* **44,** 13583–13590
57. Chang, Y. Y., Eichel, J., and Cronan, J. E., Jr. (2000) *J. Bacteriol.* **182,** 4288–4294
58. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA