Molecular Cloning and Sequencing of the Gene for Mycocerosic Acid Synthase, a Novel Fatty Acid Elongating Multifunctional Enzyme, from *Mycobacterium tuberculosis* var. *bovis* Bacillus Calmette-Guerin*†

(Received for publication, May 22, 1992)

Manjula Mathur and Pappachan E. Kolattukudy‡

From The Ohio State Biotechnology Center, The Ohio State University, Columbus, Ohio 43210

Mycocerosyl lipids are found uniquely in the cell walls of pathogenic mycobacteria. Mycocerosic acid synthase (MAS) is a multifunctional protein which catalyzes elongation of *n*-fatty acyl-CoA with methylmalonyl-CoA as the elongating agent (Rainwater, D. L., and Kolattukudy, P. E. (1985) *J. Biol. Chem.* 260, 616–623). To understand how the various domains that catalyze the reactions involved in chain elongation are organized, *mas* gene from *Mycobacterium tuberculosis* BCG was cloned. A Xgt11 library of *Alu*I partially digested genomic DNA from the organism was screened with an oligonucleotide probe designed from the N-terminal amino acid sequence of purified MAS. Using terminal segments of inserts from positive clones as the probe, the library was rescreened and the process was repeated. Sequencing of four overlapping clones revealed a contiguous sequence of 9699 base pairs (bp) of mycobacterial genome containing a 6330-bp open reading frame that could code for a protein of 2100 amino acids with a molecular mass of 225,437 daltons. The authenticity of the open reading frame as that of MAS was verified by correspondence of the amino acid sequences deduced from the gene with the directly determined amino acid sequences of the N terminus and three different internal peptide fragments. By comparing the MAS amino acid sequence with the sequences in the active site regions of known fatty acid synthases and polyketide synthases the functional domains in MAS were identified. This analysis showed that the domains were organized in the following order: β-ketoacyl synthase, acyl transferase, dehydratase-enoyl reductase, β-keto reductase, acyl carrier protein; no thioesterase-like domain could be found. These results establish MAS as the first case of an elongating multifunctional enzyme composed of two identical subunits that resemble the vertebrate fatty acid synthase in size, subunit structure, and linear organization of functional domains. Southern and Western blot analyses showed absence of *mas* gene and encoded proteins in *Mycobacterium smegmatis* and *Escherichia coli*. This result is consistent with the report that mycocerosic acid is present only in pathogenic mycobacteria.

Multimethyl-branched fatty acids occur in a variety of biological systems. The aglycone portion of macrocyclic lactone antibiotics such as erythromycin generated by Streptomyces (1) and multimethyl-branched fatty acids found in sebaceous glands (2, 3) are examples of such natural products. The enzymology and regulation of their biosynthesis is not well understood except in the case of the goose uropygial gland which generates 2,4,6,8-tetramethyl decanoic acid (4). In this system, an acyl-CoA carboxylase generates malonyl-CoA and methylmalonyl-CoA (5) but the former is converted to acetyl-CoA by a cytoplasmic decarboxylase present only in the gland (4, 6). This ensures that methylmalonyl-CoA would be the only substrate available for fatty acid synthase which is capable of synthesizing multimethyl-branched fatty acids (7).

*Mycobacterium tuberculosis* currently afflicts 30 million people causing 3 million deaths per year (8, 9) and *Mycobacterium leprae* causes leprosy that afflicts 10–12 million people (10, 11). Even though tuberculosis was thought to have been nearly eradicated from the United States, recently there has been a resurgence of tuberculosis, and increasing incidences of tuberculosis among patients with acquired immunodeficiency syndrome (AIDS) has been noted (12). These pathogenic mycobacteria contain a variety of unique fatty acids which have one to six methyl branches, at even-numbered positions at the carboxyl end, and a long n-aliphatic chain (13, 14). One such group of acids, called mycocerosic acids, is found exclusively esterified to phenolphthiocerol (15) in the cell wall of only pathogenic mycobacteria (16). A cell-free system which catalyzes the synthesis of mycocerosic acids was obtained from *M. tuberculosis* bovis BCG (17). Mycocerosic acid synthase (MAS), a multifunctional enzyme, which catalyzes the elongation of *n*-fatty acyl-CoA specifically using methylmalonyl-CoA (but not malonyl-CoA) to produce primarily the corresponding tetramethyl-branched mycocerosic acids was purified and characterized (18). It was found to be a dimer of 238-kDa protomers each with an acyl carrier protein like segment. Thus, this enzyme which elongates *n*-fatty acyl moieties with methylmalonyl-CoA is a multifunctional enzyme similar to vertebrate fatty acid synthase in size and subunit structure (19). To achieve elongation the dimeric

1 The abbreviations used are: AIDS, acquired immunodeficiency syndrome; *M. tuberculosis* bovis BCG, Bacillus Calmette-Guerin vaccine of *Mycobacterium tuberculosis* var. bovis; MAS, mycocerosic acid synthase; FAS, fatty acid synthase; ORF, open reading frame; AT, acyl transferase; bp, base pair(s); kb, kilobase(s); SDS, sodium dodecyl sulfate.

---

* This work was supported by Grant GM-18278 from the National Institutes of Health. 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.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) M95808.

‡ To whom correspondence should be addressed: The Ohio State Biotechnology Center, The Ohio State University, 206 Rightmire Hall, 1060 Carmack Rd., Columbus, OH 43210. Tel.: 614-292-5670; Fax: 614-292-5379.
enzyme must catalyze transacylation steps, condensation, keto reduction, dehydration, and enoyl reduction. How the domains that catalyze the various steps involved in this enzyme must catalyze transacylation steps, condensation, unique elongation system are arranged is not known.

In the present paper, we describe the molecular cloning and sequence of the mycocerosic acid synthase gene (mas) and present evidence that the domains are organized in this multifunctional enzyme in the following order: ketoacyl synthase acyl transferase, dehydratase, enoyl reductase, β-ketoreductase, acyl carrier protein. This is the first case of structure elucidation of a fatty acid chain elongating enzyme.

**Experimental Procedures**

**Materials**—Escherichia coli strains Y1090 and DH5αF were purchased from Stratagene and Bethesda Research Laboratories, respectively. TICE Bacillus Calmette-Guerin vaccine of M. tuberculosis var. bovis was obtained and cultivated as previously described (23). A size-selected AluI partial library was made in Xgtll. The genomic DNA was partially digested with AluI. DNA fragments with length of 2-7 kb were isolated and ligated to EcoRI linker-adapter. Western blot was done in an Applied Biosystems Model 473 protein sequencer. The purified MAS was cleaved with Staphylococcus aureus V8 protease (20). The resulting peptides were loaded on a 3% stacking and 4-15% resolving sodium dodecyl sulfate-polyacrylamide gel electrophoresis gel. After electrophoresis the proteins were electroblotted on polyvinylidene difluoride membrane (Immobilon-P, Millipore), stained with Coomassie Blue, destained, rinsed with water, and air-dried (21). Membrane segments representing three internal peptides were cut out, and the peptides were sequenced in an Applied Biosystem sequencer.

**Amino acid Sequence Analysis**—Mycocerosic acid synthase was purified as described below (18). The N-terminal sequencing was done in an Applied Biosystems Model 473 protein sequencer. The purified MAS was cleaved with Staphylococcus aureus V8 protease (20). The resulting peptides were loaded on a 3% stacking and 4-15% resolving sodium dodecyl sulfate-polyacrylamide gel electrophoresis gel. After electrophoresis the proteins were electroblotted on polyvinylidene difluoride membrane (Immobilon-P, Millipore), stained with Coomassie Blue, destained, rinsed with water, and air-dried (21). Membrane segments representing three internal peptides were cut out, and the peptides were sequenced in an Applied Biosystem sequencer.

**Preparation of Genomic DNA Library**—TICE BCG vaccine of M. tuberculosis bovis was obtained and cultivated as previously described (17). Genomic DNA from bacteria was isolated as described before (23). A size-selected AluI partial library was made in λgt11. The genomic DNA was partially digested with AluI. DNA fragments with length of 2-7 kb were isolated and ligated to EcoRI linker-adapter (according to the manufacturer’s instructions, Invitrogen), and the products were ligated with λgt11 arms and packaged using an in vitro Gigapack 11 plus packaging extract (as described by the manufacturer. Stratagene). The titer of the unamplified λgt11 genomic DNA library was 2 × 10^10 recombinant phages, which was then amplified in strain Y1090.

**Screening of the Genomic DNA Library**—A synthetic oligonucleotide, 5′-GC(TGACGTGGCATACTG/GGCTATGTTGGG-3′ with 192 redundancy, corresponding to Ala-9 through Gly-14, from the N-terminal sequence of MAS, end-labeled with [γ-32P]ATP and T4 polynucleotide kinase, was used to screen the λgt11 genomic DNA library of M. tuberculosis bovis BCG (24-26). Prehybridization conditions were 6 × SSPE (1 × SSPE: 180 mM NaCl, 1 mM EDTA, 10 mM NaH2PO4.H2O, 10 mM Na2HPO4) and 1 × Denhardt's, 0.5% SDS, 100 mg/ml denatured salmon sperm DNA, 65°C sodium pyrophosphate at 20°C, 10% SDS, 30% formamide, 0.5% sodium pyrophosphate, 30 min at room temperature, 1 × SSPE, 0.1% SDS at 20°C, and 0.2 × SSPE, 0.1% SDS, 15 min at 45°C.

**Southern Blot Analysis**—The EcoRI-digested lambda clones from λMAS1, λMAS2, λMAS3, and λMAS4 were subcloned into M13 mp18. Single stranded DNA was sequentially deleted with T4 DNA polymerase generating series of overlapping clones (24). Using the universal M13 sequencing primer as well as various additional, specifically synthesized oligonucleotide primers, the M13 mp18 clones were sequenced by the Sanger dideoxy chain termination method using [α-32P]dATP as label.

**RESULTS**

**Isolation and Sequencing of Genomic Clones**—To understand the catalytic domains of the multifunctional enzyme, mycocerosic acid synthase, we cloned and sequenced the gene for this enzyme from M. tuberculosis bovis BCG. To prepare oligonucleotide probes for detecting mas gene, the purified enzyme was subjected to N-terminal sequencing. Unlike multifunctional fatty acid synthases such as those from vertebrates (19) and M. tuberculosis bovis BCG (31) in which the N terminus is blocked, N terminus of MAS is not blocked and consequently yielded the sequence of 20 amino acid residues: H1-NESRVTVPAGMCRLPGGCOOH. Based on this sequence an oligonucleotide corresponding to Ala-9 through Gly-14 was synthesized. When a λgt11 size-selected (2-7 kb) AluI partial genomic library of M. tuberculosis bovis BCG was screened with the labeled oligonucleotide probe, two positive clones, λMAS1 and λMAS2, were obtained. EcoRI digestion of λMAS1 and λMAS2 each gave one EcoRI insert of 3.2 and 2.5 kb, respectively. These EcoRI fragments were subcloned in M13 mp18 and sequenced. Upon sequencing the two clones were found to be overlapping by 1988 bp, and the sequence of λMAS2 gave a 3′-extension of 555 bp; thus a total sequence of 3758 bp was obtained from λMAS1 and λMAS2. A 724-bp HindIII-EcoRI fragment from 3′-end of λMAS2 was used as a probe (probe B, Fig. 1) to screen the genomic library. A positive clone λMAS3, which contained a 4-kb insert, was obtained which gave four EcoRI fragments. These EcoRI fragments yielded sequences of 1918, 1297, 499, and 381 bp. A total sequence of 4077 bp was obtained from λMAS3 which was found to be overlapping with λMAS2 by 499 bp and gave an extension of 3578 bp. A 1297-bp EcoRI fragment from 3′-end of λMAS3 (probe D, Fig. 1) was then used as a probe to screen the genomic library. A positive clone λMAS4, containing a 2.8-kb insert was obtained which gave two EcoRI fragments. These two EcoRI fragments gave a sequence of 1238 and 1621 bp yielding a total sequence of 2853 bp from λMAS3 with an overlap of 400 bp with λMAS3 and a 3′-extension of 2453 bp. Both the strands of all four genomic fragments were sequenced. The nucleotide sequence of these four overlapping clones reveals a contiguous sequence of 9699 bp of the mycobacterial genome. The complete DNA sequence and deduced amino acid sequences are shown in Fig. 2.
Fig. 1. Restriction map of the three segments of cloned genomic DNA and restriction fragments used for sequencing mycoerosic acid synthase gene from M. tuberculosis bovis BCG. The solid box shows the MAS coding region. The solid lines indicate Agt11 clones λM1, λM2, λM3, and λM4, and double-headed arrows under the lambda clones represent M13 mp18 clones. All genomic sequences are connected to vector sequences through EcoRI linkers. Probes A, B, C, and D are the gene segments used for hybridization, B, BamHI; E, EcoRI; H, HindIII; P, PstI; *S, SacI; S, Smal.

---

Restriction map of the three segments of cloned genomic DNA and restriction fragments used for sequencing mycoerosic acid synthase gene from M. tuberculosis bovis BCG. The solid box shows the MAS coding region. The solid lines indicate Agt11 clones λM1, λM2, λM3, and λM4, and double-headed arrows under the lambda clones represent M13 mp18 clones. All genomic sequences are connected to vector sequences through EcoRI linkers. Probes A, B, C, and D are the gene segments used for hybridization, B, BamHI; E, EcoRI; H, HindIII; P, PstI; *S, SacI; S, Smal.
The deduced amino acid sequences (single letter code) for MAS and ORFII are beneath the nucleotide sequence. The numbers indicate the different regions.

**FIG. 2.** Nucleotide sequence of mas gene and its flanking regions. The deduced amino acid sequences (single letter code) for MAS and ORFII are beneath the nucleotide sequence. The numbers indicate the different regions.
The nucleotide sequence of the contiguous segment of 9699 bp of mycobacterial genome revealed an open reading frame of 6330 bp long with a coding capacity of 2110 amino acids. The ORF starts with an ATG at nucleotide position 1741 and ends with a TAG stop codon at position 8073. Codon usage in the mas gene indicates that a remarkably wide range of codons are used; 61 out of 64 codons are used in the synthesis of the mas gene product. Southern and Western Blot Analyses—Genomic DNA prepared from M. tuberculosis bouis BCG and rabbit antiserum prepared against MAS purified from M. tuberculosis bouis were used with 32P-labeled protein A for detection.
matis is known to be quite similar to that in *M. tuberculosis bovis* (31), and if *M. smegmatis* had a MAS it probably would have been immunologically similar to MAS from *M. tuberculosis* bovis. Thus, the results presented here are consistent with the report that mycocerosic acids, unique products of *M. tuberculosis* bouis. Thus, the results presented here are consistent with the report that mycocerosic acids, unique products of mycobacteria, are not found in *E. coli* and *M. smegmatis* (16).

Identification of Catalytic Domains in MAS—MAS is a unique fatty acid-elongating enzyme which uses methyloannyl-CoA for elongation of α-fatty acids to produce multi-methyl-branched very long chain fatty acids. The various functional domains required for synthesis of mycocerosic acid are: acyl transferase, β-ketoacyl synthase, β-ketoreductase, dehydratase, enoyl reductase, and acyl carrier protein. Such reactions are catalyzed by the multifunctional fatty acid synthases, and the amino acid sequence motifs surrounding the substrate binding sites for the various domains of such synthases have been identified (33–40). Recent studies revealed that the polypeptide synthase from *Streptomyces erythraeus*, *eryA*, also contains domains that can be identified by the presence of the amino acid sequence motifs similar to those found in the corresponding domains of fatty acid synthases (34, 41). Comparison of the amino acid sequence of MAS with the known amino acid sequence motifs of the domains of fatty acid synthases of a number of eucaryotes and prokaryotes (19, 42) and of *eryA* helped to identify the domains in MAS.

Comparison of sequence between the ketoacyl synthase active sites (Table I) in MAS with *eryA* shows 93% identity, whereas the corresponding regions in FAS from chicken, goose, rat, *S. cerevisiae*, and *E. coli* showed only 71, 71, 84, 29, and 21% identity, respectively. The highly conserved sequence GPXXXTXCTCSS (43) around the active cysteine residue of ketoacyl synthase domain of FAS and polyketide synthase that participates in thioester formation can be detected in MAS (Gly-168 to Ser-179). Table II shows that the amino acid sequence surrounding the essential serine of acyl transferase shows maximum identity with *S. erythraeus* *eryA* (80%) followed by rat FAS (45%) and chicken FAS (40%). The highly conserved GHSXG motif (41, 44) of the acyl transferase domain, where S is serine, involved in the formation of acyl-cysteine intermediate, is present in the MAS (Gly-621 to Gly-625). Table III compares the amino acid sequences around the NADPH binding site of enoyl reductase and shows that 89, 83, and 67% amino acids of MAS are identical to the corresponding regions of fatty acid synthase of rat, chicken, and *S. cerevisiae* *eryA*, respectively. Comparison of amino acids around the NADPH binding site of ketoreductase shows that 81 and 56% amino acids of MAS are identical to the corresponding regions of vertebrate FAS and polyketide synthase from *S. erythraeus*, respectively (Table IV). Common to the NADPH binding domains of many enzymes is a β-α-β-fold, centered around a highly conserved sequence Gly-X-Gly-X-X-Gly (where X is any amino acid) that constitutes a tight turn at the end of the first strand of a β-sheet and marks the beginning of the succeeding α-helix (45). This fingerprint region GXXGXXXXXXA of NADPH-dependent reductases found in the enoyl reductase and ketoreductase domains can be allocated to two positions in the MAS sequence: Gly-1568 to Ala-1578 and Gly-1773 to Ala-1784. Table V shows that 54% of the amino acid sequence around the pantetheine-binding serine of acyl carrier protein of MAS is identical to the corresponding regions of FAS of chicken, rat, and goose; 46, 38, and 23% identity were found for the corresponding regions of *module IV in ORF 2 of S. erythraea* *eryA*, FAS of *E. coli*, and FAS of *S. cerevisiae*, respectively. The pantetheine-binding serine present in acyl carrier proteins in the GLDSLXXXE motif (46) is found at position Gly-2056 to G1u-2064.

A thioesterase-like domain was not found in the MAS open reading frame. Adjacent to MAS coding region, three open reading frames were found. In the 5′-end two overlapping open reading frames 1 and 3 were found, and in the 3′-end open reading frame 2 was found (Fig. 2).

**DISCUSSION**

The comparison of the amino acid sequence of MAS with that of polyketide synthase and fatty acid synthase from various systems (32, 34–40) shows that MAS contains active site motifs of ketoacyl synthase, acyl transferase, enoyl reductase, and ketoreductase domains. The sequences are aligned to the essential cysteine site of β-ketoacyl synthase.

**Table I**

| β-Ketoacyl synthase | Amino acid sequences | Ref. |
|---------------------|---------------------|-----|
| *M. tuberculosis bovis* MAS | GPAMFDTACSSGL | This study |
| *S. erythraea* *eryA* | GPAMFDTACSSGL | 34 |
| Chicken fatty acid synthase | GPBSLTDACSSGL | 35 |
| Rat fatty acid synthase | GPBSLTDACSSGL | 36 |
| Goat fatty acid synthase | GPBSLTDACSSGL | 42 |
| *S. cerevisiae* fatty acid synthase | GPBSLTDACSSGL | 37 |
| *E. coli* fatty acid synthase | GVMY9ISSAATSA | 38 |
Molecular Cloning and Sequencing of Mycocerosic Acid Synthase

TABLE II
Comparison of the amino acid sequences in the region of essential serine site of acyl transferase

| Acyl transferase          | Amino acid sequences          | Ref.  |
|--------------------------|-------------------------------|-------|
| M. tuberculosis bovis BCG MAS | AVVHGGSGESAAAVAGALS          | This study |
| S. erythraea eryA         | AVVHGSGVGGEAASVAGALS          | 34    |
| Rat fatty acid synthase   |                               | 36    |
| Chicken fatty acid synthase | G1LIGSVEAVCGYAGDGLS       | 35    |
|                          |                               |       |

TABLE III
Comparison of the amino acid sequences in the NADPH binding region of enoyl reductase

| Enoyl reductase          | Amino acid sequences          | Ref.  |
|--------------------------|-------------------------------|-------|
| M. tuberculosis bovis BCG MAS | VLIHSATGVQQVAASIA            | This study |
| S. erythraea eryA         | VLIHAAAGGGVMAAVA            | 34    |
| Rat fatty acid synthase   |                               | 36    |
| Chicken fatty acid synthase | VIHSOSGQGQQVAISA       | 39    |

TABLE IV
Comparison of the amino acid sequences in the NADPH binding region of β-keto reductase

| β-Keto reductase          | Amino acid sequences          | Ref.  |
|--------------------------|-------------------------------|-------|
| M. tuberculosis bovis BCG MAS | SYIIITGGLGSLGFELA          | This study |
| S. erythraea eryA         | TYIIITGGLGSGGELA            | 34    |
| Rat fatty acid synthase   | SYIIITGGLGFLGELA            | 36    |
| Chicken fatty acid synthase | SYIIITGGLGGFLGELA       | 39    |

TABLE V
Comparison of the amino acid sequences in the region of the essential serine site of acyl carrier protein

| Acyl carrier protein     | Amino acid sequences          | Ref.  |
|--------------------------|-------------------------------|-------|
| M. tuberculosis bovis BCG MAS | EYGLDISLGLMELR       | This study |
| S. erythraea eryA         | DLGFDISLAAVEL4            | 34    |
| Chicken fatty acid synthase | DLGDISLGMGEVR        | 35    |
| Rat fatty acid synthase   | DLGDISLGMGEVR            | 36    |
| Goose fatty acid synthase | DLGDISLGMGEVR           | 42    |
| S. cerevisiae fatty acid synthase | LVVGKSTVNL           | 37    |
| E. coli fatty acid synthase | GDADISLTVELV          | 49    |

Thus mycocerosic acid synthase is a multifunctional protein in which all the functional domains lie on single polypeptide chain as in vertebrate FAS. The linear arrangement of the various domains in MAS does not correspond to the fatty acid synthase reaction sequence as in the vertebrate FAS. In the MAS dimer, like in the FAS dimer, the monomers may have head-to-tail arrangement (47).

The mycobacterial genes involved in fatty acid synthesis seem to have a high degree of resemblance to the gene coding for the enzyme that produces macroyclic lactone in S. erythraea (Fig. 6). The eryA gene that codes for the enzyme that synthesizes erythronolide B contains three open reading frames each coding for a fused dimeric FAS-like protein (34, 44, 48), and the mas gene shows a high degree of similarity to segments of eryA gene. We have recently purified FAS from M. tuberculosis bovis BCG, and this protein is composed of 500-kDa monomers which probably represent two fused monomers (31) as found in eryA gene.

The open reading frame of mas as deduced from nucleic acid sequence lacks a thioesterase domain that resembles the thioesterase domains of vertebrate fatty acid synthase and eryA from Streptomyces. A transferase domain, like the palmitoyl transferase of yeast, that might transfer the product as CoA ester was also not found in the mas gene. The lack of a domain that would catalyze either release of free mycocerosic acids or their CoA esters helps to explain the finding that purified MAS showed extremely low specific activity and the product was found covalently attached to the enzyme (18). This lack of release of the free product might have a biological function. In vivo mycocerosic acids are found exclusively as glycosylated phenolphthiocerol dimycocerosate in the cell wall (15). If the mycocerosic acid is released either as free acid or as CoA ester, these acids may be incorporated into other cellular lipids. Therefore, a direct transfer from MAS to phenolphthiocerol might be an effective method to target the product to the specific glycolipid. To test whether MAS can catalyze such a transfer we tested the effect of purified phenolphthiocerol on purified mas from M. tuberculosis bovis BCG on purified MAS; but the rate of methylmalonyl-CoA incorporation was enhanced only <100%, and release of mycocerosic acid from the synthase could not be detected. Therefore, we tentatively conclude that MAS probably functions in conjunction with a separate transferase that transfers the mycocerosyl groups from MAS to the hydroxyl groups of phenolphthiocerol.

Since mas gene does not appear to contain a mycocerosyl transferase domain, we searched for other open reading frames near the mas coding region. Three open reading frames were found adjacent to the mas gene. These open reading frames 1, 2, and 3 could code for 272, 355, and 298 amino acid residues identical to these in the essential serine site of acyl carrier protein.

2 Mathur and P. E. Kolattukudy, unpublished observation.
polypeptides, respectively. Hydrophobicity plots of these three proteins indicate that all three could contain transmembrane domains (data not shown). Data bank searches did not reveal adequate similarity to enable us to classify the protein products of these open reading frames as members of any known protein classes. However, the observed limited identities are tantalizing enough to suggest that one of them might be a mycocerosyl transferase. For example, open reading frame 1, that is at the 5′-side of mas gene, could code for a 272-amino acid protein that has a region homologous to the acyl trans-ferase of Streptomyces. It is possible that this protein might be anchored to the membrane and functions by directly trans-ferring the mycocerosyl group from MAS to the hydroxyl group of phenolphthiocerol. Recent immunogold labeling re-sults indicating that MAS is located in association with the membrane, whereas FAS is located in the cytoplasm, is consistent with such a hypothesis. This hypothesis is analogous to that postulated in the case of the synthesis of the cyclic peptide, gramicidin S (49, 50). In this case, at the 5′-side of a large open reading frame that codes for the multi-functional gramicidin S synthase, a small open reading frame was found that could code for thioesterase-like protein, homologous to the S-acyl fatty acid synthase thioesterase that we originally cloned from an animal tissue (51). This thio-esterase-like protein was postulated to be involved in the term-ination of the multistep gramicidin synthesis.

Many important natural products are derived from methylmalonyl-CoA only, malonyl-CoA only, or both of these elongation substrates. However, which enzyme involved in the process selectively chooses the branched or precursor and the molecular basis of the specificity are un-known. The present results raise the possibility of discovering structural factors of the enzyme that might be involved in selecting methylmalonyl-CoA as the substrate. The active site areas of the acyl transferase domain and the ketoacyl synthase domain of mycobacterial MAS showed 80 and 93% identity, respectively, to the corresponding regions of erythronolide synthase, while the corresponding regions of vertebrate FAS showed only 40 and 67% identity, respectively. On the other hand, the other active sites such as those of the reductases of MAS were much more similar to those of vertebrate FAS than to those of erythronolide synthase. Since the uniqueness of the erythronolide synthase and MAS is that they both use methylmalonyl-CoA as the substrate, the striking similarities in the acyl transferase and ketoacyl synthase domains suggest that either or both of these domains may be selective for methylmalonyl-CoA. The availability of the cloned domains would allow us to test this hypothesis with the possibility of discovering methylmalonyl-CoA-selective enzyme domains. If this approach succeeds, tailor-made branching patterns could be introduced in the future into biologically active natural products with the possibility of generating new structures with novel activities.

The present results reveal the structure of a unique multi-functional chain-elongating enzyme that catalyzes the syn-thesis of mycrocyclic acids. These unusual fatty acids are found exclusively esterified to phenolphthiocerol that is found only in pathogenic mycobacteria. Since the mycocerosyl lipids are unique constituents of the walls of these pathogenic organisms, the synthesis of mycoceric acids could be a possible target for novel antimiycobacterial drugs that could be of great benefit to many millions of people afflicted by diseases caused by mycobacteria.

Acknowledgments—We acknowledge the valuable service of Jane Tolley, The Ohio State Biochemical Instrument Center, for amino acid sequencing.

REFERENCES
1. Grisebach, H. (1979) in Antibiotics and Other Secondary Metabolites (Hutter, R., Leisinger, T., Nunsch, J., and Wehrli, W., eds) pp. 113–127, Academic Press, London
2. Jacob, J. (1970) in Chemistry and Biochemistry of Natural Waxes (Kolattukudy, P. E., ed) pp. 93–146, Elsevier, Amsterdam
3. Downing, D. T. (1976) in Chemistry and Biochemistry of Natural Waxes (Kolattukudy, P. E., ed) pp. 149–184, Elsevier, Amsterdam
4. Buckner, J. S., and Kolattukudy, P. E. (1976) in Chemistry and Biochemistry of Natural Waxes (Kolattukudy, P. E., ed) pp. 149–184, Elsevier, Amsterdam
5. Rainwater, D. L., and Kolattukudy, P. E. (1982) Arch. Biochem. Biophys. 217, 372–393
6. Kim, Y. S., Kolattukudy, P. E., and Boon, A. (1979) Comp. Biochem. Physiol. 102, 443–447
7. Bakaletz, J. S., Kolattukudy, P. E., and Rogers, L. (1978) Arch. Biochem. Biophys. 185, 152–153
8. Rainwater, D. L., and Kolattukudy, P. E. (1980) J. Biol. Chem. 260, 616–629
9. Wakó, S. J., Stoops, J. K., and Joshi, V. C. (1983) Annui. Rev. Biochem. 52, 553–579
10. Cleveland, D. W., Fischer, G. S., Kirschner, M. H., and Laemmli, U. K. (1979) J. Biol. Chem. 254, 1102–1106
11. Matsudaira, P. (1987) J. Biol. Chem. 262, 10035–10038
12. Burnett, W. H. (1981) Anal. Biochem. 112, 195–203
13. Williams, K. R. (1980) in Current Protocols in Molecular Biology (Ausubel, M. F., Brent, R., Kingston, E. R., Moore, D., Seidman, G. J., Smith, A. J., and Struhl, K., eds) Vol. 1, pp. 2.4.3–2.4.5, Wiley Interscience, New York
14. Becker, D. E., Markham, R., Fricke, A. R., Fuller, G. B., and Jensen, H. R. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 5691–5695
15. Whitehead, A. S., Goldberger, G., Woods, D. E., Markham, A. F., and Kolton, H. R. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 5387–5391
16. Hanahan, D., and Meselson, M. (1980) Gene (Amst.) 10, 63–67
17. Sanburn, S. N., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463–5467
18. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, pp. 329–329, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
19. Dale, R. M. K., McClaire, B. A., and Houchins, J. P. (1985) Plasmid 13, 51–46
20. Southern, E. M. (1975) J. Mol. Biol. 98, 503–517
21. Itoh, S., Rainwater, D. L., and Kolattukudy, P. E. (1992) Arch. Biochem. Biophys. 295, 318–329
22. Shine, J., and Dalgaard, L. (1974) Proc. Natl. Acad. Sci. U.S.A. 71, 1342–1346
23. Poulooe, A. J., and Kolattukudy, P. E. (1985) Arch. Biochem. Biophys. 220, 652–656
24. Donadie, S., Staver, M. J., McAlpine, J. B., Swanson, S. J., and Katz, L. (1991) Science 252, 675–679
25. Holzer, K. P., Liu, W., and Hammes, G. G. (1989) Proc. Natl. Acad. Sci. U.S.A. 86, 4387–4391
26. Schweizer, M., Takabayashi, K., Loux, T., Beck, K.-F., and Schregimann, R. (1989) Nucleic Acids Res. 17, 567–569
27. Mohamed, A. H., Chirala, S., Slabi, N. D., Hwang, W.-Y., and Wakó, S. J. (1988) J. Biol. Chem. 263, 12315–12320
28. Kaumann, S., Siggaard-Anderson, M., and von Wettstein-Knowles, P. (1989) Carlsberg Res. Commun. 53, 357–370
29. Chirala, S., Baskar, R., Pratarandeh, M., Stolow, D. T., Hwang, W.-Y., and Poulooe, A. J. (1991) J. Biol. Chem. 266, 539–545
30. Poulooe, A. J., and Waks, S. J. (1991) J. Biol. Chem. 266, 171–175
31. Yamauchi, A., Sekiguchi, H., Rainwater, D. L., and Kolattukudy, P. E. (1992) J. Biol. Chem. 267, 15953–15958
32. Poulooe, A. J., and Kolattukudy, P. E. (1988) J. Biol. Chem. 263, 15953–15958
33. Poulooe, A. J., Rogers, L., Cheshouragh, J. M., and Kolattukudy, P. E. (1988) J. Biol. Chem. 263, 15953–15958