Biochemistry

 Volume 41, Number 22 June 4, 2002

Accelerated Publications

ATP-Dependent L-Cysteine:1D-*myo*-Inosityl 2-Amino-2-deoxy-α-D-glucopyranoside Ligase, Mycothiol Biosynthesis Enzyme MshC, Is Related to Class I Cysteinyl-tRNA Synthetases†

Dipti Sareen, Micah Steffek, Gerald L. Newton, and Robert C. Fahey*

Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093

Received December 28, 2001

ABSTRACT: Mycothiol is a novel thiol produced only by actinomycetes and is the major low molecular weight thiol in mycobacteria. The mycothiol biosynthetic pathway has been postulated to involve ATP-dependent ligation of L-cysteine (Cys) with 1D-*myo*-inosityl 2-amino-2-deoxy-α-D-glucopyranoside; GlcN-Ins) catalyzed by MshC to produce Cys-GlcN-Ins. The ligase activity was purified ~2400-fold from *Mycobacterium smegmatis* and two proteins of slightly different $M_r \sim 47000$ were identified with MshC activity. The N-terminal sequence of the smaller protein revealed that it was coded by a gene in the databases for *M. smegmatis* and *M. tuberculosis* previously designated as *cysS2*. The larger protein was coded by the same gene in *M. smegmatis* but included an eight amino acid N-terminal extension involving a different start codon. The ligase was found to have K_m values of 40 ± 3 and 72 ± 9 μM for Cys and GlcN-Ins, respectively. The *cysS2* gene was thought to encode a second cysteinyl-tRNA synthetase in addition to *cysS* but the present results indicate that *cysS2* is actually the *mshC* gene encoding ATP-dependent Cys:GlcN-Ins ligase.

Mycothiol (MSH; AcCys-GlcN-Ins; 1D-*myo*-inosityl 2-(*N*-acetyl-L-cysteinyl)amido-2-deoxy-α-D-glucopyranoside)[1] is a novel low molecular weight thiol (*1*) produced by most actinomycetes (*2*). Like glutathione, MSH appears to play a significant role in protecting these bacteria against oxidative stress (*3*, *4*), and a disulfide reductase capable of maintaining it in the reduced state has been recently characterized (*5*−*7*). In addition, MSH functions in the detoxification of formaldehyde (*8*−*10*) and of alkylating agents (*11*). Enzymes of MSH metabolism would appear to be potential targets for new drugs directed against actinomycetes, including *Mycobacterium tuberculosis*, which produces especially large quantities of MSH (*2*).

The biosynthesis of mycothiol would seem to offer the most attractive targets for new drugs but our understanding of the enzymology of this process is still in the early stages of discovery. Upon the basis of studies with cell extracts, Bornemann et al. (*12*) proposed that the final steps of the biosynthesis involve ATP-dependent ligation of L-cysteine with GlcN-Ins (MshC) followed by acetylation of the product

† This work was supported by grants from the National Science Foundation (9981850) and from Cytoclonal Pharmaceutics Incorporated.

* Author for correspondence: E-mail rcfahey@ucsd.edu; telephone 858-534-2163; FAX 858-534-4864.

[1] Abbreviations: GlcN-Ins, 1D-*myo*-inosityl 2-amino-2-deoxy-α-D-glucopyranoside; Cys-GlcN-Ins, 1D-*myo*-inosityl 2-(L-cysteinyl)amido-2-deoxy-α-D-glucopyranoside; GSH, glutathione; HEPES, *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid; HPLC, high performance liquid chromatography; mBBr, monobromobimane; MSH and AcCys-GlcN-Ins, mycothiol or 1D-*myo*-inosityl 2-(*N*-acetyl-L-cysteinyl)amido-2-deoxy-α-D-glucopyranoside; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; RSmB, bimane derivative of corresponding thiol RSH; SAS, saturated ammonium sulfate; SDS, sodium dodecyl sulfate.

10.1021/bi012212u CCC: $22.00
Published on Web 05/10/2002

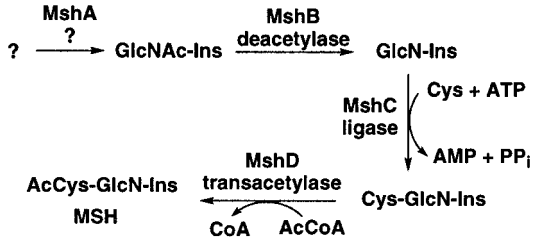

FIGURE 1: Biosynthesis pathway for mycothiol.

(Cys-GlcN-Ins) to product MSH (MshD) as illustrated in Figure 1. Production of GlcN-Ins from GlcNAc-Ins was shown to be catalyzed by a deacetylase (MshB) which was identified as coded by the open reading frame Rv1170 of *M. tuberculosis*, the first, and until now the only, mycothiol biosynthesis gene to be identified (*13*). For several years, we have sought to identify the ligase (MshC) utilizing both genetic and enzyme isolation protocols. The task proved surprisingly difficult, but we finally succeeded in purifying the enzyme from *M. smegmatis* and identifying the *mshC* gene of *M. tuberculosis*, as reported herein.

MATERIALS AND METHODS

Bacterial Culture. M. smegmatis strain mc^2155 was grown in Middlebrook 7H9 broth supplemented with 0.05% Tween 80 and 0.4% glucose at 37 °C and 250 rpm. After 28 h of cultivation, the bacterial cells were collected by centrifugation at 8000*g* for 15 min. The cell pellets were stored at −70 °C until used.

Reagents. Middlebrook 7H9 was purchased from Difco Laboratories, and glucose and Tween 80 were from Fisher. MSH was isolated from *M. smegmatis*, derivatized with monobromobimane (mBBr, Molecular probes) to form MSmB, and purified as described earlier (*14*). GlcN-Ins was prepared by hydrolyzing MSmB quantitatively, with purified *M. smegmatis* mycothiol *S*-conjugate amidase, as previously described (*11*). CysmB-GlcN-Ins was purified by preparative HPLC, after acid hydrolysis of MSmB, as described earlier (*15*).

Assays. A minor modification of the protocol described by Anderberg et al. (*15*) was used for routine measurement of ATP-dependent ligase activity. The enzyme activity was assayed in a final volume of 25 μL containing 12.5 μL of the ligase in different dilutions, 50 μM GlcN-Ins, 100 μM Cys (Calbiochem), 1 mM ATP (Sigma), 1 mM $MgCl_2$ (Fisher), 1 mM DTT (Calbiochem), 25 mM HEPES (Sigma), pH 7.5, and 35 μM each of the protease inhibitors (Sigma) *N*-α-*p*-tosyl-L-phenylalanylchloromethyl ketone and *N*-α-*p*-tosyl-L-lysinechloromethyl ketone. The mixture was incubated at 37 °C for 30 min. The assay was terminated by addition of 25 μL of 8 mM mBBr in acetonitrile and heating the mixture at 60 °C for 5 min to derivatize the thiols. The derivatization was quenched with the addition of 150 μL of 10 mM methanesulfonic acid and vortexing. HPLC analysis of CySmB-GlcN-Ins and CySmB was carried out by HPLC using a Beckman Ultrasphere IP (250 × 4.6 mm) analytical column fitted with a Brownley OD-GU 5 μ C-18 cartridge using the following linear gradients: 0 min, 100% A (0.1% TFA in water); 10 min, 100% A; 30 min, 80% A; 33 min, 100% B (7.5% methanol in acetonitrile); 36 min, 100% B; 38 min, 100% A; 50 min, 100% A (reinjection). The flow rate was 1 mL/min and fluorescence detection was accomplished with a 370 nm excitation filter and a 418−700 nm emission filter. The apparent V_{max} and K_m values were calculated by least-squares analysis of Eadie-Hofstee plots with initial rate data using KaleidaGraph 3.5 (Synergy Software).

Protein concentration was measured by the method of Bradford (*16*) using BSA as standard.

Purification of Ligase. All operations were carried out at 4 °C in the presence of 3 mM 2-mercaptoethanol unless stated otherwise. Cells (182 g wet weight) were suspended in 780 mL of 50 mM HEPES pH 7.5 in the presence of 35 μM of the protease inhibitors *N*-α-*p*-tosyl-L-phenylalanyl-chloromethyl ketone and *N*-α-*p*-tosyl-L-lysinechloromethyl ketone. The cells were disrupted by ultrasonication (Branson Sonifier 200) in an ice bath. The cell debris was removed by centrifugation at 100000*g* for 1 h at 4 °C. A solution of saturated ammonium sulfate (Fisher) was added to the supernatant to 15%, and the mixture was allowed to stand on ice for 2 h. After centrifugation at 28000*g* for 30 min, additional SAS was added to the supernatant to 50% saturation, and the mixture was stored overnight at 4 °C. After centrifugation at 28000*g* for 30 min, the protein pellet (32 g) was resuspended in 500 mL of 50 mM HEPES pH 7.5 and was desalted by passing it through Sephadex G-25 column (7.5 × 36 cm). The collected eluent (550 mL) was applied on a Toso Haas DEAE 650-M column (5.2 × 26 cm, 500 mL) equilibrated with 50 mM HEPES, pH 7.5. The enzyme was eluted with a linear gradient of 0−0.4 M NaCl in 16 column volumes of the buffer at 600 mL/h. The fractions (315−350) containing the third peak of enzyme activity were combined (670 mL) and were diluted 3-fold with Milli-Q water to lower the salt concentration. The diluted solution was applied to a hydroxyl apatite column (Bio-gel HTP from BioRad; 2.6 × 26 cm) at 120 mL/h, which was preequilibrated with 10 mM potassium phosphate (Fisher) buffer pH 6.8 containing 100 mM NaCl. The bound proteins were eluted at 240 mL/h with a linear gradient of 10−100 mM phosphate (100−0 mM NaCl) in 20 column volumes. The activity in fractions (75−125) was collected (800 mL) and diluted to 1800 mL with Milli-Q water to lower the salt concentration. The pH was adjusted to 6.4 with 1 M potassium dihydrogen phosphate. The diluted material was applied to a Reactive Brown 10 dye affinity column (Sigma, 1.5 × 11 cm, 20 mL) at 60 mL/h preequilibrated with 50 mM potassium phosphate buffer pH 6.5 and washed with buffer until no absorption was evident in the effluents (200 mL, 120 mL/h). The ligase was eluted with 50 mM phosphate buffer pH 8.0 at a flow rate of 120 mL/h. The fractions containing optimal ligase activity (12−64) were collected (104 mL). The protein was concentrated by adding solid ammonium sulfate to 80% saturation and allowing precipitation to continue on ice for 2 h. It was collected by centrifugation at 28000*g* for 30 min at 4 °C and resuspended in 50 mM HEPES pH 7.5 containing 150 mM NaCl and 3 mM 2-mercaptoethanol. This was loaded on a Sephadex G-100 column (Pharmacia, 1.5 × 95 cm, 170 mL) and eluted with the same buffer. The active fractions were collected (2 mL each) and analyzed for purity on 12.5% SDS−PAGE. Purified protein thus obtained was stored at 4 °C for further characterization.

Protein Sequencing. Amino acid sequencing was performed after electroblotting the two bands of active ligase

Table 1: Purification of *M. smegmatis* Cys:GlcN-Ins Ligase (MshC)

step	protein (mg)	total activity (nmol min^{-1})	specific activity (nmol min^{-1} mg^{-1})	yield (%)	purif factor
crude extract	9660	227	0.023	(100)	(1)
15−50% SAS	6900	1030	0.15	454	6.5
DEAE ion exchange	847	465	0.55	205	24
hydroxyl apatite	90	300	3.3	132	139
Reactive Brown 10 affinity chromatography	15	150	10	66	435
gel filtration G-100	0.2	10.8	54	4.8	2350

from 12.5% SDS−PAGE to a poly(vinylidene difluoride) membrane. The N-terminal amino acid sequence was determined on an Applied Biosystems model 494 Procise gas-phase protein sequencer at the UCSD Department of Biology Protein Sequencing Facility.

Cloning of the Ligase. M. tuberculosis H37Rv genomic DNA (kindly provided by Y. Av-Gay, University of British Columbia) was employed to amplify the *mshC* gene by PCR using the primers DF-1 (5′-GCGGATCCATGCAGTCGTG-GTATTGCCC-3′) and DR-1 (5′CCAAGCTTCTACAG-GTCCACCCCGAGCA-3′), and Platinum Pfx DNA polymerase (Gibco BRL). The PCR product was cloned into the *Bam*H I/*Hin*d III (New England Biolabs) sites of pRSETA (Invitrogen) using T4 DNA ligase (Gibco BRL). The expression plasmid was transformed into *E. coli* TOP10F′ (Invitrogen) and the presence of the insert verified by cleavage with *Bam*H I and *Hin*d III. The MshC ligase, which contains an N-terminal His_6 extension, was expressed in *E. coli* BL21 (DE3) plys S (Invitrogen) with induction overnight by 0.5 mM isopropyl-β-D-thiogalactopyranoside (Fisher) at 20 °C.

RESULTS

Purification of the Cys:GlcN-Ins ligase proved difficult and many preliminary studies were needed to finally obtain sufficiently pure protein for N-terminal sequencing. The major problems encountered were multiple interconverting forms of the enzyme activity which eluted separately on chromatography and instability of the enzyme activity in purification buffers. The method that was eventually successful depended upon selection of the correct peak of activity at each stage of purification and intensive effort to complete the protocol before activity was lost.

The first purification step (Table 1) involved a 15−50% ammonium sulfate fractionation and this resulted in a 6.5-fold increase in specific activity with very little loss of protein. After the sample was desalted on G-25, this material was chromatographed on DEAE 650-M. The elution profile (Figure 2A) revealed three major peaks of activity, the major portion of the activity residing in the first and third peaks. Experience revealed that the third peak was more readily susceptible to further purification, and the central fractions from this peak were combined for chromatography on hydroxyl-apatite (Figure 2B). Activity eluted primarily in one broad peak. The most active fractions were combined and bound to a Reactive Brown-10 dye affinity column at pH 6.5. After unbound protein was washed off, the column was eluted by increasing the pH to 8.0 over 5 column volumes (Figure 2C). The activity eluted at pH 8 in a sharp but tailing peak.

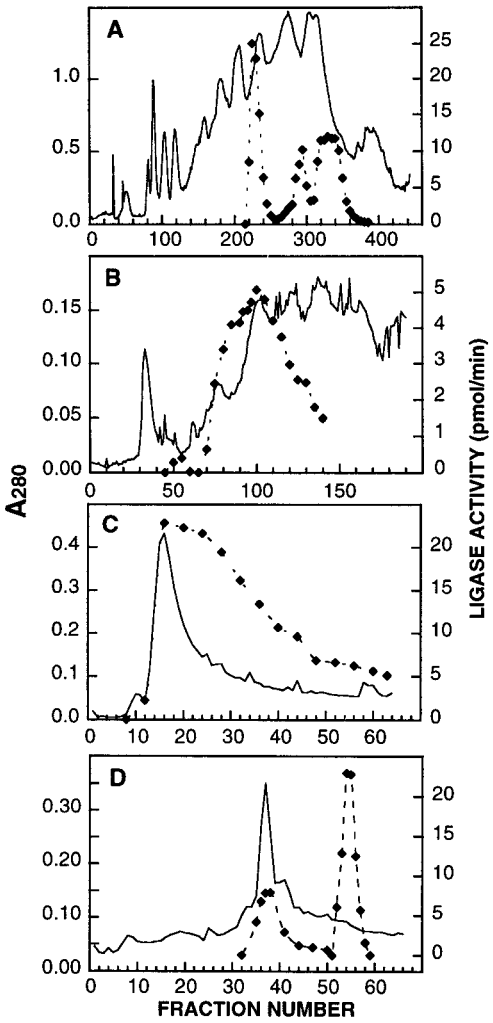

FIGURE 2: Chromatographic profiles for purification of MshC: (A) DEAE ion exchange; (B) hydroxyl apatite; (C) Reactive Brown 10 affinity chromatography; (D) Sephadex G-100 gel filtration.

The purification was monitored by SDS−PAGE (Figure 3), which suggested that only a few proteins of varying size were present after chromatography on the Brown-10 column. The peak activity from this column was pooled, precipitated with ammonium sulfate, and taken up in HEPES buffer pH 7.5 for chromatography on Sephadex G-100. The activity eluted in two peaks corresponding to M_r ~150 000 and ~89 000 (Figure 2D), the second peak exhibiting highest specific activity. Each fraction of the second peak was concentrated and examined by SDS−PAGE (Figure 3). Two closely migrating proteins of M_r ~ 47 000 were detected (Figure 3, lanes 10−14), with the larger protein rich in the earlier fractions (lanes 10 and 11) and the smaller protein dominant in the later fractions (lanes 13 and 14).

The smaller protein was submitted for N-terminal amino acid sequencing, and the sequence MQSWSAPAIPVVPGRG-PALR was obtained. When this sequence was BLAST searched against the TIGR website of unfinished *M. smegmatis* genomic sequence, a 100% identity match was obtained with the N-terminal sequence for an open reading frame found on contig 3267 (1222384−1223620) and encoding a protein of 412 amino acid residues. A BLAST search of the N-terminal sequence against the *M. tuberculosis* genome database (*17*) on Tuberculist retrieved the sequence assigned as CysS2 (Rv2130c) as the only one with significant

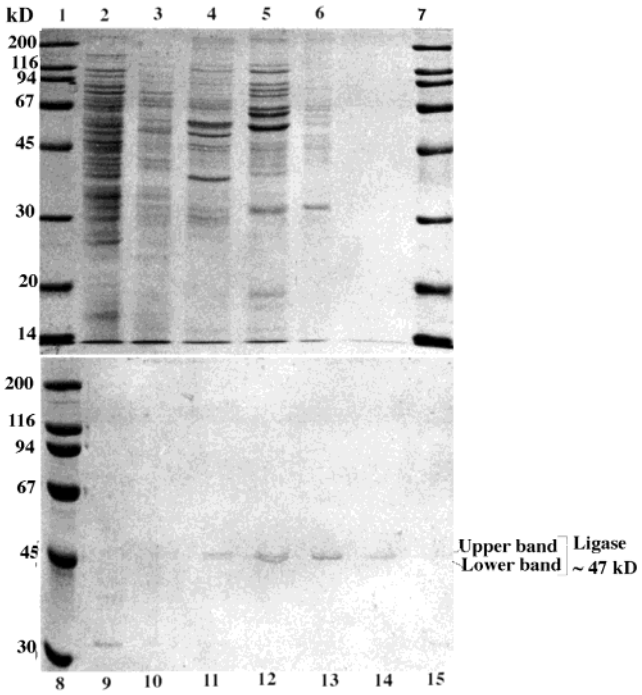

FIGURE 3: SDS−PAGE gels showing purification of MshC: (lanes 1, 7, and 8) Bio-Rad Broad Range molecular mass standards; (lane 2) crude cell free extract; (lane 3) 15−50% saturated ammonium sulfate extract; (lane 4) pooled fractions #315−350 from DEAE chromatography; (lane 5) pooled fractions #75−125 from hydroxyl apatite chromatography; (lane 6) pooled fractions #12−64 from Reactive Brown 10 chromatography; (lane 9) Sephadex G-100 fraction fraction 50 (lacking ligase activity) and (lanes 10−15) fractions 52−57, respectively, with purified ligase.

N-terminal sequence correspondence and revealed a 70% identity and 80% positive homology for the N-terminal 20 amino acid residues. The full sequences for the two proteins are compared in Figure 4.

The larger protein was also submitted for sequencing and produced the sequence (G/S/M)(E/Q)HLKVDAMQSW(S/D/P)APAIP. This sequence overlaps that of the smaller protein and when the *M. smegmatis* gene was examined upstream from the terminal Met of the smaller protein a matching sequence SEHLKVDAMQSWSAPAIP was found. However, no sequence other than CysS2 was found when Tuberculist was searched with this sequence and the upstream residues of the *M. tuberculosis cysS2* gene are not homologous to those of the *M. smegmatis* gene. Thus, the ligase of *M. smegmatis* may be translated in two forms, one having an N-terminal extension of eight residues starting with Ser and the other having Met as the N-terminal residue. Alternatively, the larger protein may have been translated and partially processed by proteases.

The remaining active enzyme fractions were utilized to estimate K_m and V_{max} values for Cys and GlcN-Ins from Eadie-Hofstee plots. With [ATP] = 1 mM and [GlcN-Ins] = 300 μM, the apparent K_m for Cys (5−400 μM, $n = 8$) was determined to be 40 ± 3 μM and the apparent V_{max} was 83 ± 3 nmol min^{-1} mg^{-1}. For [ATP] = 1 mM and [Cys] = 100 μM the apparent K_m for GlcN-Ins (5−600 μM, $n = 7$) was 72 ± 9 μM and the apparent V_{max} was 90 ± 7 nmol min^{-1} mg^{-1}.

Confirmation that *cysS2* of *M. tuberculosis* codes for MshC was obtained by using PCR to clone the gene into pRSETA. Sequencing of the cloned DNA verified that the cloning was accurate. The His_6-tagged protein was expressed in *E. coli* after induction with isopropyl-β-D-thiogalactopyranoside. Assay of the crude extract with 1 mM ATP, 0.1 mM Cys, and 50 μM GlcN-Ins gave 0.12 nmol min^{-1} mg^{-1} of ligase activity, whereas analogous measurements on *E. coli* transformed with the blank vector yielded no measurable ligase activity (<0.01 nmol min^{-1} mg^{-1}).

DISCUSSION

It appears that the difficulty in purifying the Cys:GlcN-Ins ligase derives in part from the presence of two different forms of the MshC protein and their tendency to oligomerize. Some of the different forms apparently elute separately on some chromatography columns but once separated can rearrange to regenerate multiple forms. The ligase activity eluted from Sephadex G100 as an apparent tetramer but the purified ligase was collected as what appeared to be partially resolvable α_2 and β_2 dimers. Since the *M. tuberculosis* gene did not exhibit upstream homology to the *M. smegmatis* gene in the region coding for the N-terminal extension of the larger protein, this problem may be unique to *M. smegmatis*. The significance of multiple forms of the enzyme under physiologic conditions is not clear and could not be explored with the limited amounts of purified enzymes available. This would likely be most conveniently explored if the two proteins can be cloned and purified separately.

The catalytic nature of the enzyme activity is consistent with its assignment as the ligase involved in MSH biosynthesis. The cellular concentration of Cys in *M. smegmatis* during exponential growth has been measured at ~0.15 μmol/g of residual dry weight (*15*), which corresponds to a cellular concentration of 50 μM assuming that the cellular water content is three times the residual dry weight. This is comparable to the apparent $K_m = 40$ μM for Cys determined here. The cellular level of GlcN-Ins in *M. smegmatis* was measured at 33 ± 14 μM (*15*), roughly half the K_m value of 72 ± 9 μM determined here. This value is significantly below the value of 140 ± 9 μM determined by Bornemann et al. (*12*) at [ATP] = 2 mM and [Cys] = 245 μM using a 30−50% SAS fraction prepared from a *M. smegmatis* crude extract supernatant.

Bacteria generally produce one aminoacyl-tRNA synthetase for each of the 20 natural amino acids, but some exceptions are known (*18*). However, in those cases where two aminoacyl-tRNA synthetases for the same amino acid have been identified, they have nearly identical size. This is not the case with CysS and CysS2 (MshC) which in *M. tuberculosis* are comprised of 475 and 414 residues, respectively. There are structural similarities and differences between CysS and MshC from *M. tuberculosis* (Figure 5). CysS lacks the N-terminal sequence (residues 1−19, Figure 4) partially conserved in the MshC genes. The CysS and the MshC genes possess the "HIGH" (*19*, *20*) and "KMSKS" (*21*, *22*) signatures characteristic of class I aminoacyl-tRNA synthetases (*18*). However, whereas the CysS proteins contain the actual HIGH sequence near residue 40 (Figure 5), the sequence is H(L/M)GH in the MshC proteins shown in Figure 4. The substitution of Leu or Met for Ile is found in several class I aminoacyl-tRNA synthetases (*18*). The CysS protein of *M. tuberculosis* has a 76-residue extension

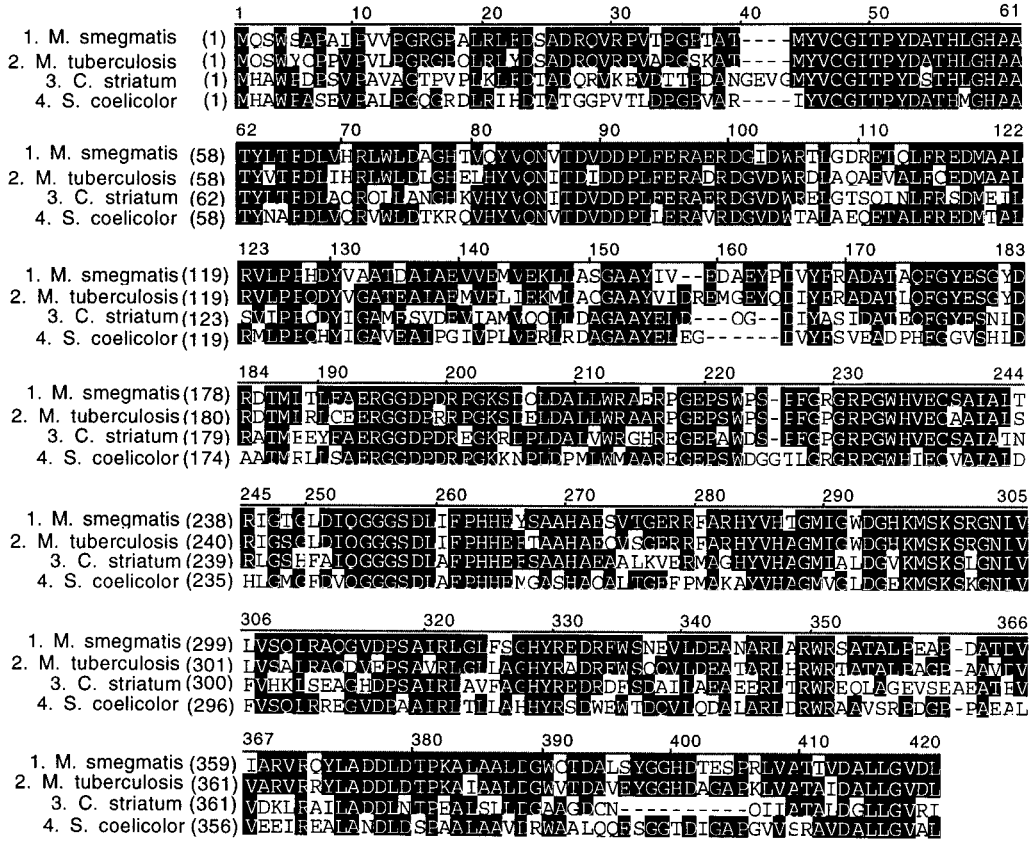

FIGURE 4: Sequence alignments for orthologs of MshC.

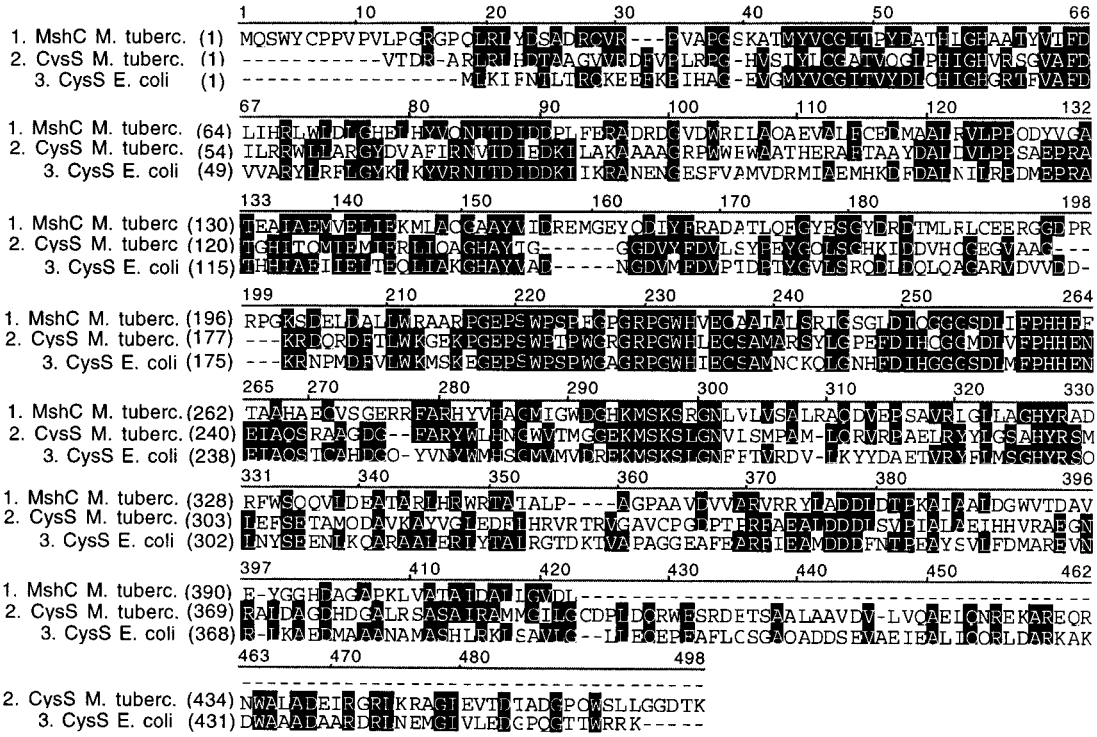

FIGURE 5: Sequence alignments of MshC and CysS of *M. tuberculosis* and CysS of *E. coli*.

beyond the C-terminus of MshC which has substantial identity (29%) with the corresponding 69 residue extension in the CysS from *E. coli* (Figure 5). Overall, the CysS of *M. tuberculosis* is 42% identical to the CysS of *E. coli* and 32% identical to MshC of *M. tuberculosis*. Finally, CysS has been normally found to be active as a monomer, whereas MshC is active in both dimer and tetramer forms. Thus, it is probable that CysS of *M. tuberculosis* is the true cysteinyl-tRNA synthetase and the protein originally classified as CysS2 is actually the Cys:GlcN-Ins ligase (MshC).

A BLAST search with the *M. tuberculosis* MshC sequence on GenBank revealed additional homologues in *Streptomyces coelicolor* A3 (*2*) (accession no. CAC36366) and *Corynebacterium striatum* (accession no. AAG03366). The se-

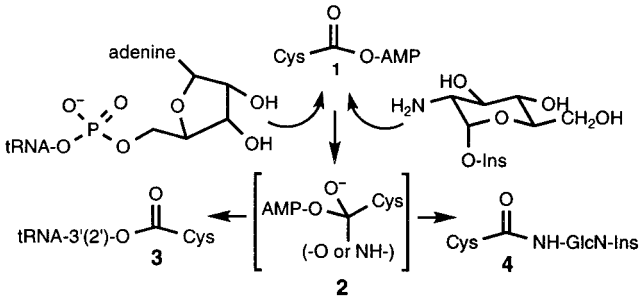

FIGURE 6: Chemistry of cysteinyl-tRNA synthetase and Cys:GlcN-Ins ligase reactions.

quences for these MshC proteins are included in Figure 4 and have 54−55% identity to the *M. tuberculosis* sequence. Moreover, orthologs of *M. tuberculosis* MshC were also found (tblastn) at the Sanger Institute in *M. leprae* (82% identity, ref 23), *M. bovis* (96% identity; www.Sanger.org/Projects/M_bovis), and *Corynebacterium diphtheriae* (54% identity; www.Sanger.org/Projects/C_diphtheriae), and at TIGR in *M. avium* (81% identity). All of these organisms belong to genera of bacteria that have been shown to produce MSH (*2*).

It seems logical that *mshC* originated from the *cysS* gene following a gene duplication. The enzyme mechanism of the aminoacyl-tRNA synthetases involves activation of cysteine to produce a cysteinyl-AMP intermediate (*18*). The activated cysteinyl group is subsequently transferred to the 2′ or 3′ ribosyl hydroxyl at the 3′ terminus of t-RNA^{cys} to produce the charged cysteinyl ester on t-RNA^{cys} (Figure 6). The initial step, formation of AMP−Cys (**1**), is the same for the ligase reaction. The ligase mechanism differs chemically by recognition of GlcN-Ins in place of the ribose of t-RNA^{cys} and in the attack of the amino group, in place of a hydroxyl group, upon AMP−Cys (Figure 6). A general base on the enzyme presumably functions to remove a proton from the hydroxyl or amino group leading to formation of a tetrahedral intermediate (**2**) which decomposes to form an ester (**3**), in the case of cysteinyl-tRNA synthetase, or an amide (**4**) in the case of the ligase. The later process is chemically more favorable owing to the greater nucleophilicity of amines over alcohols and the greater thermodynamic stability of amides relative to esters. These features, inherent in the chemistry of the reactions, make evolution of the ligase from a cysteinyl-tRNA synthetase a highly plausible route.

The ligase gene, *mshC*, is the second gene of MSH biosynthesis to be identified, the first being that coding for GlcNAc-Ins deacetylase, *mshB* (*13*). Identification of the genes for mycothiol biosynthesis is the first step in evaluating their utility as targets for development of new drugs and vaccines directed against infections by actinomycetes. An important next step is the characterization of mutants in which these genes have been inactivated and this is the subject of current studies.

ACKNOWLEDGMENT

We thank Sara Anderberg, Mamta Rawat, Ken Nguyen, Morrell Leghorn, Jr., and Kristyl Sasaki for their participation in the initial efforts to purify the ligase. We thank Jack Kyte, Yossef Av-Gay, and Robert Bartsch for helpful discussions concerning this project. We are indebted to The Institute for Genomic Research (http.//wwwq.tigr.org) for providing preliminary sequence data for *M. smegmatis* and *M. avium* (funded by the National Institute of Allergy and Infectious Diseases). We also thank the Sanger Institute, in collaboration with S. Cole and G. Hewinson, for making preliminary sequence data available for *M. bovis*, and, in collaboration with A. Efstratiou, for preliminary sequence data on *C. diphtheriae*, both projects funded by Beowulf Genomics.

REFERENCES

1. Fahey, R. C. (2001) *Annu. Rev. Microbiol. 51*, 333−356.
2. Newton, G. L., Arnold, K., Price, M. S., Sherrill, C., delCardayré, S. B., Aharonowitz, Y., Cohen, G., Davies, J., Fahey, R. C., and Davis, C. (1996) *J. Bacteriol. 178*, 1990−1995.
3. Newton, G. L., Bewley, C. A., Dwyer, T. J., Horn, R., Aharonowitz, Y., Cohen, G., Davies, J., Faulkner, D. J., and Fahey, R. C. (1995) *Eur. J. Biochem. 230*, 821−825.
4. Newton, G. L., Unson, M., Anderberg, S., Aguilera, J. A., Oh, N. N., delCardayre, S., Davies, J., Av-Gay, Y., and Fahey, R. C. (1999) *Biochem. Biophys. Res. Commun. 255*, 239−244.
5. Patel, M., and Blanchard, J. (1999) *J. Am. Chem. Soc. 120*, 11538−11539.
6. Patel, M. P., and Blanchard, J. S. (1999) *Biochemistry 38*, 11827−11833.
7. Patel, M. P., and Blanchard, J. S. (2001) *Biochemistry 40*, 3119−3126.
8. Norin, A., Van Ophem, P. W., Piersma, S. R., Persson, B., Duine, J. A., and Jornvall, H. (1997) *Eur. J. Biochem. 248*, 282−289.
9. Misset-Smits, M., van Ophem, P. W., Sakuda, S., and Duine, J. A. (1997) *FEBS Lett. 409*, 221−222.
10. Duine, J. A. (1999) *Biofactors 10*, 201−206.
11. Newton, G. L., Av-Gay, Y., and Fahey, R. C. (2000) *Biochemistry 35*, 10739−10746.
12. Bornemann, C., Jardine, M. A., Spies, H. S. C., and Steenkamp, D. J. (1997) *Biochem. J. 325*, 623−629.
13. Newton, G. L., Av-Gay, Y., and Fahey, R. C. (2000) *J. Bacteriol. 182*, 6958−6963.
14. Newton, G. L., Fahey, R. C., Cohen, G., and Aharonowitz, Y. (1993) *J. Bacteriol. 175*, 2734−2742.
15. Anderberg, S., Newton, G. L., and Fahey, R. C. (1998) *J. Biol. Chem. 273*, 30391−30397.
16. Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
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., Barrell, B. G., and et al. (1998) *Nature 393*, 537−544.
18. Meinnel, T., Mechulam, Y., and Blanquet, S. (1995) in *tRNA: Structure, Biosynthesis, and Function* (RajBhandary, D. S. a. U., Ed.) pp 251−292, ASM Press, Washington, DC.
19. Webster, T., Tsai, H., Kula, T. M., Mackie, G. A., and Schimmel, P. (1984) *Science 226*, 1315−1317.
20. Burbaum, J. J., Starzyk, R. M., and Schimmel, P. (1990) *Proteins 7*, 99−111.
21. Hountondji, C., Lederer, F., Dessen, P., and Blanquet, S. (1986) *Biochemistry 25*, 16−21.
22. Hountondji, C., Schmitter, J. M., Beauvallet, C., and Blanquet, S. (1990) *Biochemistry 29*, 8190−8198.
23. Cole, S. T., Eiglmeier, K., Parkhill, J., James, K. D., Thomson, N. R., Wheeler, P. R., Honore, N., Garnier, T., Churcher, C., Harris, D., Mungall, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R. M., Devlin, K., Duthoy, S., Feltwell, T., Fraser, A., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Lacroix, C., Maclean, J., Moule, S., Murphy, L., Oliver, K., Quail, M. A., Rajandream, M. A., Rutherford, K. M., Rutter, S., Seeger, K., Simon, S., Simmonds, M., Skelton, J., Squares, R., Squares, S., Stevens, K., Taylor, K., Whitehead, S., Woodward, J. R., and Barrell, B. G. (2001) *Nature 409*, 1007−1011.

BI012212U