Biochemistry of the Initial Steps of Mycothiol Biosynthesis

Received for publication, May 17, 2006, and in revised form, July 25, 2006. Published, JBC Papers in Press, August 28, 2006, DOI 10.1074/jbc.M604724200

Gerald L. Newton, Philong Ta, Krzysztof P. Bzymek, and Robert C. Fahey

From the Department of Chemistry and Biochemistry, University of California, La Jolla, California 92093

Mycothiol is the major thiol produced by mycobacteria and is required for growth of Mycobacterium tuberculosis. The final three steps in the biosynthesis of mycothiol have been fully elucidated but the initial steps have been unclear. A glycosyltransferase, MshA, is required for production of the mycothiol precursor, 1-O-(2-acetamido-2-deoxy-α-D-glucopyranosyl)-D-myoinositol, but its substrates and immediate products were unknown. In this study, we show that the N-acetylglucosamine donor is UDP-N-acetylgalactosamine and that the N-acetylglucosamine acceptor is 1-O-myoinositol 1-phosphate. The reaction generates UDP and 1-O-(2-acetamido-2-deoxy-α-D-glucopyranosyl)-D-myoinositol 3-phosphate. Using cell-free extracts of Mycobacterium smegmatis, little activity was obtained with myo-inositol, 1-O-myoinositol 1-phosphate, or myo-inositol 2-phosphate as the N-acetylglucosamine acceptor. A phosphatase, designated MshA2, is required to dephosphorylate 1-O-(2-acetamido-2-deoxy-α-D-glucopyranosyl)-D-myoinositol 3-phosphate to produce 1-O-(2-acetamido-2-deoxy-α-D-glucopyranosyl)-D-myoinositol. The latter is deacetylated, ligated with cysteine, and the cysteinyl amino group acetylated by acetyl-CoA to complete the mycothiol biosynthesis pathway. Uptake and concentration of myo-[(14)C]inositol is rapid in Mycobacterium smegmatis and leads to production of radiolabeled inositol 1-phosphate and mycothiol. This demonstrates the presence of a myo-inositol transporter and a kinase that generates 1-O-myoinositol 1-phosphate. The biochemical pathway of mycothiol biosynthesis is now fully elucidated.

Mycothiol and other mycobacteria and other actinomycetes do not synthesize glutathione but produce another thiol, mycothiol (MSH), which has many of the same functions as glutathione (1). Mycothiol (Fig. 1) is comprised of N-acetylcysteine amide-linked to 1-O-(2-amino-2-deoxy-α-D-glucopyranosyl)-D-myoinositol (GlcN-Ins) (2–4). Mycothiol is maintained in the reduced state by an NADPH-dependent reductase (5, 6) and is involved in the detoxification of a variety of thioreactive agents (7–11).

The enzymes of mycothiol biosynthesis and metabolism have been of interest as potential targets for drugs in the treatment of tuberculosis. The final three steps of MSH biosynthesis have been elucidated (Fig. 1) and require the genes encoding a deacetylase MshB (Rv1170 gene in Mycobacterium tuberculosis) (12), a ligase MshC (Rv2130c) (13), and mycothiol synthase, a transacetylase MshD (Rv0819) (14). A glycosyltransferase MshA (Rv0486) was also identified by transposon mutagenesis, but the biochemical reaction catalyzed was not defined (15). Crystal structures for the MshB (16, 17) and MshD (18, 19) enzymes have been reported.

As the genes for mycothiol biosynthesis have been identified in studies using nonpathogenic Mycobacterium smegmatis, attempts were undertaken to generate the corresponding mutants in M. tuberculosis. Inactivation of the mshB gene does not fully block MSH biosynthesis because an alternative deacetylase activity leads to modest GlcN-Ins production and limited MSH synthesis in both M. smegmatis and M. tuberculosis (7, 20). MshC is essential for MSH biosynthesis in M. smegmatis, which does not require MSH for growth (9). On the other hand, the mshC gene is essential for growth of M. tuberculosis, indicating that MSH itself is essential for M. tuberculosis (21). Inactivation of the mshD gene in M. smegmatis (22) and M. tuberculosis (23) does not fully prevent MSH production. In the mshD mutants low levels of MSH are produced by a chemical transacylation reaction of acetyl-CoA with Cys-GlcN-Ins that accumulates at high levels in both mutants. Another thiol, N-formyl-Cys-GlcN-Ins is also produced in the mshD mutants and functions as a poor substitute for mycothiol in some biochemical reactions. The altered thiol status is thought to be responsible for the inability of mshD mutants to survive and grow in the macrophage (24). MshA is essential for synthesis of MSH in M. smegmatis (15) and we have recently shown that the mshA gene is required for growth of M. tuberculosis. Thus, two enzymes of mycothiol biosynthesis, MshA and MshC, have been identified as potentially important TB drug targets, and a third enzyme, MshD, shows significant potential.

MshA belongs to the retaining glycosyltransferase family 4 (25), a group that includes enzymes such as sucrose-phosphate synthase that catalyze the production of disaccharides and...
Mycothiol biosynthesis begins with a glycosyltransferase, MshA, required for production of GlcNAc-Ins. The product, Cys-GlcN-Ins, is acetylated by the MshA-catalyzed reaction are the subject of the current study. The possible inositol moieties of the mycobacterial cell wall (26) are devoid of GlcNAc-Ins, a pseudo-disaccharide intermediate in mycothiol biosynthesis (15). This suggests that the substrates for MshA include a nucleotide donor for the GlcNAc moiety and an inositol phosphate acceptor. UDP-GlcNAc is the source of GlcNAc for the unique linkage region of the mycobacterial cell wall (Fig. 1). The possible inositol acceptors used by MshA will be constrained by the stereochemistry of the inositol moiety of GlcNAc-Ins (27) and the phosphate residue blocks the known site for transfer of the GlcNAc moiety. This is the most commonly available inositol monophosphate and is produced from phosphoryl myo-inositol-1-phosphate synthase (28).

In the present study we elaborate the details of the reaction catalyzed by MshA. We identify the substrates and products of the MshA-catalyzed reaction, describe a simple assay for MshA activity, and show that an additional enzyme, a phosphatase designated MshA2, is required to complete the mycothiol biosynthesis pathway.

**Experimental Procedures**

**Preparation of Cell-free Extracts**—Unless otherwise specified the reagents were reagent or higher grade from Fisher. *M. smegmatis* wild-type strain mc²155 and mshB mutant Myco504 (20) from starter cultures were expanded to three 1-liter cultures in Middlebrook 7H9 medium (Difco) containing 0.05% Tween 80 and 0.4% glucose (with 20 μg/ml kanamycin for Myco504), shaken at 250 rpm and 37 °C. The mshA mutant mshA::Tn5 (15) of *M. smegmatis* was cultured in the same medium with 20 μg/ml kanamycin. Exponential phase cultures (A₆₀₀ = 0.7–1.0) were harvested by centrifugation (5,000 × g, 4 °C, 30 min). The cells were resuspended in 2 volumes of extraction buffer containing 25 mM HEPES, pH 7.5, 100 mM NaCl, 5 mM MgCl₂, 1 mM 2-mercaptoethanol (Sigma), and 35 μM each of the protease inhibitors Nα-p-tosyl-L-phenylalanyl chloromethyl ketone (Sigma) and Nα-p-tosyl-L-lysine chloromethyl ketone (Sigma). The cells were extracted at 4 °C with 4 passes through an Aminco French press. The extract was clarified by centrifugation (30,000 × g, 4 °C, 30 min) and the supernatant dialyzed twice against 100 volumes of extraction buffer. The extract was then centrifuged (30,000 × g, 4 °C, 30 min) to remove insoluble protein. The final extract typically contained 10–14 mg/ml protein as determined by the method of Bradford (Bio-Rad).

**Determination of MSH, GlcN-Ins, and GlcNAc-Ins**—The levels of MSH, GlcN-Ins, and GlcNAc-Ins were determined in exponential (A₆₀₀ = 0.5–0.8) and stationary phase (A₆₀₀ = 2.4–3.1) cultures of *M. smegmatis* mc²155 and mshB mutant strain Myco504 cultured as described above. Triplicate cell pellets were extracted in 50% acetonitrile containing mBBr for HPLC analysis of MSH as previously described (7). Triplicate cell pellets were also extracted in 50% acetonitrile without mBBr and processed before (GlcN-Ins determination) or after (GlcNAc-Ins determination) deacetylation with MshB (7).

**MshA Standard Assay Conditions**—Unless stated otherwise, MshA assays were performed with the undiluted dialyzed cell extract using 1 mM inositol compounds and 1 mM UDP-GlcNAc (Sigma) at 37 °C. The extracts produced stable activity (0.3–0.8 nmol/min/mg) for at least 2 weeks when stored at 4 °C. Stock solutions of the inositol compounds and UDP-GlcNAc were made at 20–50 mM in water and stored at −70 °C. UDP-GlcNAc, myo-inositol (Ins), 1-D-Ins-1-P, and Ins-2-P were obtained from Sigma. 1-D-Ins-1-P was obtained from Bachem and 1-D-Ins-1-P was enzymatically synthesized from glucose 6-phosphate using 1-L-mylo-inositol-1-phosphate synthase (28) as described under the supplemental materials.

For a standard MshA assay, a preheated (37 °C) 250-μl sample of the extract was added to an inositol compound, followed by UDP-GlcNAc to initiate the reaction. The reaction was sampled four times by removing a 50-μl sample and adding it to 50 μl of preheated (60 °C) acetonitrile. After a 10-min incubation at 60 °C, the sample was chilled on ice and centrifuged (16,000 × g, 3 min). For the GlcN-Ins assay, 7.5 μl of the supernatant was removed for labeling by AccQ-Fluor as described previously (7, 29). To the remaining supernatant, 4 μl of 5 M ammonium formate (adjusted to pH 5 with formic acid) was added to precipitate additional protein and mildly acidify the sample for HPLC. The sample was clarified by centrifugation (16,000 × g, 3 min), the supernatant was dried in a Savant SpeedVac, and the residual material was dissolved in 70 μl of water. A 50-μl sample was injected without dilution for HPLC analysis of UMP, UDP, and UDP-GlcNAc as detailed under the supplemental materials.

**Incorporation of Radiolabel from [Glucosamine-6-³H]Uridine Diphosphate N-Acetyl-d-glucosamine**—To identify the direct product of MshA, extracts of *M. smegmatis* mshB mutant Myco504 (20) were examined for the incorporation of [³H]glucosamine from UDP-[³H]GlcNAc. This extract lacks MshB and should terminate the mycothiol biosynthetic pathway at GlcNAc-Ins. A dialyzed extract of Myco504 was prepared as described above with a protein content of 11 mg/ml. The MshA assay was conducted with 1.5 ml of preheated extract, 1 mM 1-DL-Ins-1-P, 1 mM UDP-GlcNAc, and 2.5 μCi of uridine.
Initial Steps of Mycothiol Biosynthesis

diphosphate N-acetyl-D-glucosamine, [glucosamine-6-3H] uridine diphosphate N-acetyl-D-glucosamine (PerkinElmer Life Sciences, NET-434), in a total volume of 1.6 ml. Samples (200 μl) were removed at 0, 10, 20, 30, 60, 90, and 150 min and added to 200 μl of acetonitrile preheated to 60 °C. The samples were incubated for 10 min at 60 °C and then iced to precipitate protein. The samples were clarified by centrifugation (16,000 × g for 5 min) and the supernatant was removed to a separate microcentrifuge tube. For GlcN-Ins analysis an aliquot (7.5 μl) was assayed using exponentially growing cells in Middlebrook 7H9 medium (see above). One ml of 30 mM myo-inositol containing 0.51 μCi of [14C]inositol (myo-[U-14C]inositol, NEC-606, PerkinElmer Life Sciences) was added to a 1-liter culture of cells (37 °C) at A990 = 0.91 to give a final inositol concentration in the medium of 30 μM at zero time. Quadruplicate samples (45 ml) were chilled on ice at 30, 60, 90, 120, and 180 min and pelleted by centrifugation (5,000 × g, 15 min, 4 °C). The culture density was A990 = 1.5 at 3 h, the last sampling time. Four 1-ml samples of the medium for each sampling time were mixed with 8 ml of Scintiverse scintillation mixture (Fisher) and counted for 14C content. At each sampling time, two samples were centrifuged and the pellets extracted at 60 °C in 50% acetonitrile, aqueous 20 mM HEPES, pH 7.5, for analysis on HPLC Program A (supplemental materials). Pellets from two additional samples were extracted in 50% acetonitrile, aqueous 20 mM HEPES, pH 8.0, containing 2 mM monobromobimane (mBBr, Molecular Probes) for mycothiol analysis. After centrifugation, the supernatants were removed to 1.5-ml microcentrifuge tubes. Each pellet was washed with 1 ml of 50% acetonitrile/water (60 °C) and the wash was discarded. The pellets were dried to a constant weight in tarred microcentrifuge tubes in a vacuum oven, weighed, and transferred to glass scintillation vials. The pellets were incubated overnight at 37 °C with 0.5 ml of 10% sodium dodecyl sulfate and 0.1 ml of 3.8 M NaOH to solubilize the pellet and bleach the bimane label. The pellet sample was neutralized with 0.4 ml of 6 M HCl and 16 ml of Scintiverse scintillation mixture was added for counting.

[14C]Inositol-labeled thiols were analyzed by HPLC with fluorescence detection of the mBBr derivatives as described by Koledin et al. (14). One-min fractions (1 ml) were collected, mixed with 8 ml of Scintiverse scintillation mixture, and counted for [14C]inositol. Unretained 14C-labeled compounds eluted at 3 min and include mycothiol biosynthesis intermediates such as inositol, GlcNAc-Ins, and GlcN-Ins that do not label with mBBr. The mycothiol-bimane derivative (MSmB) eluted at 20 min; specific activity was estimated from mycothiol content based upon fluorescence quantitation with MSmB standards and counts from the MSmB containing fractions.

RESULTS

Substrates for MshA—Previous studies had shown that GlcNAc-Ins is an intermediate in MSH biosynthesis (12) and that the mshA gene encoding a glycosyltransferase was required for GlcNAc-Ins production in M. smegmatis (15). With the sequencing of the M. smegmatis genome now complete, the mshA gene can be identified as MSMEG0924. To pursue identification of the substrates of MshA we utilized the assay outlined in Scheme 1 in conjunction with cell-free extracts of M. smegmatis mc2155. Enzymatic reactions in the extract involving MshA, MshB, and possibly other unspecified enzymes generate GlcN-Ins. The amino group of GlcN-Ins was fluorescently labeled with AccQ-Fluor, allowing quantitative determination of GlcN-Ins by HPLC (7, 29). An extract of M. smegmatis mc2155 was prepared and the crude extract dialyzed to remove mycothiol and its precursors. This extract was used to test UDP-GlcNAc as the donor with various GlcNAc acceptors for production of GlcN-Ins (Fig. 2A). Endogenous MshB sufficed to fully deacetylate GlcNAc-Ins, and addition of supplemental

M. smegmatis Uptake and Incorporation of [14C]Inositol— The uptake and metabolism of [14C]inositol by M. smegmatis mc2155 was assayed using exponentially growing cells in Middle-
purified MshB did not increase GlcN-Ins values. The results show that no product is generated in the absence of acceptor or with Ins as acceptor. Maximal product formation was found using commercially available 1DL-Ins-1-P as acceptor, whereas 1D-Ins-1-P and Ins-2-P produced only minor amounts of GlcN-Ins. These results suggest that UDP-GlcNAc is a donor and 1L-Ins-1-P is the optimal acceptor in the MshA-catalyzed reaction.

A Direct HPLC Assay for MshA Based Upon Measurement of UDP Production—An assay of MshA activity that did not depend upon multiple enzyme activities was desirable. Having confirmed that UDP-GlcNAc was a donor in the MshA-catalyzed reaction, it appeared that determination of UDP production by HPLC would be a viable way to directly measure MshA activity. An HPLC protocol was developed based upon tetrabutylammonium ion-pairing chromatography that separates various nucleotide phosphates, including UDP, as well as sugar phosphates (supplemental materials Fig. S1), and was used to confirm the findings of the GlcN-Ins assay. Sample HPLC analyses of a crude extract reaction mixture are shown in supplemental Fig. S2. Fig. 2B shows that the substrate profile, as determined from UDP production on aliquots of the reaction analyzed in Fig. 2A, is qualitatively similar to that of Fig. 2A. However, the UDP assay is inherently less sensitive due to the higher background signal and UV detection. With extracts of mc2155 using 1DL-Ins-1-P as substrate, there is no lag in the production of UDP (Fig. 2B) that is seen in the production of GlcN-Ins (Fig. 2A). Because the expected initial product is GlcNAc-(\(\alpha\))-1L-Ins-1-P, which is alternatively designated GlcNAc-(\(\alpha\)-1,1)-1D-Ins-3-P and which we will abbreviate GlcNAc-Ins-P, subsequent dephosphorylation and deacetylation steps are required to generate GlcN-Ins. A significant lag in GlcN-Ins production (Fig. 2A) is therefore reasonable.

A similar UDP assay with extracts of a M. smegmatis mshA mutant strain mshA::Tn5 (15) using 1 mM each of UDP-GlcNAc and 1DL-Ins-1-P gave a specific activity of \(\frac{3 \text{ pmol min}^{-1}}{\text{mg}^{-1}}\), or \(<1\%\) of that of strain mc2155 (Fig. 2B) after subtraction of the control without 1DL-Ins-1-P. This confirms that the observed reaction is catalyzed by MshA (MSMEG0924) in the crude extract.

Optimal Substrate Concentrations for Assay of MshA—To optimize sensitivity in the assay, apparent \(K_m\) values were determined for UDP-GlcNAc and 1DL-Ins-1-P. The apparent \(K_m\) value for UDP-GlcNAc was 0.17 \(\pm\) 0.02 mM (supplemental Fig. S3A) and a \(K_m\) value of 0.24 \(\pm\) 0.06 mM was found for 1DL-Ins-1-P (supplemental Fig. S3B). Because the 1D-isomer is almost inactive as a substrate and comprises half of the mixture, the actual apparent \(K_m\) value for 1L-Ins-1-P is estimated to be 0.12 mM. However, this assumes that the 1D-isomer has no inhibitory effect upon MshA. This was subsequently determined to be the case when it became possible to produce authentic 1L-Ins-1-P from glucose 6-phosphate using inositol-1-phosphate synthase from Archaeoglobus fulgidus, kindly provided by Mary Roberts (28). When the 1L- and 1DL-Ins-1-P substrates were compared at concentrations below the \(K_m\) value, the former was almost twice as reactive (Fig. 3). After correcting for the background rate without inositol compounds as acceptor, the relative rates calculated at 30 min were: 1L-Ins-1-P,
Initial Steps of Mycothiol Biosynthesis

The content (μmol/g of residual dry weight) of mycothiol and its precursors was determined for wild-type strain mc2155 and the mshB mutant Myco504 during exponential and stationary phase as detailed under “Experimental Procedures.” Mean ± S.D. of triplicate determinations.

| Quantity       | mc2155 Exponential | mc2155 Stationary | Myco504* Exponential | Myco504* Stationary |
|----------------|--------------------|-------------------|----------------------|---------------------|
| GlcNAc-Ins     | <0.2               | <0.07             | 2.6 ± 0.2            | 6.2 ± 0.5           |
| GlcN-Ins       | 0.8 ± 0.3          | 0.05 ± 0.01       | <0.0006              | 0.026 ± 0.006       |
| MSH            | 9.6 ± 0.7          | 12.0 ± 0.05       | 0.41 ± 0.05          | 0.92 ± 0.04         |

* Cysteine was a major thiol, 0.72 ± 0.07 and 0.18 ± 0.01 μmol/g of residual dry weight, respectively, in exponential and stationary phase Mc2155 Myco504.

100 ± 6%; 1DL-Ins-1-P, 52 ± 5%; Ins-2-P, 10 ± 3%; 1D-Ins-P, 9 ± 3%; Ins, 2 ± 6%.

Determination of the $K_m$ value for 1L-Ins-1-P under conditions identical to those of supplemental Fig. S3 produced a value of 0.15 ± 0.01 mM. The apparent $K_m$ for UDP-GlcNAc using 1 mM 1L-Ins-1-P was 0.17 ± 0.04 mM, identical to that found using 1DL-Ins-1-P as substrate (Fig. S3A). When the reaction velocity for a 5-fold diluted crude extract was examined with 1 mM UDP-GlcNAc and either 1 mM 1L-Ins-1-P or 2 mM 1DL-Ins-1-P as substrates, the rates were found to be 0.70 ± 0.02 and 0.67 ± 0.01 nmol min⁻¹ mg⁻¹, respectively. The rate for 2 mM 1DL-Ins-1-P was 95% of that of 1 mM 1L-Ins-1-P, with overlapping errors, indicating that there was ≤5% inhibition by 1 mM 1D-Ins-1-P. Thus, 1D-Ins-1-P is neither a good substrate nor an inhibitor of MshA. This allows the use of the commercially available 1DL-Ins-1-P as substrate for MshA assays without complications.

The HPLC analysis was also used to determine the UDP-GlcNAc content for M. smegmatis mc2155 during exponential growth. A value of 2.0 ± 0.5 μmol/g of residual dry weight (n = 3) was obtained. Using a value of 4 μl of cellular water/mg of residual dry weight, this value translates to 0.50 ± 0.13 mM UDP-GlcNAc in the cell. Based upon the apparent $K_m$ value of 0.17 mM for UDP-GlcNAc determined for MshA activity in cell-free extracts, it appears that cellular levels of UDP-GlcNAc will not limit MshA activity.

Detection of Intermediates Leading to GlcN-Ins—With the donor and acceptor for the glycosyltransfer step established, the next task was to identify the expected intermediate GlcNAc-Ins-P. It was known that GlcNAc-Ins is a good substrate for MshB (12, 30) so we expected that the major pathway was from GlcNAc-Ins-P to GlcNAc-Ins and then deacetylation by MshB to produce GlcN-Ins. To simplify the analysis and maximize intermediate accumulation we utilized a mutant lacking MshB (Myco504). In extracts from this mutant, dephosphorylation of GlcNAc-Ins-P would produce GlcNAc-Ins as the end product, whereas if another deacetylase were available to act on this intermediate then a second phosphorylated intermediate (GlcN-Ins-P) might be formed with production of GlcN-Ins as a final product (Scheme 2).

To search for the GlcN-containing intermediates, the reaction of radiolabeled UDP-GlcN[6-3H]Ac with 1L-Ins-1-P was examined. To analyze the products we needed an HPLC protocol that would provide a good separation of potential phosphorylated sugar intermediates and yet permit analysis of eluents by mass spectrometry when the reaction was run with unlabeled substrate. Program B, a modified version of program A, was developed for this purpose. In program B, unphosphorylated sugars such as GlcNAc, GlcNAc-Ins, and inositol are unretracted, whereas GlcNAc-1-P and Ins-1-P are substantially retained, and UDP-GlcNAc and UDP are eluted later in the gradient.

Fig. 4 shows the results for HPLC analysis of a reaction catalyzed by extracts of Myco504 between radiolabeled UDP-GlcNAc and 1DL-Ins-1-P with counting of radioactivity in the eluted fractions. At zero reaction time the expected strong peak for UDP-GlcNAc was found at a retention time of ~30 min together with a small, unretracted peak at ~3 min and an additional peak at ~14 min reflecting impurities in the radiolabeled UDP-GlcNAc. After 60 min of reaction, there was a strong increase in the peak at ~3 min (U3), a modest increase in the...
FIGURE 4. HPLC analysis of UDP-[3H]GlcNAc-labeled products in the MshA catalyzed reaction. A dialyzed cell-free extract of mshB mutant of \( M. \) smegmatis (Myco504) \( \text{(20)} \) was used to truncate the mycothiol biosynthesis pathway at GlcNAc-Ins and increase the levels of GlcNAc-Ins precursors produced by MshA. The extract (1.5 ml) was assayed with 1 mM 1DL-Ins-1-P and 1 mM UDP-GlcNAc containing 2.5 \( \mu \)Ci of UDP-Glc[6-3H]Ac in a total assay volume of 1.6 ml. The extract was sampled at intervals and analyzed for GlcN and GlcNAc by AccQ-Fluor labeling and HPLC (not shown). The remaining sample was processed for analysis by tetrabutylammonium ion-pairing HPLC with GlcNAc-Ins at all time points, contained lesser amounts of GlcN-Ins by AccQ-Fluor labeling and HPLC (not shown). The remaining sample was processed for analysis by tetrabutylammonium ion-pairing HPLC with levels in the assays (Fig. S2). The GlcN and GlcNAc-Ins content of fraction U3 was determined by HPLC analysis before and after deacetylation with purified MshB and subsequent fluorescent labeling of the amino group with AccQ-Fluor. As expected for the \( mshB \) mutant, the GlcN-Ins content was very low (\( \leq 0.03 \mu M \)). However, after treatment with MshB the samples gave a substantial increase in GlcN-Ins analyses, corresponding to a high GlcNAc-Ins content in the original samples. This accounts for the major fraction of the U3 content at late times.

The variation in the composition of the reaction with reaction time is shown in Fig. 5. U3 was comprised largely of GlcNAc-Ins at all time points, contained lesser amounts of GlcN (\( \leq 0.2 \) \( \mu M \)), and no significant amount of GlcN-Ins (\( \leq 0.03 \) \( \mu M \)). The level of U3 increased by 240 \( \mu M \), showing that this peak contained the major radiolabeled product that proved to be GlcNAc-Ins at 220 \( \mu M \) at the final time point (Fig. 5). In contrast, the peaks for GlcNAc-1-P and U14 reached an early maximum and then declined. This indicates that GlcNAc-1-P and U14 are intermediates that are converted to another product or other products. At 150 min 300 \( \mu M \) UDP-GlcNAc was utilized and 280 \( \mu M \) UDP (from \( A_{260} \) monitoring) was produced.

Intermediate U14 Is GlcNAc-Ins-P—To identify component U14 the reaction catalyzed by \( mshB \) mutant Myco504 was conducted for 20 min using unlabeled UDP-GlcNAc and 1DL-Ins-1-P, and fractions were collected during HPLC separation. Each fraction was concentrated and individually reanalyzed by HPLC with mass spectral analysis (m/z 200 – 800). Fractions collected at 15 and 16 min showed the expected m/z of 462 for GlcNAc-Ins-P and these were combined, and the mixture concentrated and rechromatographed. Fig. 6B shows the HPLC analysis of this sample with specific ion mass spectral analysis at m/z 259 (Ins-1-P), 300 (GlcNAc-1-P), and 462 (GlcNAc-Ins-P). As expected, a strong m/z 259 signal was observed at 15.5

peak at \( \pm 14 \) min (U14), and a significant peak appeared at \( \sim 19 \) min corresponding to GlcNAc-1-P. The later presumably arises from the action of a pyrophosphatase on UDP-GlcNAc. The other product of the pyrophosphatase, UMP, is evident at low levels in the assays (Fig. S2). The GlcN and GlcNAc-Ins content of fraction U3 was determined by HPLC analysis before and after deacetylation with purified MshB and subsequent fluorescent labeling of the amino group with AccQ-Fluor. As expected for the \( mshB \) mutant, the GlcN-Ins content was very low (\( \leq 0.03 \mu M \)). However, after treatment with MshB the samples gave a substantial increase in GlcN-Ins analyses, corresponding to a high GlcNAc-Ins content in the original samples. This accounts for the major fraction of the U3 content at late times.

The variation in the composition of the reaction with reaction time is shown in Fig. 5. U3 was comprised largely of GlcNAc-Ins at all time points, contained lesser amounts of GlcN (\( \leq 0.2 \) \( \mu M \)), and no significant amount of GlcN-Ins (\( \leq 0.03 \) \( \mu M \)). The level of U3 increased by 240 \( \mu M \), showing that this peak contained the major radiolabeled product that proved to be GlcNAc-Ins at 220 \( \mu M \) at the final time point (Fig. 5). In contrast, the peaks for GlcNAc-1-P and U14 reached an early maximum and then declined. This indicates that GlcNAc-1-P and U14 are intermediates that are converted to another product or other products. At 150 min 300 \( \mu M \) UDP-GlcNAc was utilized and 280 \( \mu M \) UDP (from \( A_{260} \) monitoring) was produced.
GlcNAc-Ins-P as the product of the MshA-catalyzed reaction and indicate that its further metabolism proceeds by dephosphorylation to GlcNAc-Ins and subsequent deacetylation by MshB to produce GlcN-Ins.

Import of [14C]Ins by M. smegmatis and Incorporation into MSH—The present HPLC methods were utilized to analyze the products generated in M. smegmatis mc²155 from uptake and metabolism of radiolabeled Ins. The distribution of counts between the medium, the pellet, and the supernatant fractions is shown in Fig. 7A. Over the 3-h incubation ~90% of the radioactivity disappeared from the medium. The counts remained largely constant in the supernatant fraction but steadily increased in the pellet until the medium Ins was largely depleted, presumably reflecting incorporation into cell wall components. Analysis of the mBBr-labeled samples by fluorescence and by counting the eluted fractions allowed determination of the specific activity of the cellular MSH. This increased sharply during the first 90 min and much more slowly afterward (Fig. 7B). Inositol, GlcNAc-Ins, and GlcN-Ins elute with the unretained fraction in this chromatography, and the counts in this fraction were maximal at the first measurement and declined steadily thereafter.

HPLC analyses on program A of the samples prepared at 60 and 120 min without mBBr labeling are shown in Fig. 8. The unretained fraction eluting at ~3 min contains Ins and possibly other uncharged (GlcNAc-Ins) or cationic (GlcN-Ins) Ins derivatives and exhibits a marked decrease with time as the radiolabeled Ins is incorporated into products. The mycothiol biosynthesis intermediates GlcNAc-Ins and GlcN-Ins are typically a fraction of the mycothiol content and represent a small contribution to the unretained Ins counts. The MSH in this sample is largely oxidized to mycothiol disulfide under the conditions of the extraction and elutes at 9–10 min as established with authentic mycothiol disulfide monitored by absorbance at 220 nm. Inositol monophosphate is observed at 21–22 min at a level that remains constant from 60 to 120 min. A transient peak (U35) was maximal at 60 min and its elution time suggests that it is an inositol diphosphate. U35 might also be a more hydrophobic inositol derivative such as phosphatidylinositol, an established component of mycobacteria (26, 31). Although we could detect soybean phosphatidylinositol (m/z 834) by direct injection on the mass detector (ES⁺ mode), it was not eluted from the HPLC column under the conditions used in these assays. This indicates that U35 is not a phosphatidylinositol.

The time dependence of these components is shown in Fig. 7C. The counts determined in mycothiol disulfide, derived from oxidation of MSH during the extraction process, follow the same pattern as found for the specific activity of MSH determined as MSmB (Fig. 7B). This is expected because they represent independent measurements of MSH content. The counts associated with Ins-P and U35 reach an early peak and then decline slowly, indicating that these components are intermediates that reach a steady-state concentration. Based upon the specific activity of the [14C]Ins employed, the steady-state cellular concentration of [14C]Ins-P was estimated at 0.2 mM. These studies establish that Ins is actively transported into M.

![Figure 6](image_url)
DISCUSSION

The results of the present study establish the pathway for mycothiol biosynthesis as shown in Fig. 9. The GlcNAc acceptor for MshA is 1L-Ins-1-P, known to be produced from Glc-6-P in a reaction catalyzed by inositol-1-phosphate synthase (Ino1, Fig. 9) (32). UDP-GlcNAc is an effective GlcNAc donor for MshA and is known to be a donor for biosynthesis of glycolipids 1 and 2, precursors of the linkage region of the cell wall (26). UDP-GlcNAc is present at a concentration in \emph{M. smegmatis} well above its apparent \(K_m\) value for MshA and is therefore considered to be the physiological GlcNAc donor. Other NDP-GlcNAc donors may also exhibit activity with MshA but there

\emph{smegmatis}, converted to inositol monophosphate, and utilized in the production of mycothiol.

FIGURE 8. HPLC analysis of extracts from \emph{M. smegmatis} mc²155 incubated with \([\text{14C}]\)Ins. Samples from the experiment described in the legend to Fig. 7 were analyzed by HPLC program A (supplemental Fig. S1) and fractions were collected for scintillation counting. The analyses shown are for supernatant fractions prepared after 60 (○) and 120 min (★) incubation of cells with \([\text{14C}]\)Ins.
are no precedents for such donors. None are commercially available but isolation of ADP-GlcNAc from corn has been reported (35).

With UDP-GlcNAc identified as a donor, a direct assay of MshA activity based upon UDP determination by HPLC became possible. This assay produced similar results to those based upon GlcN-Ins production but lacked the initial lag-time associated with a multistep process (Fig. 2B). Stereochromically pure 1L-Ins-1-P was prepared from glucose 6-phosphate using inositol-1-phosphate synthase (28) and was used to establish that 1L-Ins-1-P is the immediate substrate of MshA (Fig. 3) and that 1D-Ins-1-P does not inhibit the enzyme at concentrations used in the assay. Thus, the commercially available 1D-Ins-1-P can be utilized to assay MshA without complications. The apparent Km values determined for 1D-Ins-1-P, 1L-Ins-1-P, and UDP-GlcNAc should prove useful for setting assay conditions.

From the nature of the substrates for MshA we predict that the product in addition to UDP is GlcNAc-Ins-P. Direct evidence for production of GlcNAc-Ins-P as a reaction intermediate leading largely to GlcNAc-Ins was obtained in studies using the msbB mutant Myco504 (20) to block deacetylation of GlcNAc-Ins. Because GlcNAc-Ins accumulates to a high level in this mutant (Table 1), the main pathway leading to GlcN-Ins in the wild-type cells must be via GlcNAc-Ins (Scheme 2).

Although 1D-Ins-1-P and Ins-2-P were poor substrates using the definitive GlcN-Ins assay (Fig. 2A), it is instructive to consider how their low activity is generated. Dephosphorylation of these substrates to Ins and subsequent rephosphorylation to generate 1L-Ins-1-P can be ruled out for several reasons. First, Ins itself is inactive in this assay and, second, the dialyzed extracts are devoid of nucleotide triphosphates required for Ins phosphorylation. It is difficult to understand how 1D-Ins-1-P can directly serve as an acceptor, as this would generate a product with the wrong stereochemistry. In addition, upon dephosphorylation it would produce 1-O-(2-acetamido-2-deoxy-α-D-glucopyranosyl)-L-myoinositol (GlcNAc-(α1,1)-1L-Ins) and Nicholas et al. (36) have shown that this stereoisomer is inactive with MshB, whereas the 1-O-(2-acetamido-2-deoxy-α-D-glucopyranosyl)-L-myoinositol (GlcNAc-(α1,1)-1D-Ins) is fully active. Thus, to explain the activity of 1D-Ins-1-P in the MshA assay we are forced to postulate the presence of an epimerase activity in the extract that converts this isomer to 1L-Ins-1-P.

The activity of Ins-2-P may result from the presence of an isomerase that converts it to the 1L-Ins-1-P or it may function as a poor GlcNAc acceptor generating GlcNAc-(α1,1)-1D-Ins-2-P, which is dephosphorylated to generate GlcNAc-Ins. This process can make only a minor contribution to the production of MSH in cell-free extracts. It is not known whether Ins-2-P is even produced in mycobacteria; if it is not then this activity is irrelevant in cells.

An additional mycothiol biosynthesis enzyme, the GlcNAc-Ins-P phosphatase designated MshA2, must be postulated to complete the mycothiol biosynthetic pathway (Fig. 9). The gene encoding this enzyme is the one missing item of genetic information for mycothiol biosynthesis. Also, the one or more IMP activities involved in dephosphorylation of 1L-Ins-1-P remain to be identified. Several genes encoding homologs of the human imp gene have been identified in the M. tuberculosis genome, including suhB (Rv2701c), Rv3137, impA, and cysQ (37), and these are potential candidates to encode MshA2 as well as IMP. The suhB gene has been cloned and expressed, and SuhB purified (37). SuhB was shown to have high activity with 1D-Ins-1-P, and lower activity with several other sugar phosphates, but activity with 1L-Ins-1-P was not tested (37). Similarly CysQ was recently reported to have IMP activity although only the 1D-isomer of Ins-1-P was apparently tested as substrate (38). The stereospecificity of IMPs from human (39), Escherichia coli, and A. fulgidus (40) has been examined with 1L-Ins-1-P and 1D-Ins-1-P and all were found to hydrolyze both substrates, so this may also be the case for the mycobacterial IMP.

The 1L-Ins-1-P required for MSH biosynthesis is normally produced by inositol-1-phosphate synthase (Ino1) but can also be derived from exogenous inositol, as shown in the earlier studies of Gammon et al. (41) and confirmed in the present work. The present results for intact cells establish that inositol is taken up against a concentration gradient, implicating an active transport system. [14C]Ins rapidly reaches a cellular level of ≥1.2 mM and then falls 20-fold during the 3-h incubation, largely being incorporated into the pellet (cell wall) fraction. In the supernatant fraction, [3H]Ins is rapidly converted to [14C]Ins-P and reaches a steady level at 30 min and then
remains largely constant for the next 90 min (Fig. 7C). The steady-state level of soluble cellular \(^{14}C\)Ins-P is estimated as ~0.2 mM, a value comparable with the apparent \(K_m\) determined for 1L-Ins-1-P (0.15 mM). During this period the counts associated with cellular Ins fall sharply and those in MSH increase, consistent with the processing of Ins via Ins-P to MSH. An unknown inositol derivative, U35, is transiently observed in the supernatant fraction from 30 to 90 min. This retention time corresponds to those of diphosphate compounds such as UDP and UDP-GlcNAc, suggesting that U35 may be a soluble inositol diphosphate. This compound could also be a more hydrophobic inositol derivative, albeit not a phosphatidylinositol. Gammon et al. (41) also reported a hydrophobic myo-\(^{14}C\)inositol-labeled compound in similar inositol uptake experiments with \(M.\) smegmatis.

To accommodate these results there must be two proteins, an inositol active transport protein and an inositol kinase (designated InoK), as indicated in Fig. 9. Partially purified myo-inositol kinases from wheat germ (42) and \(D\). \(D\)ictyostelium (43) have been shown to generate 1L-Ins-1-P but no gene encoding this enzyme has been identified. Prior to the present work, the only established function for 1L-Ins-1-P in mycobacteria was that of the intermediate in the biosynthesis of Ins. Now that 1L-Ins-1-P has been shown to be the starting substrate for MSH biosynthesis, a significant role for myo-inositol kinase in mycobacteria can be identified as the activation of Ins from exogenous sources, or from degradative processes, for utilization by MshA in MSH biosynthesis.

The primary transporter of myo-inositol in \(B.\) \(B\)acillus subtilis has been shown by directed gene knockout to be encoded by the \(iolT\) gene (formerly called \(ydjK\)) (44). When the \(iolT\) protein sequence (473 residues) was used in a blastp search of the TB genome on TubercuList, the only strong match was Rv3331, which was 26% identical in a 432-residue overlap and annotated as a probable sugar-transport integral membrane protein. An analogous search of the \(M.\) \(M\) sme2natis genome at the TIGR website identified MSMEG 5539 and MSMEG 4185 as strong potential homologs with 31% identity in 446- and 456-residue overlaps, respectively. Rv3331 was most similar to MSMEG 5539, being 24% identical in a 384-residue overlap. Rv3331 appears to be the best candidate for a protein to transport Ins in \(M.\) \(M\) tuberculosis.

The majority of the counts from \(^{14}C\)inositol are associated with the pellet fraction that contains the cell wall (Fig. 7A). Haites et al. (31) have examined the organic soluble lipid compounds in some detail for wild-type \(M.\) \(M\) sme2natis and for an \(ino1\) mutant following pulse labeling with \(^{3}H\)Ins. Incorporation of \(^{3}H\)Ins was first observed in phosphatidylinositol and later in phosphatidylinositol mannosides. They show that transfer of the cells to inositol-free medium leads to rapid degradation of the phosphatidylinositol and apolar phosphatidylinositol mannosides that are substantially utilized to synthesize polar phosphatidylinositol mannosides and lipooarabinomannan. Whether degradation of the membrane-bound phosphatidylinositol by phospholipases C or D can serve as a source of inositol for mycothiol biosynthesis is unclear.

In summary, the biochemical pathways involved in mycothiol biosynthesis have now been fully identified. The initial substrates for MshA, 1L-Ins-1-P, and UDP-GlcNAc, and a new intermediate, GlcNAc-Ins-P, have been identified. Dephosphorylation of GlcNAc-Ins-P by MshA2 generates GlcNAc-Ins, the substrate of MshB, and the remaining steps in the pathway leading to MSH are as previously elucidated. 1L-Ins-1-P is normally generated by Ino1 from glucose-6-P, but can also be produced by active transport of Ins from the external environment followed by phosphorylation. The genes encoding MshA2, the inositol kinase producing 1L-Ins-1-P from Ins, and the Ins transport protein need to be identified experimentally.

Acknowledgments—We thank Mary Roberts of Boston College for a generous gift of \(A.\) \(A\) fulgidus 1L-myoinositol-1-phosphate synthase and Yoosef Av-Gay and Helen Billman-Jacobe for kindly providing the \(M.\) \(M\) sme2natis \(M\) yc0504 mshB mutant. We are indebted to Nancy Buchmeier and Carole Bewley for critical review of the manuscript.

REFERENCES

1. Newton, G. L., and Fahey, R. C. (2002) Arch. Microbiol. 178, 388–394
2. Newton, G. L., Bewley, C. A., Doyer, T. J., Horn, R., Aharonowitz, Y., Cohen, G., Davies, J., Faulkner, D. J., and Fahey, R. C. (1995) Eur. J. Biochem. 230, 821–825
3. Sakuda, S., Zhou, Z.-Y., and Yamada, Y. (1994) Biosci. Biotech. Biochem. 58, 1347–1348
4. Spies, H. S., and Steenkamp, D. J. (1994) Eur. J. Biochem. 224, 203–213
5. Patel, M. P., and Blanchard, I. S. (1999) Biochemistry 38, 11827–11833
6. Patel, M. P., and Blanchard, I. S. (2001) Biochemistry 40, 3119–3126
7. Buchmeier, N. A., Newton, G. L., Koledin, T., and Fahey, R. C. (2003) Mol. Microbiol. 47, 1723–1732
8. Newton, G. L., Av-Gay, Y., and Fahey, R. C. (2000) Biochemistry 39, 10739–10746
9. Vetting, M. W., Yu, M., and Blanchard, J. S. (2003) Biochemistry 42, 12067–12076
10. Vogt, R. N., Steenkamp, D. J., Zheng, R., and Blanchard, J. S. (2003) Biochem. J. 374, 657–666
11. Newton, G. L., Av-Gay, Y., and Fahey, R. C. (2000) J. Bacteriol. 182, 6958–6963
12. Sareen, D., Steffek, M., Newton, G. L., and Fahey, R. C. (2002) Biochemistry 41, 6885–6890
13. Sareen, D., Steffek, M., Newton, G. L., and Fahey, R. C. (2002) Arch. Microbiol. 178, 331–337
14. Koledin, T., Newton, G. L., and Fahey, R. C. (2002) Arch. Microbiol. 178, 1131–1141
15. Vetting, M. W., Roderick, S. L., Yu, M., and Blanchard, J. S. (2003) Protein Sci. 12, 1954–1959
16. Maynes, J. T., Garen, C., Cherney, M. M., Newton, G. L., Arad, D., Av-Gay, Y., Fahey, R. C., and James, M. N. (2003) J. Biol. Chem. 278, 47166–47170
17. McCarthy, A. A., Peterson, N. A., Knijff, R., and Baker, E. N. (2004) J. Mol. Biol. 335, 1131–1141
18. Vetting, M. W., Roderick, S. L., Yu, M., and Blanchard, J. S. (2002) Proc. Natl. Acad. Sci. U. S. A. 103, 33919–3400
19. Spies, H. S., and Steenkamp, D. J. (1994) Eur. J. Biochem. 224, 203–213
20. Vetting, M. W., Yu, M., and Blanchard, J. S. (2003) Biochemistry 42, 12067–12076
21. Maynes, J. T., Garen, C., Cherney, M. M., Newton, G. L., Arad, D., Av-Gay, Y., Fahey, R. C., and James, M. N. (2003) J. Biol. Chem. 278, 47166–47170
22. McCarthy, A. A., Peterson, N. A., Knijff, R., and Baker, E. N. (2004) J. Mol. Biol. 335, 1131–1141
23. Vetting, M. W., Roderick, S. L., Yu, M., and Blanchard, J. S. (2002) Proc. Natl. Acad. Sci. U. S. A. 103, 33919–3400
24. Rengarajan, J., Bloom, B. R., and Rubin, E. J. (2005) Proc. Natl. Acad. Sci. U. S. A. 103, 8327–8332
25. Campbell, J. A., Davies, G. J., Bulone, V., and Henrissat, B. (1997) Biochem. J. 326, 929–939

Initial Steps of Mycothiol Biosynthesis
Initial Steps of Mycothiol Biosynthesis

26. Mikusova, K., Mikus, M., Besra, G. S., Hancock, I., and Brennan, P. J. (1996) J. Biol. Chem. 271, 7820–7828
27. Nicholas, G. M., Eckman, L. L., Kovac, P., Otero-Quintero, S., and Bewley, C. A. (2003) Bioorg. Med. Chem. 11, 2641–2647
28. Chen, L., Zhou, C., Yang, H., and Roberts, M. F. (2000) Biochemistry 39, 12415–12423
29. Anderberg, S. J., Newton, G. L., and Fahey, R. C. (1998) J. Biol. Chem. 273, 30391–30397
30. Newton, G. L., Ko, M., Ta, P., Av-Gay, Y., and Fahey, R. C. (2006) Protein Expression Purif. 47, 542–550
31. Haites, R. E., Morita, Y. S., McConville, M. J., and Billman-Jacobe, H. (2005) J. Biol. Chem. 280, 10981–10987
32. Bachhawat, N., and Mande, S. C. (1999) J. Mol. Biol. 291, 531–536
33. Brennan, P. J., and Nikaido, H. (1995) Annu. Rev. Biochem. 64, 29–63
34. Bornemann, C., Jardine, M. A., Spies, H. S. C., and Steenkamp, D. J. (1997) Biochem. J. 325, 623–629
35. Dankert, M., Passeron, S., and Recondo, E. (1972) An. Asoc. Quim. Argent. 60, 257–271
36. Nicholas, G. M., Eckman, L. L., Newton, G. L., Fahey, R. C., Ray, S., and Bewley, C. A. (2003) Bioorg. Med. Chem. Lett. 11, 601–608
37. Nigou, J., Doer, L. G., and Besra, G. S. (2002) Biochemistry 41, 4392–4398
38. Gu, X., Chen, M., Shen, H., Jiang, X., Huang, Y., and Wang, H. (2006) Biochem. Biophys. Res. Commun. 339, 897–904
39. Bone, R., Frank, L., Springer, J. P., Pollack, S. J., Osborne, S. A., Atack, J. R., Knowles, M. R., McAllister, G., Ragan, C. I., Broughton, H. B., Baker, R., and Fletcher, S. R. (1994) Biochemistry 33, 9460–9467
40. Morgan, A. J., Wang, Y. K., Roberts, M. F., and Miller, S. J. (2004) J. Am. Chem. Soc. 126, 15370–15371
41. Gammon, D. W., Hunter, R., Steenkamp, D. J., and Mudzunga, T. T. (2003) Bioorg. Med. Chem. Lett. 13, 2045–2049
42. Loewus, M. W., Sasaki, K., Leavitt, A. L., Munsell, L., Sherman, W. R., and Loewus, F. A. (1982) Plant Physiol. 70, 1661–1663
43. Stephens, L. R., Kay, R. R., and Irvine, R. F. (1990) Biochem. J. 272, 201–210
44. Yoshida, K., Yamamoto, Y., Omae, K., Yamamoto, M., and Fujita, Y. (2002) J. Bacteriol. 184, 983–991