Cloning, expression and rapid purification of active recombinant mycothiol ligase as B1 immunoglobulin binding domain of streptococcal protein G, glutathione-*S*-transferase and maltose binding protein fusion proteins in *Mycobacterium smegmatis*

Maria-Teresa Gutierrez-Lugo [a], Gerald L. Newton [b], Robert C. Fahey [b], Carole A. Bewley [a,*]

[a] *Laboratory of Bioorganic Chemistry, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, 9000 Rockville Pike, Bldg 8, Room 1A-02, Bethesda, MD 20892-0820, USA*
[b] *Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093, USA*

Abstract

Mycothiol ligase (MshC) is a key enzyme in the biosynthesis of mycothiol, a small molecular weight thiol found in *Mycobacteria* spp. and other actinomycetes. Mycothiol plays a fundamental role in these organisms by helping to provide protection from the effects of reactive oxygen species and electrophiles, including many antibiotics. It has recently been demonstrated that the MshC gene and more generally the production of mycothiol are essential to *Mycobacterium tuberculosis*, indicating that MshC may represent a novel target for new classes of antituberculars. Because MshC cannot be expressed heterologously in *Escherichia coli* and isolation from *Mycobacterium smegmatis* is impractical, we have optimized the *E. coli–M. smegmatis* shuttle vector pACE for cloning and recombinant expression of MshC (under control of an acetamidase-inducible promoter). To improve expression levels and simplify purification, we further constructed three N-terminal-MshC fusion proteins where N-terminal tags included the B1 domain of streptococcal protein G (to give GB1-MshC), glutathione-*S*-transferase (to give GST-MshC) and maltose binding protein (to give MBP-MshC), for expression in *M. smegmatis*. By expressing all three fusion proteins in a mutant strain of *M. smegmatis* mc^{2}155, namely I64 L205P MshC *M. smegmatis* which lacks mycothiol ligase activity, we demonstrate in vivo mycothiol ligase activity for each construct. Recombinant GST-MshC and MBP-MshC were isolated in one step by affinity chromatography in a yield of 0.7 and 1.2 mg fusion protein/L and exhibited specific activities of 9 nmol min^{-1} mg^{-1} and 25 nmol min^{-1} mg^{-1}, respectively.
© 2006 Elsevier Inc. All rights reserved.

Keywords: MshC; Mycothiol; Fusion proteins; Affinity purification; Mycothiol ligase; Complementation; *Mycobacterium tuberculosis*; Shuttle vector

Introduction

Mycothiol (MSH, AcCys-GlcN-Ins)[1] is the dominant low molecular weight thiol that is so far unique to actinomycetes [1,2] and has been shown to be present in all major groups of actinomycetes [1]. Mycothiol comprises one unit each of *N*-acetylcysteine (AcCys), 1-D-*myo*-inositol (Ins) and 2-amino-2-deoxy-D-glucopyranose (GlcN). Mycothiol is

---

* Corresponding author. Fax: +1 301 402 0008.
  E-mail address: caroleb@mail.nih.gov (C.A. Bewley).

[1] *Abbreviations used:* MSH, AcCys-GlcN-Ins, 1-D-*myo*-inosityl-2-(*N*-acetyl-L-cysteinyl) amido-2-deoxy-α-D-glucopyranoside; GlcN-Ins, 1-D-*myo*-inosityl-2-amino-2-deoxy-α-D-glucopyrano-side; Cys-GlcN-Ins, 1-D-*myo*-inosityl-2-(L-cysteinyl) amido-2-deoxy-D-glucopyranoside; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; GB1, B1 immunoglobulin binding domain of streptococcal protein G; GST, glutathione-*S*-transferase; MBP, maltose binding protein; DTT, dithiothreitol; SDM, site directed mutagenesis; DAD, diode array detector; API-ES, atmospheric pressure electrospray ionization; MSD, mass selective detector; ORF, open reading frame.
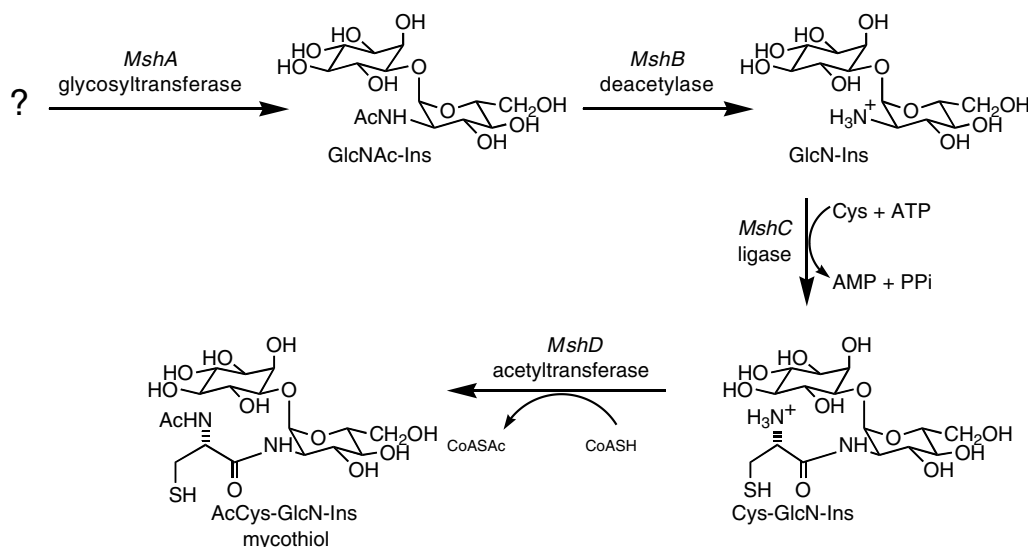
Fig. 1. Biosynthetic scheme of mycothiol.

assumed to be the functional equivalent of glutathione in mycobacteria and has been shown to be involved in protecting *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* from toxic oxidants and antibiotics [3,4]. Owing to the demonstration that mycothiol-deficient mutants become hypersensitive to some first line antituberculars, interest in identifying inhibitors of mycothiol biosynthesis and mycothiol-assisted detoxification has continued to rise [5].

Mycothiol biosynthesis occurs in at least four steps and involves the gene products of *mshA* through *mshD*. The third committed step in mycothiol biosynthesis involves transfer of L-cysteine (Cys) to GlcN-Ins to produce Cys-GlcN-Ins and is catalyzed by the enzyme mycothiol ligase, or *mshC* (L-Cys:1D-*myo*-inosityl 2-amino-2-deoxy-α-D-glucopyranoside ligase), in an ATP-dependent manner [2,6] (Fig. 1). Recently, the *mshC* gene has been shown to be essential for growth in *M. tuberculosis* by targeted gene disruption [7] and by transposon site hybridization [8]. When a second copy of *mshC* was inserted into the genome prior to targeted inactivation of the native *mshC* gene, viable mutants were isolated [7]. Since MshC has been shown to be required for MSH biosynthesis [4], these studies indicate that mycothiol is essential for growth in *M. tuberculosis* strains provoking further study of this enzyme as a therapeutic target.

The initial report describing the identification and characterization of MshC relied on laborious isolation methods from large scale cultures of *M. smegmatis* that were further hampered by the production of two MshC protein products that differed by addition of eight amino acids on the N-terminus [6]. In contrast to heterologous production of most proteins that can be expressed well in *Escherichia coli*, there is ample evidence that the proteins derived from *M. tuberculosis* are refractory to these conditions [9–12]. Because a prerequisite for screening small molecule libraries to identify inhibitors of MshC will rely on reproducible production of relatively large amounts of active enzyme, we have optimized

the *E. coli*-mycobacterial shuttle vector pACE [13,14] to provide a robust cloning and expression vector system for production of MshC using *M. smegmatis* as the expression host. Fusion to three unrelated proteins through the N-terminus of MshC increases production and greatly facilitates purification. We believe these expression vectors will be of general value for production of proteins derived from *M. tuberculosis*.

Materials and methods

Construction of mshC-tagged expression plasmids pACE/GB1-MshC, pACE/GST-MshC and pACE/MBP-MshC

The MshC gene was previously amplified from total genomic DNA of *M. tuberculosis* strain H37Rv and cloned into the *Bam*HI/*Hin*dIII sites of pRSETA (Invitrogen) [6]. The resulting plasmid pRSETA/MshC was used as a template for PCR amplification of MshC using *Taq* PCR master mix kit according to the manufacturer's instructions (Qiagen) together with an appropriate pair of forward and reverse primers designed for each of three fusion protein constructs. The primer sequences used are shown in Table 1 and numbered 1, 2 and 4. The amplified MshC gene products were subcloned into the *Bam*HI/*Xho*I sites of pGEV2 [15], *Nco*I/*Bam*HI of pET-42a (Novagen) or *Bam*HI/*Hin*dIII sites of pMALc2X (New England Biolabs) using a quick ligation kit (New England Biolabs) to generate constructs encoding GB1, GST-H$_6$-S and MBP fusion proteins, respectively. Note that insertion of the MshC gene into the *Nco*I/*Bam*HI sites of pET-42a generates a fusion construct in which GST and MshC are separated by a 6-His and S tag-linker, referred to above as GST-H$_6$-S and as GST from here on. The corresponding constructs, referred to as pGEV/MshC (for expression of GB1-MshC), pET42a/MshC-2 (for expression of GST-MshC after frame correction of pET42a/MshC by primer set 3 listed in Table 1) and pMALc2X/MshC (for expression of MBP-MshC), were subsequently used as tem-

Table 1
Summary of primers and templates used in this study and plasmids generated

Primer set	Primer ID	Primer sequence[a,b]	Restriction site	Template	Construct
1	5′GEV-MshC	5′-gcttatGGATCCatgcagtcgtgg-3′	*Bam*HI	pRSETA/MshC	pGEV/MshC
	3′GEV-MshC	5′-agttcaCTCGAGctacaggtccacc-3′	*Xho*I		
2	5′42a-MshC	5′-gcttatCCATGGatgcagtcgtgg-3′	*Nco*I	pRSETA/MshC	pET42a/MshC
	3′42a-MshC	5′-agttcaGGATCCctacaggtccacc-3′	*Bam*HI		
3	5′42aX-MshC[c]	5′-ggtattgagggacgcggg<u>ccatgg</u>atgcagtcgtgg-3′	N/A*	pET42a/MshC	pET42a/MshC-2
	3′42aX-MshC[c]	5′-ccacgactgcat<u>ccatgg</u>cccgcgtccctcaatacc-3′	N/A*		
4	5′GEV-MshC	5′-gcttatGGATCCatgcagtcgtgg-3′	*Bam*HI	pRSETA/MshC	pMALc2x/MshC
	3′MBP-MshC	5′-agttcaAAGCTTctacaggtccacc-3′	*Hin*dIII		
5	5′pACEMshC-1	5′-gataagagaaagggagtcca<u>catatg</u>ggatccatgcagtcgtgg-3′	*Nde*I	pACE/MshC	pACE/MshC-1
	3′pACEMshC-1	5′-ccacgactgcatggatcc<u>catatg</u>tggactccctttctcttatc-3′			
6	5′pACEMshC-2	5′-gaaaggcagtccacatatg<u>actagt</u>ggatccatgcagtcgtgg-3′	*Spe*I	pACE/MshC-1	pACE/MshC-2
	3′pACEMshC-2	5′-ccacgactgcatggatcc<u>actagt</u>catatgtggactccctttc-3′			
7	5′GB1	5′-gtcaggCATATGcagtacaagctt-3′	*Nde*I	pGEV	pACE/GB1-MshC
	3′GB1	5′-tacacgGGATCCacgcggaaccag-3′	*Bam*HI		
8	5′GST-MshCf	5′-aagcgcggatcccatatgtccct-3′	*Nde*I	pET42a/MshC-2	pACE/GST-MshC
	3′GST-MshCf	5′-tgctgtAACGTTctacaggtccaccccgagcag-3′	*Acl*I		
9	5′MBP	5′-cagcgaCATATGaaaatcgaagaa-3′	*Nde*I	pMALc2X/MshC	pACE/MBP-MshC
	3′GST-MshCf	5′-tgctgtAACGTTctacaggtccaccccgagcag-3′	*Acl*I		

[a] Capital letters indicate the added restriction sites in the PCR amplification.
[b] Underlined letters indicate the added restriction sites by site directed mutagenesis.
[c] Primers were designed to delete a single nucleotide located 5′ to the *Nco*I restriction site.
* Not applicable.

plates for PCR amplification of individual tags and/or the complete fusion protein sequences for cloning into the pACE shuttle vector (Table 1) [13,14].

Prior to insertion of genes encoding tags or complete fusion proteins into the vector, additional cloning sites *Nde*I and *Spe*I were introduced into pACE/MshC by site directed mutagenesis (Stratagene) using the primer sets 5 and 6 to produce pACE/MshC-1 and pACE/MshC-2, respectively. Separate plasmid preparations of pACE/MshC-2 were digested with *Nde*I/*Cla*I or *Nde*I/*Bam*H1 and dephosphorylated with Antarctic phosphatase (NEB) for cloning of *Nde*I/*Acl*I digested GST-MshC and MBP-MshC fragments, amplified from pET42a/MshC-2 and pMALc2x/MshC using the set of primers 8 and 9, respectively; or a *Nde*I/*Bam*HI digested GB1 fragment amplified from pGEV2 using primer set **7** (Table 1). Ligation of these fragments into complementary linear pACE/MshC-2 was performed using T4 DNA ligase (NEB) at 16 °C overnight. The ligation mixtures were purified by ethanol precipitation [16] and electroporated into Electromax *E. coli* DH10B (Invitrogen) according to the manufacturer's instructions. The constructs, identified as pACE/GB1MshC, pACE/GSTMshC and pACE/MBPMshC, were verified by restriction analysis and DNA sequencing (Fig. 2). All sequences have been deposited in GenBank.

Transformation and isolation of recombinant I64 M. smegmatis strains

The different expression plasmids were electroporated into the I64 L205P MshC *M. smegmatis* mutant (referred to as I64 *M. smegmatis*) that is devoid of cysteine ligase activity and therefore mycothiol production [4]. Preparation of I64 *M. smegmatis* competent cells and electroporation procedures were performed following published protocols [17]. Briefly, I64 *M. smegmatis* was grown in 7H9/OADC/Tween (comprising Middlebrook 7H9 (DIFCO) media supplemented with 10% OADC (Oleic acid, albumin, dextrose, catalase, BBL) and 0.05% Tween 20) until midlog phase. After harvesting, cells were washed twice with cold 10% glycerol by resuspension and centrifugation. I64 *M. smegmatis* competent cells in 10% cold glycerol (0.06 ml) were transformed by electroporation in the presence of 1 μg of DNA (1.2 kV) in prechilled 1 mm gap cuvettes using a Multiporator®/Electroporator 2510 (Eppendorf, Hamburg, Germany). Cells were harvested with 1 ml of warm 7H9/OADC/Tween and incubated for 2 h at 37 °C with shaking (225 rpm). An aliquot (25 μl) of this culture was plated on 7H9/OADC/Tween agar plates containing 75 μg/ml hygromycin (Invitrogen). Plates were incubated at 37 °C for three days to obtain the transformed recombinant strains.

Expression of recombinant GB1-MshC, GST-MshC and MBP-MshC fusion proteins

The selected I64 *M. smegmatis* recombinant strains were grown in 25 ml 7H9/OADC/Tween containing 75 μg/ml hygromycin for 72 h at 37 °C with shaking (225 rpm). A 1-ml aliquot of this starter culture was transferred to 100 ml of 7H9 containing 1% glucose, 0.05% Tween 20 and 75 μg/ml hygromycin and incubated overnight at
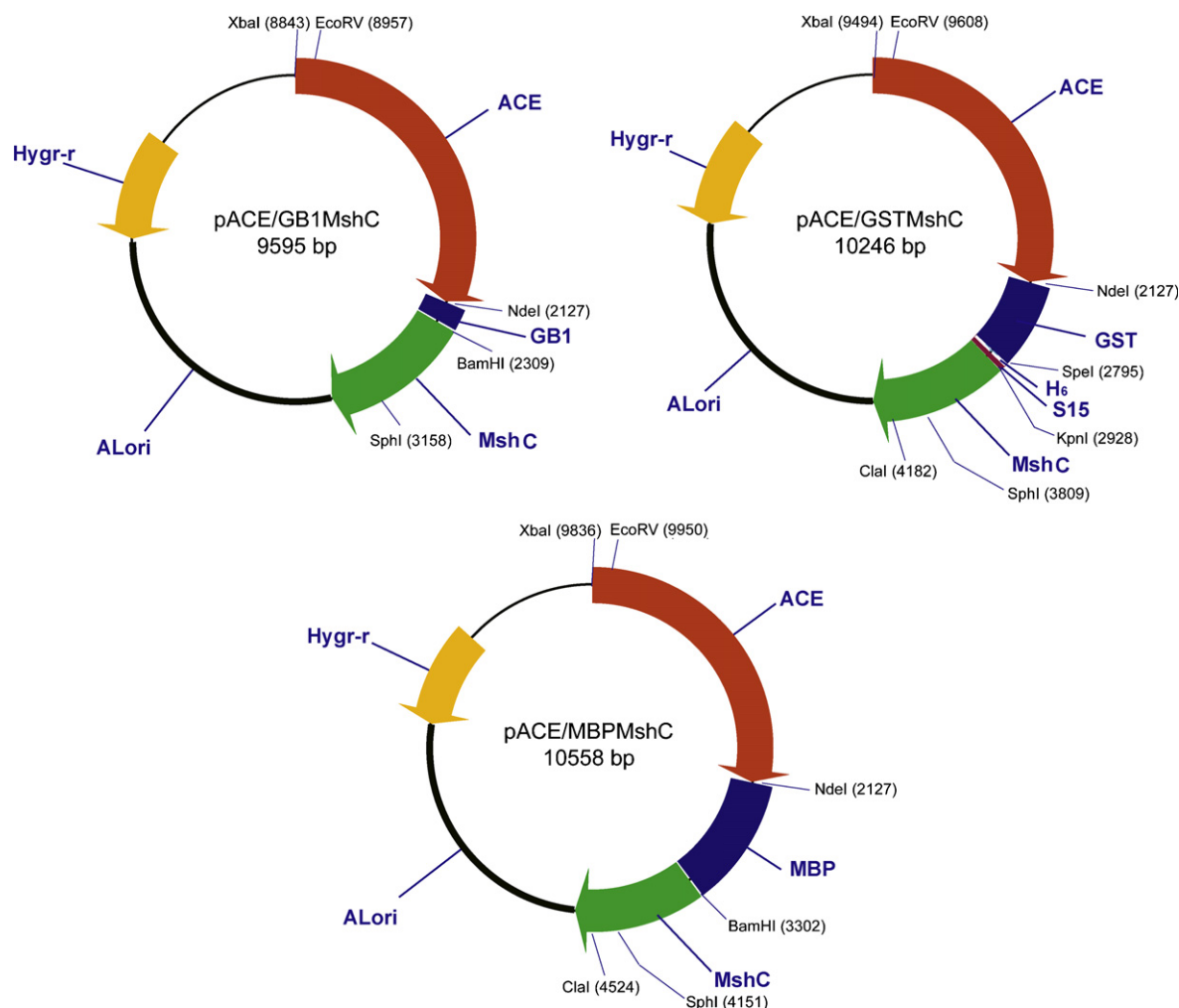
Fig. 2. Maps of mycobacterial expression vectors generated in this study encoding N-terminal-tagged MshC. Unique restriction sites are labeled and their locations indicated in parentheses. Proteins are labeled in bold and denoted by thick lines on the maps with the N-terminal tag and MshC colored blue and green, respectively.

37 °C with shaking. After incubation, cells were harvested by centrifugation at $8000g$ for 15 min, 4 °C (Beckman Avanti J-20 XP centrifuge) and resuspended to an A_{600} of 0.3 in fresh 7H9 medium containing 0.4% acetamide (Sigma), 0.05% Tween 20 and 75 µg/ml hygromycin (referred to as expression medium) and incubated at 37 °C with shaking (225 rpm). Cells were harvested after 22–26 h by centrifugation ($8000g$, 4 °C for 15 min) and stored at −80 °C until further use.

Frozen cell pellets were thawed and resuspended to 25% w/v cell suspension in cold lysis buffer containing 50 mM HEPES, pH 7.5 (Cellgro), 3 mM β-mercaptoethanol (Aldrich), 5 mM MgCl$_2$ (Quality biological Inc.) and protease inhibitors (Complete, EDTA-free protease inhibitor cocktail tablets, Roche). The cells were disrupted by ultrasonication in an ice bath using an Ultrasonic Processor XL sonicator (Misonix Inc.) for 7 cycles (3 s on, 3 min off) each. Cell lysates were centrifuged at $3000g$ for 1 h at 4 °C, and supernatants were transferred to clean tubes. Cell lysates were analyzed by SDS–PAGE for protein expression (Phastgels™ Gradient 10–15, Amersham Biosciences) and

visualized by staining with Coomassie Blue (Coomassie™ Blue R-350, Amersham Biosciences).

Detection of ligase activity by means of mycothiol production by recombinant I64 M. smegmatis pACE/GB1-MshC, pACE/GST-MshC and pACE/MBP-MshC strains

I64 *M. smegmatis* recombinant strains were grown in expression medium (100 ml) under the conditions described above. The cultures were incubated for up to 55 h and aliquots of 1–12 ml (depending on cell density) were taken during incubation to give cell pellets of approximately 10 mg wet weight. Each aliquot was transferred to a tared 1.5 ml microcentrifuge tube and cells pelleted at $12,100g$ for 10 min at 4 °C (Sorvall Biofuge Fresco); cell pellets were stored at −20 °C until further use.

Pellets were extracted and analyzed for *total* mycothiol content, and mycothiol production was quantified employing a modified method described previously by Newton et al. [1]. Briefly, cell pellets were lysed by vortexing in 160 µl of reducing lysis buffer (50 mM HEPES 7.5 in 50%

CH₃CN:H₂O, 4 mM DTT) followed by incubation at 60 °C for 30 min with constant vortexing. Reduced thiols were derivatized by addition of 40 μl of 100 mM monobromobimane in acetonitrile and kept at 60 °C for 10 min with occasional vortexing, and then placed on ice and quenched with 75 μl of 0.5 M methanesulfonic acid (Fluka). Samples were centrifuged at 12,100g at 4 °C for 10 min. Supernatants (2 μl) were analyzed for total mycothiol content by LC–MS plus fluorescence detection using the following chromatographic conditions: Waters Symmetry C18 5 μm column (4.6 × 150 mm) fitted with a QR 5 μm C18 guard cartridge (4.6 mm × 1 cm), gradient elution (flow rate of 0.5 ml min⁻¹) of 5% MeOH/95% 0.15% formic acid in water to 100% MeOH in 15 min followed by a 2-min wash with 100% MeOH. The column was reequilibrated with 5% MeOH/95% 0.15% formic acid for 3 min prior to each injection. Detection of derivatized thiols was accomplished using a fluorescence detector (Agilent Technologies) with excitation at 398 nm and emission at 480 nm. Under these chromatographic conditions mycothiol bimane eluted at 9.9 min (*m/z* 677, 699, 715). Total mycothiol content (μmol/g wet weight) in each aliquot was calculated by comparing peak areas against a mycothiol calibration curve.

Purification and characterization of GST-MshC and MBP-MshC from recombinant I64 M. smegmatis strains

To obtain recombinant GST-MshC or MBP-MshC fusion proteins, respective I64 *M. smegmatis* strains I64pACE/GST-MshC and I64pACE/MBP-MshC were grown in 4 × 500 ml (1 L flasks) of expression medium (described above) inoculated with 3–4 ml each of a starter culture grown in 7H9 supplemented with 1% glucose, 0.05% Tween 20 and 75 μg/ml hygromycin. The cultures (A_{600} 0.3) were incubated at 37 °C (225 rpm) for 22–26 h and cell lysates were prepared by sonication at 4 °C for 7 cycles (3 s on, 3 min off) each, followed by centrifugation at 4 °C for 1 h at 18,000g. All subsequent purification steps were carried out at 4 °C.

The resulting protein extracts were diluted 1:1 v/v with appropriate binding buffer before loading onto an affinity resin column equilibrated in the same buffer. For GST-MshC, supernatants were diluted with GST bind/wash buffer (Novagen) and loaded onto a column containing 12 ml GST Sepharose™ affinity resin (Novagen, 1.6 × 7.5 cm). MBP-MshC supernatants were diluted with 50 mM HEPES, 200 mM NaCl, pH 7.5, and loaded onto a column containing 20 ml amylose resin (New England Biolabs, 2.6 × 5 cm). Using an ÄKTAprime plus chromatographic system (GE Healthcare), columns were washed with 10 column volumes of binding buffer followed by elution with three column volumes of 10 mM reduced glutathione (Novagen) in binding buffer for GST-MshC (flow rate 0.8 ml min⁻¹) or 10 mM maltose (Sigma) in 50 mM HEPES, pH 7.5, 200 mM NaCl for MBP-MshC (flow rate of 1 ml min⁻¹). All binding and elution buffers also contained 5 mM MgCl₂, 3 mM β-mercaptoethanol and 5 mM benzamidine (Sigma) as protease inhibi-

tor. Fractions (2 ml) were collected during the elution step and analyzed by SDS–PAGE for detection of fusion proteins. Fractions containing the desired recombinant protein were pooled and stored as 20% glycerol stocks at −80 °C. Protein concentrations were determined by A_{280} measurements and calculated using theoretical extinction coefficients of 127,825 M⁻¹ cm⁻¹ for GST-MshC and 145,565 M⁻¹ cm⁻¹ for MBP-MshC (http://us.expasy.org/tools/protparam.html) [18].

Identity of the purified recombinant proteins was confirmed by mass spectrometry analysis using an Agilent Technologies LCMS equipped with DAD detector and API-ES ionization source. Proteins were eluted from a C3 reverse phase column (Zorbax 300SB-C3, 2.1 × 150 mm, 5 μm, Agilent) using a linear gradient of 5–95% acetonitrile in 5% acetic acid over 25 min at a flow rate of 0.2 ml min⁻¹ at 40 °C, and detected by UV at 280 nm and positive API-ES. The retention times for GST-MshC and MBP-MshC under these conditions were 41.1 and 40.5 min, respectively. Deconvolution of the mass spectra indicated molecular weights of 77.28 kDa and 88.68 kDa, in excellent agreement with the predicted masses of 77.2758 and 88.6791 kDa.

Detection of mycothiol ligase activity

Mycothiol ligase activity of the I64 *M. smegmatis* transformed strains was measured by a modified method reported by Sareen et al. [6]. Enzyme activity was assayed in 0.2 ml micro tubes at a final volume of 25 μl containing 25 mM HEPES, pH 7.5, 100 μM Cys (Calbiochem), 50 μM GlcN-Ins, prepared by cleavage of mycothiol bimane with mycothiol-*S*-conjugate amidase [19], 1 mM ATP (Sigma), 1 mM MgCl₂, 1 mM DTT (ICN Biomedicals), and 12 μl of different dilutions of protein extracts or affinity-purified fractions. The reaction mixture was incubated at 37 °C for 30 min. After incubation, thiols were derivatized by addition of 25 μl of 8 mM monobromobimane (Sigma) in acetonitrile and incubated at 60 °C for 10 additional minutes. Finally, reactions were quenched by addition of 50 μl of cold 20 mM methanesulfonic acid. The resulting mixture (5–10 μl) was analyzed by LCMS using a Waters Symmetry C18 5 μM column (4.6 × 150 mm) fitted with a QR 5 μm C18 cartridge (4.6 mm × 1 cm). The gradient elution using 5 mM ammonium formate, pH 6.5, and MeOH consisted of four steps at a flow rate of 0.5 ml min⁻¹: 5–12% MeOH in 4 min, 12–25% MeOH in 6 min, 25–100% in 5 min, followed by a 3-min wash with 100% methanol. Detection of derivatized thiols was accomplished using a fluorescence detector with excitation at 398 nm and emission at 480 nm. Under these chromatographic conditions Cys-GlcN-Ins bimane showed a retention time of 14.5 min (*m/z* 635, 657, 673).

Results and discussion

Construction of MshC expression plasmids

The production of recombinant MshC is a major prerequisite for the screening of potential mycothiol ligase inhibi-

tors. However, the isolation of mycothiol ligase from *M. smegmatis* proved to be an arduous task [6] such that isolation of sufficient quantities of active enzyme from *M. smegmatis* would be impractical. We therefore sought to generate a strain that expresses significant amounts of soluble and active recombinant mycothiol ligase that is amenable to affinity purification. Expression of a target protein fused with another protein of intrinsically high solubility is a well known strategy to produce soluble recombinant proteins [20–22]. We began these efforts by first cloning the MshC gene (Rv2130c, 1245 bp) from *M. tuberculosis* into pGEV2 [15], pET42a and pMALc2X plasmids under the control of T7 or *tac* promoters to generate respective fusion proteins GB1-MshC, GST-MshC and MBP-MshC for expression in *E. coli*. However, despite enhanced expression in *E. coli* relative to the MshC gene on its own, all fusion proteins were produced primarily as insoluble inclusion bodies. Attempts to improve expression and solubility of these proteins, including use of codon optimized *E. coli* strains and induction with varying amounts of isopropyl β-D-thiogalactoside (IPTG) at different cell densities alone or in combination with varying temperatures, were unsuccessful.

Recombinant mycobacterial proteins have proved notoriously difficult to express in *E. coli* [9–12]. *Mycobacteria* have a high GC content (65.6% in *M. tuberculosis*) and unique codon preferences, and the encoded proteins tend to be rich in glycine, alanine, proline and arginine, all of which can be detrimental for overexpression in *E. coli* [9]. Since *E. coli* was not found to be a suitable host for the production of recombinant MshC, we investigated the possibility of producing the same recombinant fusion proteins in the fast growing non-pathogenic bacterium *M. smegmatis*. For this study we selected the I64 *M. smegmatis* mutant as a host strain for expression of these MshC fusion proteins. Because this mutant lacks MshC activity and is deficient in mycothiol production [4], we foresaw the potential of using mycothiol production as an indicator of our constructs' activity. To generate mycobacterial compatible expression vectors, the constructs generated for expression of tagged MshC in *E. coli*, namely pGEV/MshC (GB1-MshC), pET42a/MshC-2 (GST-MshC) and pMALc2X/MshC (MBP-MshC), were next employed as templates for the amplification of the fragments to be cloned into the mycobacterial expression vector pACE (8191 bp) [13,14]. The set of expression vectors were generated by either inserting a PCR product of MshC with its N-terminal tag into pACE or by inserting the N-terminal tag alone into pACE/MshC. The first strategy proved successful in the construction of the GST-MshC and MBP-MshC fusion proteins while for GB1-MshC it was necessary to insert the GB1 fragment into the pACE/MshC plasmid.

In all cases, MshC was first cloned into the two unique cloning sites *Bam*HI/*Cla*I of pACE. Once the plasmid pACE/MshC (9412 bp) was obtained, two additional unique cloning sites, *Nde*I and *Spe*I, were introduced by two separate site directed mutagenesis reactions. The result-

ing vector named pACE/MshC-2 (9424 bp) was then digested with *Nde*I and *Cla*I (for GST-MshC and MBP-MshC) and with *Nde*I/*Bam*HI (for introduction of GB1). It is important to note that the MshC gene also contains a *Cla*I restriction site at position 1217. For purposes of keeping the MshC gene intact, the GST-MshC and MBP-MshC PCR products were generated with an *Acl*I restriction site which is compatible with *Cla*I. After cloning into this site it is converted into a non-cleavable *Acl*I/*Cla*I hybrid. However, when pACE/MshC is digested with *Cla*I, there remains a possibility that the restriction enzyme will cut MshC within the gene at position 1217. Consequently, even after gel purification, it is possible that some of the linear vector may contain a small fragment of the MshC gene corresponding to base pairs 1217–1245. Sequencing of the GST-MshC and MBP-MshC expression plasmids showed that neither contains the additional fragment. The resulting expression vectors were identified as pACE/GB1MshC (9595 bp), pACE/GSTMshC (10246 bp) and pACE/MBPMshC (10588 bp) (Fig. 2).

Expression of tagged mycothiol ligase constructs in I64 M. smegmatis and evaluation of their ligase activity

With the intent of improving protein expression and simplifying the purification of recombinant MshC (45.6 kDa), the MshC gene from *M. tuberculosis* (Rv2130c, 1245 bp) was fused at the N-terminus with GB1 (180 bp, 6.7 kDa), GST (837 bp, 31.7 kDa, which includes C-terminal 6-His and S-tags) and MBP (1173 bp, 43.0 kDa) tags using the pACE vector that controls protein expression through the acetamidase promoter. pACE was previously generated by amplification of the acetamidase gene from *M. smegmatis* [13,14]. The acetamidase promoter consists of 4 ORFs identified as AmiS (642 bp), AmiD (285 bp), AmiA (429 bp) and AmiC (1074 bp) that are involved in regulation and transport. The acetamidase from *M. smegmatis* is an inducible enzyme and enables the organism to use several amides, such as acetamide as sole carbon sources [23,24]. Upon addition of 0.4% acetamide to the medium, the three fusion proteins were expressed at maximum levels after 22–26 h of incubation and were present in the soluble fraction. Analysis of the resulting protein extracts on SDS–PAGE indicated that the three fusion proteins migrated at the expected molecular weight and that the fusion tags are well tolerated in the I64 *M. smegmatis* mutant (Fig. 3). Expected molecular weights of GB1-MshC, GST-MshC and MBP-MshC fusion proteins are 52.4 kDa, 77.3 kDa and 88.7 kDa, respectively.

In addition to the fusion proteins, it appears that upon induction with acetamide, each of the three transformed strains expresses two other major proteins migrating at ~45 kDa and ~60 kDa (Fig. 3, lanes 2–6). The band at ~45 kDa may represent acetamidase, whose molecular weight is calculated to be 43 kDa, as there is precedent for levels of this enzyme increasing up to 100-fold in the presence of suitable substrate such as acetamide [23,24]. In the
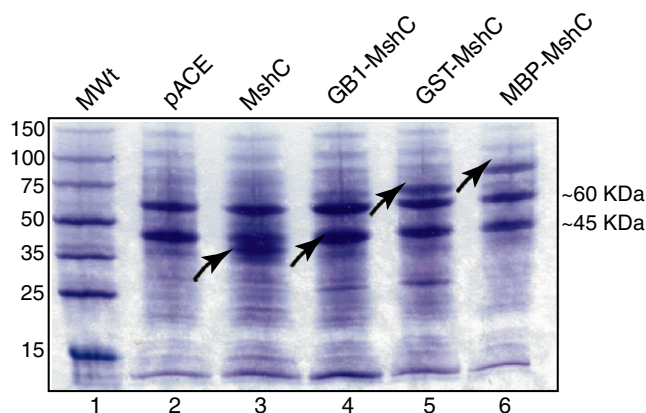
Fig. 3. SDS–PAGE analysis of protein extracts. Lanes 1–6 contain a molecular weight marker and protein extracts of I64pACE, I64pACE/ MshC, I64pACE/GB1MshC, I64pACE/GSTMshC and I64pACE/ MBPMshC, respectively. Molecular weights in kDa are labeled to the left of lane 1 and acetamidase-induced MshC constructs are marked with arrows. The strong bands at approximately 45 and 60 kDa that are present in lanes 2–6 appear upon induction with acetamide.

Table 2
Specific activity of protein extracts and purified fractions obtained from selected strains in expression medium[a]

Strain/construct	Specific activity ($nmol\ min^{-1}\ mg^{-1}$)
M. smegmatis mc^{2}155	0.06
M. smegmatis I64pACE	nd
M. smegmatis I64pACE/MshC	0.2
M. smegmatis I64pACE/GB1-MshC	0.1
M. smegmatis I64pACE/GST-MshC	0.3
M. smegmatis I64pACE/MBP-MshC	0.3

[a] Specific activity of supernatants is based on total protein determined by the Bradford protein assay.

case of the GB1-MshC construct, this fusion protein migrated together with the band of 45 kDa (Fig. 3, lane 4). However, expression of GB1-MshC was unambiguously demonstrated by measurement of its ligase activity (Table 2) and by complementation of the mycothiol deficient I64 mutant (Fig. 4). Ligase activity for all protein extracts is summarized in Table 2. All three of the fusion proteins showed ligase activity and the highest levels were observed for GST-MshC and MBP-MshC, both of which exhibited specific activities of $0.3\ nmol\ min^{-1}\ mg^{-1}$. These results indicate that the fusion tags do not interfere with the activity of mycothiol ligase. However, if desired each of the fusion proteins can be cleaved within the linker separating MshC and its N-terminal tag. The GB1 and MBP fusion proteins contain thrombin and factor Xa cleavage sites, respectively, while the GST fused protein contains both.

Complementation of the mycothiol deficient I64 M. smegmatis mutant

The I64 *M. smegmatis* mutant used in this study was generated previously by chemical mutagenesis of *M. smegmatis* mc^{2}155. The I64 mutant strain was shown to be deficient in mycothiol production producing only 1% of the
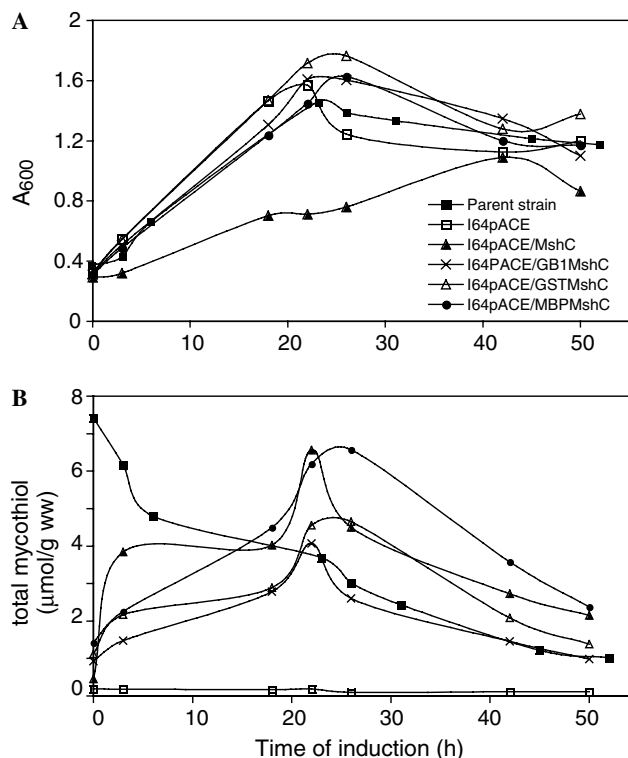

Fig. 4. Complementation of the mycothiol deficient I64 *M. smegmatis* mutant. (A) Growth rate measured as cell density (A_{600}) versus time (h). (B) Total mycothiol content of wild type mc^{2}155 and I64 *M. smegmatis* recombinant strains after transfer to expression medium and induction with acetamide. Starter cultures of all strains were grown in nutrient-rich 7H9/OADC/Tween prior to inoculating a larger volume of expression medium.

amount of mycothiol as its parent strain (0.1 μmol/g versus 10 μmol/g dry weight, respectively), and I64 *M. smegmatis* exhibits increased sensitivity to free radicals, alkylating agents and a variety of antibiotics. MshC activity in this mutant strain was shown to be $\leqslant 2\%$ that of its parent strain. Sequencing of the *mshC* gene revealed that a single amino acid substitution, Leu 205 to Pro, is responsible for the reduced mycothiol production and its associated phenotype [4]. The I64 *M. smegmatis* strain was thus chosen for expression of recombinant *M. tuberculosis* MshC to prevent any interference from endogenous wild type *M. smegmatis* MshC and to facilitate complementation experiments.

To determine whether the recombinant MshC fusion proteins were active, I64 pACE/GB1-MshC, I64 pACE/ GST-MshC, and I64 pACE/MBP-MshC were transformed into I64 *M. smegmatis* and acetamidase-induced mycothiol ligase activity and mycothiol production of these recombinant strains was measured in parallel with those of the I64 *M. smegmatis* mutant (I64pACE), its parent strain *M. smegmatis* mc^{2}155, and non-tagged MshC (I64 pACE/ MshC). The levels of mycothiol of all the strains were determined at multiple stages of growth (up to 55 h) in expression medium under antibiotic pressure, except for the parent strain mc^{2}155 which was grown in the absence of

antibiotics. The I64pACE strain produced on average around 0.1 μmol of total mycothiol per gram wet weight, averaging less than 5% of the levels produced by the parent strain during growth (Fig. 4). In general, all strains containing constructs of tagged and non-tagged MshC produced mycothiol levels between 2 and 69 times higher (depending on time from induction) than the I64pACE mutant, or levels at least as great as those of the parent strain grown under identical conditions (Fig. 4B). The three different recombinant strains had a maximum peak of mycothiol production at 22–26 h after induction. After this point, the levels of mycothiol decline in parallel with the growth curve, most probably due to the consumption of the acetamide used as a source of nitrogen and carbon (Fig. 4A). These results indicated that the three fusion proteins are functional in vivo and complement the ligase activity in the I64 *M. smegmatis* mutant. Of the three recombinant strains, I64pACE/MBP-MshC produced the highest levels of mycothiol, namely 6.6 μmol/g wet weight at 26 h, which is a 2.2-fold increase relative to the parent strain.

Purification of rGST-MshC and rMBP-MshC

On the basis of the observed complementation of the mycothiol deficient I64 *M. smegmatis* mutant by the three constructs (Fig. 4) and evaluation of their ligase activity (Table 2), I64pACE/GST-MshC and I64pACE/MBP-MshC were selected for affinity purification of recombinant protein. These recombinant strains were grown on a 2-L scale and processed as described in "Materials and methods" section. The entire protocol for expression and purification required 5 days for growth and induction and approximately 8 h to generate cleared cell lysates and to purify the proteins by affinity chromatography on glutathione or amylose resin columns. As seen in the SDS–PAGE denaturing gels in Fig. 5, affinity purifications of the GST- and MBP-tagged MshC proteins were very effective and yielded fractions containing fusion proteins of the expected molecular weights $M_r \sim 77$ and ~ 88 kDa that were at least 90% pure (as determined by SDS–PAGE and LC–MS). Starting from a 2-L culture of *M. smegmatis* I64pACE/GST-MshC or I64pACE/MBP-MshC, 0.7 mg/L of pure rGST-MshC and 1.2 mg/L of pure rMBP-MshC were obtained (Table 3). The molecular weights of the affinity-purified GST-MshC and MBP-MshC fractions were confirmed by LC–MS giving observed masses of 77.28 kDa and
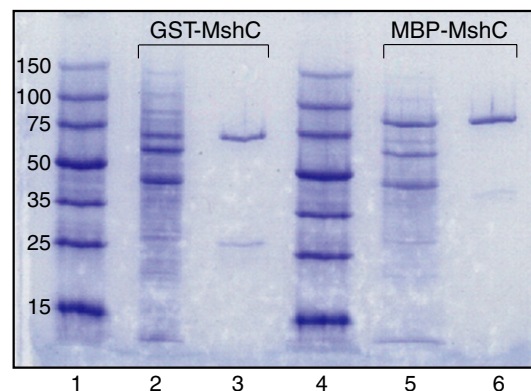

Fig. 5. SDS–PAGE analysis of protein extracts and affinity-purified recombinant GST-MshC and MBP-MshC. Lanes 1 and 4, molecular weight markers (kDa); lanes 2 and 5, protein extracts of I64pACE/GST-MshC and I64pACE/MBPMshC; lanes 3 and 6, affinity-purified GST-MshC and MBP-MshC.

88.68 kDa, respectively, in excellent agreement with the calculated molecular weights of 77,275.8 Da and 88,679.1 Da. Evident from the SDS–PAGE and detectable by LC–MS analysis, two minor species with respective masses of 31.8 and 42.9 kDa are observed in the GST-MshC and MBP-MshC affinity-purified fractions (Fig. 5, lanes 3 and 6). As the predicted masses of GST and MBP are 31.8 and 42.7 kDa respectively, these impurities likely correspond to the GST and MBP tags as a result of proteolysis within the spacer separating the tags and MshC.

Mycothiol ligase activities of total protein extracts and combined affinity-purified fractions for rGST-MshC and rMBP-MshC were measured using the protocol described in "Materials and methods" section and yielded final specific activities of 9 nmol min^{-1} mg^{-1} for purified rGST-MshC and 25 nmol min^{-1} mg^{-1} for purified rMBP-MshC. Because the greatest recovery and specific activity was observed with rMBP-MshC (Table 3), this construct will likely be chosen for larger scale protein production and used in screening efforts aimed at identifying small molecule inhibitors of MshC. In anticipation of such efforts, we have monitored at approximately two-week intervals the activity of rMBP-MshC stored as 20% glycerol stocks at −80 °C. Thus far, we have observed no detectable loss of activity upon storage under these conditions for at least three months (data not shown), further indicating that these constructs and protocols will allow for production of active enzyme in quantities sufficient to support future studies.

Table 3
Purification of recombinant GST-MshC and MBP-MshC proteins[a]

Step	Total protein (mg)	Total activity (nmol min^{-1})	Specific activity (nmol min^{-1} mg^{-1})	Recovery (%)
Total protein extract				
GST-MshC	118.1	36.5	0.3	
rGST-MshC	0.7	6.3	9.0	17
MBP-MshC	82.7	26.9	0.3	
rMBP-MshC	1.2	22.3	25.0	83

[a] Specific activity of supernatants is based on total protein determined by the Bradford protein assay.

Concluding remarks

In summary, we have demonstrated that recombinant MshC fusion proteins can be expressed in *M. smegmatis* using the pACE vector under the control of the acetamidase promoter. The GB1, GST and MBP tags are well tolerated in the I64 *M. smegmatis* mutant and the resulting MshC fusion proteins are functional in vivo. Acetamidase-induced expression of each of these constructs greatly improved the yield of MshC relative to the naturally occurring levels produced by the wild type strain *M. smegmatis* mc^{2}155. The yields obtained with this system for the GST-MshC and MBP-MshC protein ranged from 0.7 to 1.2 mg of pure and active protein/L. The specific mycothiol ligase activities measured for rGSTMshC and rMBPMshC were 9 nmol min^{-1} mg^{-1} and 25 nmol min^{-1} mg^{-1}, respectively. In particular, recombinant GST- and MBP-tagged MshC constructs can be produced at levels that will support screening efforts aimed at identifying inhibitors of mycothiol ligase. If desired, MshC may also be cleaved from its N-terminal tag by incubation with proteases Factor Xa (MBP-MshC and GST-MshC) or thrombin (GST-MshC) with negligible losses in quantity or activity (M.T.G.L. and C.A.B., unpublished data). This expression system represents a robust method for expression of recombinant mycobacterial proteins that cannot be expressed in other heterologous systems such as *E. coli*.

Acknowledgments

References

[1] G.L. Newton, K. Arnold, M.S. Price, C. Sherrill, S.B. Delcardayre, Y. Aharonowitz, G. Cohen, J. Davies, R.C. Fahey, C. Davis, Distribution of thiols in microorganisms: mycothiol is a major thiol in most actinomycetes, J. Bacteriol. 178 (1996) 1990–1995.

[2] G.L. Newton, R.C. Fahey, Mycothiol biochemistry, Arch. Microbiol. 178 (2002) 388–394.

[3] N.A. Buchmeier, G.L. Newton, T. Koledin, R.C. Fahey, Association of mycothiol with protection of *Mycobacterium tuberculosis* from toxic oxidants and antibiotics, Mol. Microbiol. 47 (2003) 1723–1732.

[4] M. Rawat, G.L. Newton, M. Ko, G.J. Martinez, R.C. Fahey, Y. Av-Gay, Mycothiol-deficient *Mycobacterium smegmatis* mutants are hypersensitive to alkylating agents, free radicals, and antibiotics, Antimicrob. Agents Chemother. 46 (2002) 3348–3355.

[5] G.M. Nicholas, C.A. Bewley, Inhibitors of mycothiol-*S*-conjugate amidase and related genes, Curr. Med. Chem.–Anti-Infective Agents 3 (2004) 221–231.

[6] D. Sareen, M. Steffek, G.L. Newton, R.C. Fahey, ATP-dependent L-cysteine:1D-myo-inosityl2-amino-2-deoxy-α-D-glucopyranoside ligase, mycothiol biosynthesis enzyme MshC, is related to class I cysteinyl-tRNA synthetases, Biochemistry 41 (2002) 6885–6890.

[7] D. Sareen, G.L. Newton, R.C. Fahey, N.A. Buchmeier, Mycothiol is essential for growth of *Mycobacterium tuberculosis* Erdman, J. Bacteriol. 185 (2003) 6736–6740.

[8] C.M. Sassetti, D.H. Boyd, E.J. Rubin, Genes required for mycobacterial growth defined by high density mutagenesis, Mol. Microbiol. 48 (2003) 77–84.

[9] S. Daugelat, J. Kowall, J. Mattow, D. Bumann, R. Winter, R. Hurwitz, S.H.E. Kaufmann, The RD1 proteins of *Mycobacterium tuberculosis*: expression in *Mycobacterium smegmatis* and biochemical characterization, Microb. Infect. 5 (2003) 1082–1095.

[10] R. Vincentelli, S. Canaan, V. Campanacci, C. Valencia, D. Maurin, F. Frassinetti, L. Scappucini-Calvo, Y. Bourne, C. Cambillau, C. Bignon, High-throughput automated refolding screening of inclusion bodies, Protein Sci. 13 (2004) 2782–2792.

[11] A. Korepanova, F.P. Gao, Y. Hua, H. Qin, R.K. Nakamoto, T.A. Cross, Cloning and expression of multiple integral membrane proteins from *Mycobacterium tuberculosis* in *Escherichia coli*, Protein Sci. 14 (2005) 148–158.

[12] S. Ahmad, M.M. Ali, A.S. Mustafa, Construction of a modified vector for efficient purification of recombinant *Mycobacterium tuberculosis* proteins expressed in *Escherichia coli*, Protein Expres. Purif. 29 (2003) 167–175.

[13] K.A.L. De Smet, K.E. Kempsell, A. Gallagher, K. Duncan, D.B. Young, Alteration of a single amino acid residue reverses fosfomycin resistance of recombinant MurA from *Mycobacterium tuberculosis*, Microbiology 145 (1999) 3177–3184.

[14] T. Parish, E. Mahenthitalingam, P. Draper, E.O. Davis, M.J. Colston, Regulation of the inducible acetamidase gene of *Mycobacterium smegmatis*, Microbiology 143 (1997) 2267–2276.

[15] J.F. Huth, C.A. Bewley, B.M. Jackson, A.G. Hinnebusch, G.M. Clore, A.M. Gronenborn, Design of an expression system for detecting folded protein domains and mapping macromolecular interactions by NMR, Protein Sci. 6 (1997) 2359–2364.

[16] J. Sambrook, E.F. Fritsch, T. Maniatis, Commonly used techniques in molecular cloning, in: second ed., Molecular Cloning A Laboratory Manual, vol. 3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989, pp. E10–E14.

[17] M.L. Beggs, J.T. Crawford, K.D. Eisenach, Isolation and sequencing of the replication region of *Mycobacterium avium* plasmid pLR7, J. Bacteriol. 177 (1995) 4836–4840.

[18] C.G. Stanley, P.H. von Hippel, Calculation of protein extinction coefficients from amino acid sequence data, Anal. Chem. 182 (1989) 319–326.

[19] G.L. Newton, Y. Av-Gay, R.C. Fahey, A novel mycothiol-dependent detoxification pathway in *Mycobacteria* involving mycothiol *S*-conjugate amidase, Biochemistry 39 (2000) 10739–10746.

[20] Y-P. Shih, W.-M. Kung, J.-C. Chen, C.-H. Yeh, A.H.-J. Wang, T.-F. Wang, High throughput screening of soluble proteins, Protein Sci. 11 (2002) 1714–1719.

[21] P. Braun, Y. Hu, B. Shen, A. Halleck, M. Koundinya, E. Harlow, J. LaBaer, Proteome-scale purification of human proteins from bacteria, Proc. Natl. Acad. Sci. USA 99 (2002) 2654–2656.

[22] M. Hammarström, N. Hellgren, S. Van Den Berg, H. Berglund, T. Härd, Rapid screening for improved solubility of small proteins produced as fusion proteins in *Escherichia coli*, Protein Sci. 11 (2002) 313–321.

[23] E. Mahenthiralingam, P. Draper, E.O. Davis, M.J. Colston, Cloning and sequencing of the gene which encodes the highly inducible acetamidase from *Mycobacterium smegmatis*, J. Gen. Microbiol. (1993) 575–583.

[24] G. Roberts, D.G.N. Muttucumaru, T. Parish, Control of the acetamidase gene of *Mycobacterium smegmatis* by multiple regulators, FEMS Microbiol. Lett. 221 (2003) 131–136.