research articles Journal of proteome research

Proteomic Definition of the Cell Wall of *Mycobacterium tuberculosis*

Lisa M. Wolfe, Spencer B. Mahaffey, Nicole A. Kruh, and Karen M. Dobos*

Department of Microbiology, Immunology and Pathology, 1682 Campus Delivery, Colorado State University, Fort Collins, Colorado 80523

Received June 14, 2010

The cell envelope of *Mycobacterium tuberculosis* (*Mtb*) is complex and diverse; composed of proteins intermingled in a matrix of peptidoglycan, mycolic acids, lipids, and carbohydrates. Proteomic studies of the *Mtb* cell wall have been limited; nonetheless, the characterization of resident and secreted proteins associated with the cell wall are critical to understanding bacterial survival and immune modulation in the host. In this study, the cell wall proteome was defined in order to better understand its unique biosynthetic and secretion processes. *Mtb* cell wall was subjected to extraction with organic solvents to remove noncovalently bound lipids and lipoglycans and remaining proteins were solubilized with either SDS, Guanidine-HCl, or TX-114. These extracts were analyzed by two-dimensional gel electrophoresis and mass-spectrometry and resulted in the identification of 234 total proteins. The lipoproteome of *Mtb*, enriched in the TX-114 extract, was further resolved by multidimensional chromatography and mass spectrometry to identify an additional 294 proteins. A query of the 528 total protein identifications against Neural Network or Hidden Markov model algorithms predicted secretion signals in 87 proteins. Classification of these 528 proteins also demonstrated that 35% are involved in small molecule metabolism and 25% are involved in macromolecule synthesis and degradation building upon evidence that the *Mtb* cell wall is actively engaged in mycobacterial survival and remodeling.

Keywords: *Mycobacterium tuberculosis* • proteome • cell wall • mass spectrometry

Introduction

Despite efforts to eradicate tuberculosis, it remains one of the most successful and deadly human diseases. The tubercle bacilli resides within dendritic cells and macrophages, and in an immunocompetent host, the infection is walled off within a granuloma where it can remain dormant for years. While many factors contribute to its success, it is the thick, waxy cell wall of the bacillus that prevents dehydration, affords protection against varying levels of acidity and the detrimental affects of free radicals. In many respects, it is the unique architecture of the cell wall itself that makes *Mycobacterium tuberculosis* (*Mtb*) infection relatively difficult to treat with antibiotics. Since it is also a reservoir for many proteins and nonproteinaceous antigens, which are secreted into the extracellular milieu to stimulate and/or suppress the host immune response,[1,2] its definition can be exploited for vaccine development.

For decades, the macromolecular features of the mycobacterial cell wall, including the mycolic acid and arabinogalactan core, have been studied in detail.[3] Structurally, the cell wall of *Mtb* is composed of a distinct inner core of mycolic acid-arabinogalactan-peptidoglycan (mAGP). In addition to the covalently attached lipids and carbohydrates, it is well-known that the free lipids, lipoglycans, and phosphotidyl inositols that reside in the outer core of the cell wall play key roles in modulation of the host immune response.[4] Specifically, the molecules, lipoarabinomannan (LAM), lipomannan (LM), and phosphatidyl inositol mannoside (PIM) are known to aid the process of host immune evasion.[2] In addition, virulence lipids such as trehalose dimycolate/monomycolate (TDM/TMM), phthiocerol dimycocerosate (PDIM), and sulfolipids (SL) and the protein machinery, such as MmpLs, required for their export are intercalated in the cell wall.[5]

Numerous cell wall associated proteins, including many lipoproteins and lipoglycoproteins, have also been described.[6−8] For example, the TLR2 agonists, lpqH (19 kDa), pstS1 (38 kDa), and lprG (Rv14llc), are all found in the cell wall,[7−11] where they function to regulate the action of macrophages and dendritic cells.[12] PstS1 also plays a role in bacterial escape from the host macrophage through apoptosis.[13] Many other lipoproteins of unknown function are identified in the secreted proteome of *Mtb* culture filtrate.[14−16]

The study of the secreted proteome of *Mtb* was driven by the search for novel immunodominant antigens, drug targets, and biomarkers for disease.[17−20] To build upon this work, a number of studies employed advances in proteomic technologies such as two-dimensional gel electrophoresis (2DGE) and liquid chromatography mass spectrometry (LC−MS/MS) to further mine additional subcellular compartments of *Mtb*—cytosol, membrane, and cell wall.[21−26] Historically, the proteins within the cell wall have been difficult to resolve and identify by traditional 2DGE methods.[19,27] Mawuenyega et al., employing two-dimensional liquid chromatography coupled with mass spectrometry (2DLC/MS) were the most successful at defining the cell wall proteome with the identification of 306

* Corresponding author. E-mail: kdobos@colostate.edu.

10.1021/pr1005873
Published on Web 09/09/2010

proteins.[24] In this study, we set out to comprehensively describe the *Mtb* cell wall proteome in an effort to exploit additional proteins that may play a role in host-pathogen interactions and define new potential drug targets via discovery of unique biosynthetic or metabolic processes. We used a combination of detergent extraction, 2DGE, multidimensional liquid chromatography, and mass spectrometry to achieve our goal.

Methods

Preparation of *Mtb* Cell Wall. *Mtb* strain H37Rv was cultured in 2 L of glycerol alanine salts (GAS) medium[28] in roller bottles for 14 d at 37 °C with gentle agitation. Cells were harvested,[24] washed with phosphate-buffered saline (PBS), pH 7.4, and inactivated by γ-irradiation. Cells were disrupted in a French press, free lipids were removed, and the cell wall was obtained as previously described.[29] Briefly, 1 g of lyophilized cell wall was subjected to two extractions of 2 h each followed by one 18 h extraction with chloroform/methanol (2:1, v/v) at a ratio of 30 mL/g of cell wall. Extractions were performed at 22 °C with agitation. Centrifugation at 27 000*g* for 30 min was performed to collect cell wall material. The 2:1 extracted cell wall was dried under N_2 and further extracted twice for 2 h and one 18 h extraction with chloroform/methanol/water (10:10:3, v/v/v) each at 22 °C. The fully delipidated cell wall was dried under N_2 and resuspended in PBS, pH 7.4. Cell wall protein was quantified by bicinchoninic acid (BCA) assay (Thermo Pierce).

Detergent Extractions of Cell Wall Proteins. Cell wall protein (CWP) was solubilized by one of three methods using (i) 6 M guanidine HCl (GuHCl), (ii) 2% sodium dodecyl sulfate (SDS), or (iii) 4% TritionX-114 (TX-114). For method (i), 75 mg of CWP was incubated at 22 °C for 4 h with agitation. The sample was centrifuged at 27 000*g* for 30 min. Solubilized proteins were exchanged into 0.01 M NH_4HCO_3, and the protein amount was determined by BCA assay. The average protein recovery following GuHCl extraction of CWP was 40%. For method (ii), SDS soluble proteins were generated by extraction of 100 mg of CWP with 2% SDS in PBS (w/v) as described previously.[29] Briefly, the sample was bath sonicated for 3 h at 90 °C. The sample was centrifuged at 27 000*g* at 22 °C for 30 min, collected, and centrifuged again. The fully cleared supernatant was then subjected to paired-ion extraction for removal of SDS.[30] Proteins were concentrated by centrifugation as above and pellets washed with acetone at −20 °C for 4 h. The final pellet was resuspended in 0.01 M NH_4HCO_3 and protein amount was determined by BCA. The average protein recovery following SDS extraction of CWP was 60%. In method (iii), a stock solution of 32% TX-114 in PBS was added to 300 mg of CWP to a final concentration of 4% detergent. Primary extraction occurred at 4 °C for 16 h. The extract was allowed to biphase at 37 °C for 30 min and was fully separated by centrifugation at 27 000*g* for 30 min. Each phase was back extracted two additional times for 2 h each. The TX-114 detergent phases were pooled, and proteins collected by cold acetone precipitation.[31] Final TX-114 proteins were resuspended as above and protein amount was determined by BCA assay. Protein recovery following TX-114 extraction of CWP was 5%.

Two-Dimensional Gel Electrophoresis (2DGE). Samples of 200 μg of each CWP preparation were solubilized for 9−14 h in rehydration buffer (7 M urea/2 M thiourea, 1% amidosulfobetaine-14 (ASB-14), 1% 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPS), 1 mM dithiothreitol (DTT), 0.25% NP-40, 0.625% ZOOM carrier ampholytes 3−10, and 1.9% ZOOM carrier ampholytes 4−7 (Invitrogen). Solubilized CWP was applied on ZOOM precast immobilized pH 4−7 linear gradient strips (7.0 cm; Invitrogen) according to manufacturer's instructions. Focusing was achieved using a stepwise voltage gradient of 200, 450, 700, and 1000 V, for 10 min each followed by focusing at 2000 V for 2 h. SDS-PAGE of the isoelectric focusing (IEF) strips was performed using ZOOM 4−12% Bis-Tris SDS-PAGE gradient gels (Invitrogen). The gels were stained by Coomassie blue R-250 (Bio-Rad). Images were captured using a Gel-Doc System (Bio-Rad, Hercules, CA), and spot detection was performed using Delta 2D software (Greifswald, Germany).

Two-Dimensional Liquid Chromatography. A total of 5.0 mg of TX-114 CWP was digested with modified trypsin (Roche Diagnostics) at a ratio of 1:20 (E/S), in 0.1 M Urea and 20 mM methylamine. The digest was desalted using Sep-Pak Light C18 cartridge (Waters, Inc.) and concentrated under vacuum. The resultant peptides were separated by strong cation-exchange (SCX) chromatography using a polysulfylethyl A column (460 μm × 200 mm, 300 Å; Poly LC, Inc.) connected to a Waters Alliance analytical HPLC with a 2487 UV detector. The digest was applied in buffer A (5 mM K_2PO_4, 20% acetonitrile (ACN), pH 3.0) using a gradient of 0−80% buffer B (A with 0.5 M KCl) over 75 min with a flow rate of 1 mL/min. The elution of peptides was monitored at 214 nm, and fractions (3 mL each) were pooled based on UV absorbance. Each fraction (6 total) was concentrated under vacuum and resuspended into reverse phase buffer A (0.1% TFA in H_2O). Individual SCX pools were subjected to further separation by reversed-phase HPLC (RP-HPLC) using a monomeric C_{18} column (4.6 mm × 150 mm, Vydac). Peptides were eluted using a gradient of 0−50% reverse-phase buffer B (90% ACN in A) over 40 min at a flow rate of 1 mL/min. Peptides were manually collected based on UV absorbance at 214 nm. A total of 139 RP-HPLC peptide fractions were collected and concentrated under vacuum.

Mass Spectrometry. Coomassie blue stained spots were excised from 2-DE gels and subjected to in-gel digestion[32,33] with modified trypsin (Roche Diagnostics). Digests and 2DLC peptide fractions were resolved by liquid chromatography−mass spectrometry (LC−MS) using either an LCQ (2DGE) or LTQ (2DLC) as described previously.[34,50] Tandem mass spectra were extracted, charge state deconvoluted, and deisotoped by BioWorks version 3.2 (Thermo Finnigan, San Jose, CA). All MS/MS samples were analyzed using Sequest (Thermo Finnigan; version 27, rev. 12) Sequest was set up to search the TB genome database (version 2.0, GenBank accession no. AL123456, 3912 entries) assuming the digestion enzyme trypsin, a fragment ion mass tolerance of 1.0 Da, a parent ion tolerance of 2.5 Da, and 3 allowable missed cleavages. Oxidation of methionine and acrylamide adduct of cysteine (2DGE spots) were specified in Sequest as variable modifications.

Criteria for Peptide Identification. Peptide identifications were accepted if they exceeded specific database search engine thresholds. Sequest identifications required at least deltaCn scores of greater than 0.3 and XCorr scores of greater than 1.5, 2.2, 2.5, and 2.5 for singly, doubly, triply, and quadruply charged peptides (Supplementary Table S1).

Criteria for Protein Identification. Scaffold (version Scaffold-01_05_21, Proteome Software, Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 90% probability as specified by the

Peptide Prophet algorithm.[35] Protein identifications were accepted if they could be established at greater than 90% probability as assigned by the Protein Prophet algorithm[35] and contained at least two unique peptides. Protein identifications at the lower threshold of acceptance were manually inspected for spectra quality. The protein probability false discovery rate for the CWP identifications was 3.5%.[36] Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony (Supplementary Table S1).

SignalP and LipoP Interrogation. FASTA files for each annotated *Mtb* protein were collected and organized into small groups of 50. Each group was then submitted to the SignalP server (http://www.cbs.dtu.dk/services/SignalP/)[37−39] using the gram-positive option. The output was saved in HTML format and a Perl script written to combine the results from all HTML files into one file and then converted to tab delimited format. This generated a complete list of results that were sorted in a spreadsheet to create the subsets of Neural Network positive, Hidden Markov Model positive, and the union of the NN and HMM sets. These lists were input into a java application along with the list of proteins isolated from the cell wall fractions. This application simply compared the predicted list to those isolated and identified the proteins found in the cell wall that were also predicted in each subset as having a signal peptide. To predict lipoproteins, the *Mtb* peptide fasta file from above was broken down into 1000 sequences per file. These were run through the downloadable version of LipoP (http://www.cbs-.dtu.dk/services/LipoP/). The output files were combined in a spreadsheet, and from the sorted data, a list was extracted with every signal peptidase II cleavage site generating a predicted lipoprotein list. This list was input into a Java application that would compare the cell wall proteins, identifying the predicted lipoproteins from each cell wall fraction. The positive results verified that both Signal P models had predicted a signal peptide.

Western Blot Analysis. Protein preparations from the three CWP fractions resolved by 2DGE or SDS-PAGE were transferred to PVDF membranes and incubated with mouse monoclonal antibodies (α-GlcB [Rv1837c], α-PstS1 [Rv0934; IT-23], α-LprG [Rv1411c], α-Ald [Rv2780], α-LpqH [Rv3763; IT-19]; all available through the NIH, NIAID Contract, Tuberculosis Vaccine Testing and Research Materials) and polyclonal antibody against fructose bisphosphate aldolase (α-Fba [Rv0363c], kindly provided by Dr. Mary Jackson, Colorado State University) in TBS and 0.25% Tween 80 (TBST). After incubation, membranes were rinsed with TBST and developed using Amersham ECL Advance Western Blotting Detection Kit (GE Healthcare Life Science, Piscataway, NJ) as per manufacturers' instructions. Protein immunoblots were visualized with the Typhoon 9400 multiplex fluorescent imager (GE Healthcare Life Science, Piscataway, NJ).

Results

To maximize the resolution and protein identification of cell wall proteins, the delipidated cell wall of *Mtb* was subjected to extraction based on solubility within three reagents—an anionic (SDS), a chaotrophic (GuHCl), and a nonionic (TX-114) detergent. 2DGE of all CWP fractions demonstrated the majority of proteins resolving within a range of MW 25−75 kDa (Supplementary Figure 1). On average, 210 spots were resolved for GuHCl and SDS CWP fractions and 170 from TX-114. From the 2DGE analysis, 290 proteins were identified with 122, 131, and 37 proteins identified in the GuHCl, SDS, and TX-114 fractions,

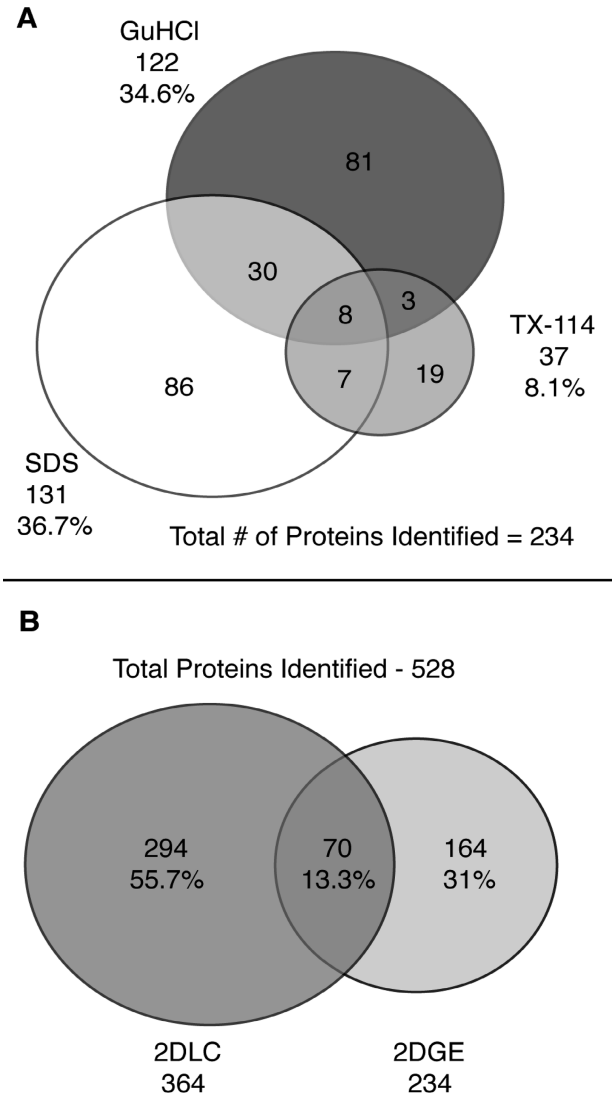

Figure 1. (A) Identification of cell wall proteins by two-dimensional gel electrophoresis and two-dimensional liquid chromatography. Detergent extraction and separation of cell wall protein subsets via 2DGE demonstrated the majority of proteins to be resolved using SDS and GuHCl. The lipoprotein-rich TX-114 fraction was less amenable to gel separation. See text for details. (B) Cumulatively, separation of cell wall proteins by 2DGE resulted in the identification of 234 proteins. Separating the TX-114 protein subset by multidimensional chromatography (2DLC−MS/MS) resolved an additional 294 proteins.

respectively. For all fractions, over half were unique to each subset (Figure 1A). Analysis of the data demonstrated 80% of proteins had known functions (Figure 2). In contrast, the TX-114 extracted fraction was less amenable to gel separation most likely due to its increased hydrophobic content. A multidimensional chromatography approach[40,41] was employed to better resolve proteins in this fraction, in order to exploit any unique or uncharacterized protein families within the TX-114 subset. While 2DGE was insufficient in identifying a significant number of proteins in the TX-114 detergent extract (Figure 1A), SCX chromatography combined with reverse phase chromatography of a digest of this CWP preparation allowed the resolution of 364 proteins, including an additional 294 proteins not found in any of the 2DGE preparations (Figure 1B).

Validation of a few selected proteins, by Western blot analysis of samples resolved by 2DGE (Supplementary Figure 1) and

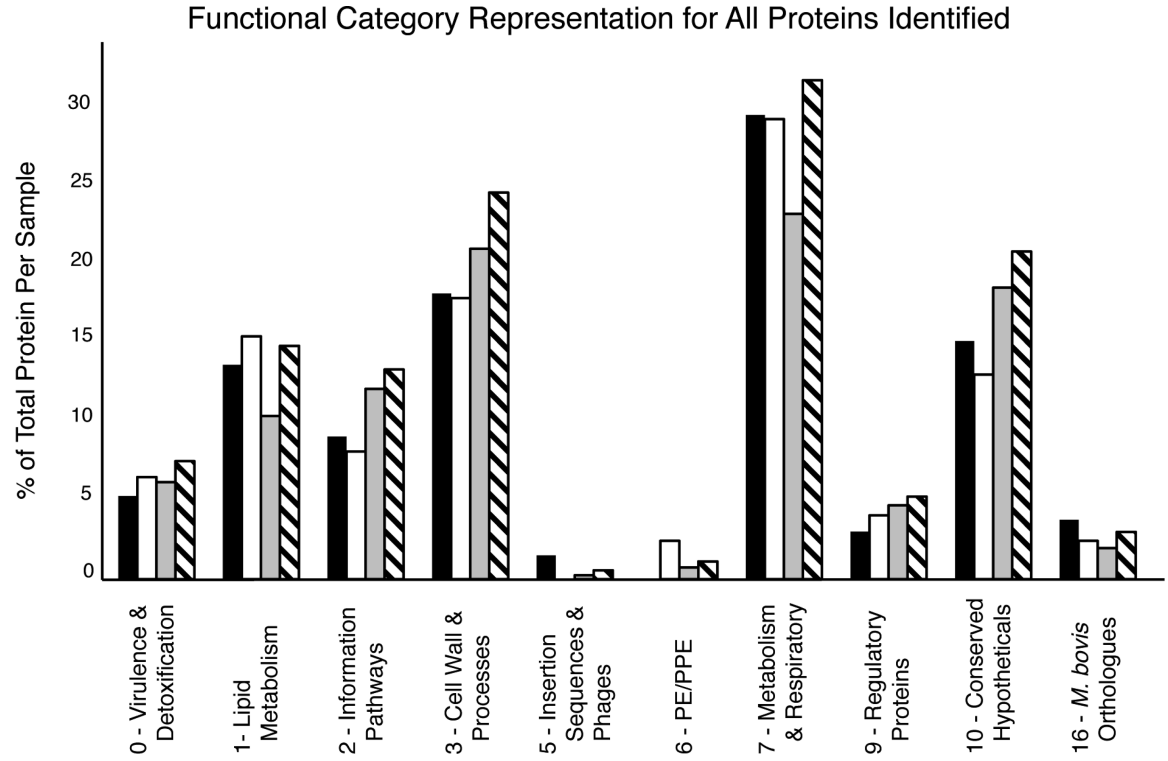

Figure 2. The functional category distribution by biological sample. Proteins identified for each extraction method (GuHCl, SDS, TX114) were sorted by functional category as a percentage of total proteins per sample. Representation of all protein groups was identified with the exception of groups 4 and 8 (data not shown). Seventy percent of the proteins identified in the cell wall are associated with lipid metabolism (group 1), cell wall processes (group 3), and intermediary metabolism (group 7). In addition, 21% are conserved hypotheticals. Black, SDS; white, GuHCl; gray, TX114; hash, total.

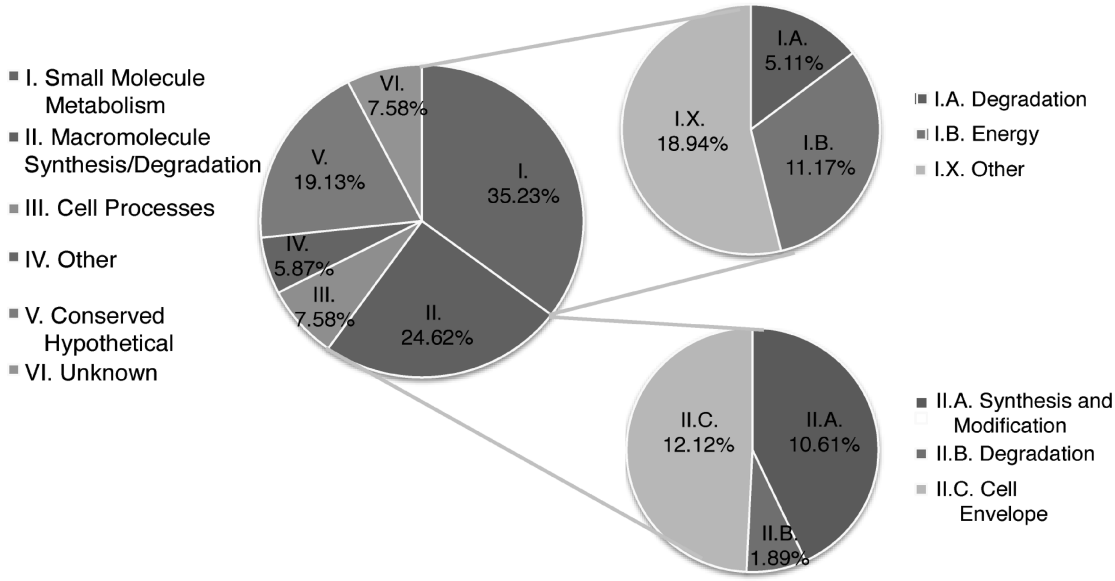

Figure 3. The functional category distribution of the 528 identified proteins. Assignments were made based on the Sanger Institute gene database. The distributions are among the major functional groups and the subgroups within functional groups I and II. The percentage for each subgroup indicates the percentage of the total number of identified proteins in its major functional group. I.X is defined by proteins involved in central intermediary metabolism, amino acid biosynthesis, polyamine synthesis, biosynthesis of cofactors, prosthetic groups, and carriers, lipid biosynthesis, polyketide and non-ribosomal peptide synthesis, and broad regulatory functions.

SDS-PAGE (Supplementary Figure 2), corroborated the identification of these proteins by mass spectrometry.

All identified proteins were grouped by functional category as defined by Institute Pasteur, which demonstrated *Mtb* protein families present in each extract. Proteins in categories 3 (cell wall and cell wall processes) and 7 (intermediary metabolism) were consistently overrepresented among the CWP preparations (Figure 2). Combining the data sets of both gel and non-gel based protein identifications led to the detection of 528 proteins (Supplementary Table S2). These were further classified into functional groups as defined by the Sanger Institute (Figure 3A–C). One hundred and five of the

Table 1. Cell Wall Protein Identifications Unique to This Study

Rv no.	protein name	protein function	class ID TUBERCULIST	class ID SANGER	CWP fraction[a]
Rv0007	Rv0007	POSSIBLE CONSERVED MEMBRANE PROTEIN	3	V	D
Rv0014c	pknB	TRANSMEMBRANE SERINE/THREONINE-PROTEIN KINASE B PKNB (PROTEIN KINASE B)	9	I.J.3	D
Rv0018c	pstP	POSSIBLE SERINE/THREONINE PHOSPHATASE PSTP	9	I.J.3	D
Rv0092	ctpA	PROBABLE CATION TRANSPORTER P-TYPE ATPASE A CTPA	3	III.A.2	B
Rv0142	Rv0142	CONSERVED HYPOTHETICAL PROTEIN	10	VI	A, B
Rv0327c	cyp135A1	POSSIBLE CYTOCHROME P450 135A1 CYP135A1	7	IV.F	B
Rv0361	Rv0361	PROBABLE CONSERVED MEMBRANE PROTEIN	3	II.C.5	D
Rv0431	Rv0431	PUTATIVE TUBERCULIN RELATED PEPTIDE	3	II.C.2	D
Rv0497	Rv0497	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	3	II.C.5	D
Rv0512	hemB	PROBABLE DELTA-AMINOLEVULINIC ACID DEHYDRATASE HEMB	7	I.G.12	C
Rv0614	Rv0614	CONSERVED HYPOTHETICAL PROTEIN	10	V	A
Rv0622	Rv0622	POSSIBLE MEMBRANE PROTEIN	3	II.C.5	B
Rv0638	secE1	PROBABLE PREPROTEIN TRANSLOCASE SECE1	3	III.D	D
Rv0712	Rv0712	CONSERVED HYPOTHETICAL PROTEIN	10	V	B
Rv0771	Rv0771	POSSIBLE 4-CARBOXYMUCONOLACTONE DECARBOXYLASE (CMD)	7	II.B.6	D
Rv0858c	Rv0858c	PROBABLE AMINOTRANSFERASE	7	IV.H	A
Rv0892	Rv0892	PROBABLE MONOOXYGENASE	7	I.B.7	B
Rv0899	ompA	OUTER MEMBRANE PROTEIN A OMPA	3	II.C.2	D
Rv0902c	prrB	TWO COMPONENT SENSOR HISTIDINE KINASE PRRB	9	I.J.2	D
Rv0926c	Rv0926c	CONSERVED HYPOTHETICAL PROTEIN	10	V	A
Rv0954	Rv0954	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	3	II.C.2	D
Rv1016c	lpqT	PROBABLE CONSERVED LIPOPROTEIN LPQT	3	II.C.1	D
Rv1029	kdpA	PROBABLE POTASSIUM-TRANSPORTING ATPASE A CHAIN KDPA	3	III.A.2	A
Rv1096	Rv1096	POSSIBLE GLYCOSYL HYDROLASE	7	I.A.1	D
Rv1100	Rv1100	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv1128c	Rv1128c	CONSERVED HYPOTHETICAL PROTEIN	5	IV.B.2	B
Rv1151c	Rv1151c	PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN	9	I.J.1	A
Rv1184c	Rv1184c	POSSIBLE EXPORTED PROTEIN	3	V	D
Rv1196	PPE18	PPE FAMILY PROTEIN	6	IV.C.2	D
Rv1209	Rv1209	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv1230c	Rv1230c	POSSIBLE MEMBRANE PROTEIN	3	II.C.5	B
Rv1256c	cyp130	PROBA BLE CYTOCHROME P450 130 CYP130	7	IV.F	B
Rv1266c	pknH	PROBABLE TRANSMEMBRANE SERINE/THREONINE-PROTEIN KINASE H PKNH (STPK H)	9	I.J.3	B,D
Rv1273c	Rv1273c	PROBABLE DRUGS-TRANSPORT TRANSMEMBRANE ATP-BINDING PROTEIN ABC TRANSPORTER	3	II.C.5	B
Rv1330c	Rv1330c	CONSERVED HYPOTHETICAL PROTEIN	10	V	B
Rv1361c	PPE19	PPE FAMILY PROTEIN	6	IV.C.2	A
Rv1379	pyrR	PROBABLE PYRIMIDINE OPERON REGULATORY PROTEIN PYRR	9	I.J.1	D
Rv1393c	Rv1393c	PROBABLE MONOOXYGENASE	7	I.B.7	C
Rv1442	bisC	PROBABLE BIOTIN SULFOLIDE REDUCTASE BISC (BDS reductase) (BSO reductase)	7	I.G.1	A,D
Rv1450c	PE_PGRS27	PE-PGRS FAMILY PROTEIN	6	IV.C.1.b	A
Rv1451	ctaB	PROBABLE CYTOCHROME C OXIDASE ASSEMBLY FACTOR CTAB	7	I.B.6.a	B
Rv1466	Rv1466	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv1586c	Rv1586c	PROBABLE PHIRV1 INTEGRASE	5	IV.B.3	D
Rv1594	nadA	PROBABLE QUINOLINATE SYNTHETASE NADA	7	I.G.7	B
Rv1599	hisD	PROBABLE HISTIDINOL DEHYDROGENASE HISD (HDH)	7	I.D.5	D
Rv1602	hisH	PROBABLE AMIDOTRANSFERASE HISH	7	I.D.5	B
Rv1623c	cydA	PROBABLE INTEGRAL MEMBRANE CYTOCHROME D UBIQUINOL 0XIDASE (SUBUNIT I) CYDA	7	I.B.6.c	D
Rv1659	argH	PROBABLE ARGININOSUCCINATE LYASE ARGH	7	I.D.1	D
Rv1666c	cyp139	PROBABLE CYTOCHROME P450 139 CYP139	7	IV.F	A
Rv1748	Rv1748	HYPOTHETICAL PROTEIN	16	VI	D
Rv1777	cyp144	PROBABLE CYTOCHROME P450 144 CYP144	7	IV.F	A
Rv1801	PPE29	PPE FAMILY PROTEIN	6	IV.C.2	A
Rv1836c	Rv1836c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv1957	Rv1957	HYPOTHETICAL PROTEIN	16	VI	D
Rv1969	mce3D	MCE-FAMILY PROTEIN MCE3D	0	IV.A	A
Rv2047c	Rv2047c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv2063c	Rv2063c	HYPOTHETICAL PROTEIN	0	VI	B
Rv2095c	Rv2095c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv2122c	hisE	PROBABLE PHOSPHORIBOSYL-AMP PYROPHOSPHATASE HISE	7	I.D.5	D
Rv2171	lppM	PROBABLE CONSERVED LIPOPROTEIN LPPM	3	II.C.1	D
Rv2198c	mmpS3	PROBABLE CONSERVED MEMBRANE PROTEIN MMPS3	3	II.C.4	D
Rv2205c	Rv2205c	CONSERVED HYPOTHETICAL PROTEIN	10	V	B
Rv2211c	gcvT	PROBABLE AMINOMETHYLTRANSFERASE GCVT (GLYCINE CLEAVAGE SYSTEM T PROTEIN)	7	I.C.1	D
Rv2223c	Rv2223c	PROBABLE EXPORTED PROTEASE	3	II.C.2	B
Rv2247	accD6	ACETYL/PROPIONYL-CoA CARBOXLASE (BETA SUBUNIT) ACCD6	1	I.H.1	A,D

Table 1. Continued

Rv no.	protein name	protein function	class ID TUBERCULIST	class ID SANGER	CWP fraction[a]
Rv2260	Rv2260	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv2307c	Rv2307c	CONSERVED HYPOTHETICAL PROTEIN	10	V	B
Rv2466c	Rv2466c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv2518c	lppS	PROBABLE CONSERVED LIPOPROTEIN LPPS	3	II.C.1	D
Rv2556c	Rv2556c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv2712c	Rv2712c	HYPOTHETICAL PROTEIN	16	VI	A
Rv2720	leXA	REPRESSOR LEXA	9	I.J.1	D
Rv2789c	fadE21	PROBABLE ACYL-CoA DEHYDROGENASE FADE21	1	I.A.3	A
Rv2864c	Rv2864c	POSSIBLE PENICILLIN-BINDING LIPOPROTEIN	3	II.C.3	B
Rv2880c	Rv2880c	CONSERVED HYPOTHETICAL PROTEIN	10	V	A
Rv2881c	cdsA	PROBABLE INTEGRAL MEMBRANE PHOSPHATIDATE CYTIDYLYLTRANSFERASE	1	I.H.3	C
Rv2931	ppsA	PHENOLPTHIOCEROL SYNTHESIS TYPE-I POLYKETIDE SYNTHASE PPSA	1	I.I	A
Rv2936	drrA	PROBABLE DAUNORUBICIN-DIM-TRANSPORT ATP-BINDING PROTEIN ABC TRANSPORTER DRRA	3	III.A.6	D
Rv2938	drrC	PROBABLE DAUNORUBICIN-DIM-TRANSPORT ABC TRANSPORTER DRRC	3	III.A.6	A,B
Rv3012c	gatC	PROBABLE GLUTAMYL-TRNA(GLN) AMIDOTRANSFERASE (GLU-ADT SUBUNIT C)	2	II.A.3	D
Rv3083	Rv3083	PROBABLE MONOOXYGENASE (HYDROYLASE)	7	I.B.7	C
Rv3086	adhD	PROBABLE ZINC-TYPE ALCOHOL DEHYDROGENASE ADHD (ALDEHYDE REDUCTASE)	7	I.B.7	B
Rv3107c	agpS	POSSIBLE ALKYLDIHYDROXYACETONEPHOSPHATE SYNTHASE AGPS (ALKYL-DHAP SYNTHASE)	1	I.B.7	B
Rv3136	PPE51	PPE FAMILY PROTEIN	6	IV.C.2	D
Rv3206c	moeB1	PROBABLE MOLYBDENUM COFACTOR BIOSYNTHESIS PROTEIN(MPT-SYNTHASE SULFURYLASE)	7	I.G.4	D
Rv3211	rhlE	PROBABLE ATP-DEPENDENT RNA HELICASE RHLE	2	II.A.7	D
Rv3243c	Rv3243c	HYPOTHETICAL PROTEIN	16	VI	A
Rv3281	Rv3281	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv3282	Rv3282	CONSERVED HYPOTHETICAL PROTEIN	10	V	A
Rv3373	echA18	PROBABLE ENOYL-CoA HYDRATASE(ENOYL HYDRASE) (UNSATURATED ACYL-CoA HYDRATASE)	1	I.A.3	A
Rv3416	whiB3	TRANSCRIPTIONAL REGULATORY PROTEIN WHIB-LIKE WHIB3	9	I.J.1	A
Rv3421c	Rv3421c	CONSERVED HYPOTHETICAL PROTEIN	10	V	C
Rv3448	Rv3448	PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN	3	II.C.5	D
Rv3496c	mce4D	MCE-FAMILY PROTEIN MCE4D	0	IV.A	C
Rv3515c	fadD19	PROBABLE FATTY-ACID-CoA LIGASE FADD19 (FATTY-ACID-CoA SYNTHETASE)	1	I.A.3	B
Rv3534c	Rv3534c	PROBABLE 4-HYDROXY-2-OXOVALERATE ALDOLASE (HOA)	7	II.B.6	C
Rv3565	aspB	POSSIBLE ASPARTATE AMINOTRANSFERASE ASPB (TRANSAMINASE A) (ASPAT)	7	I.D.2	B
Rv3615c	Rv3615c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv3651	Rv3651	CONSERVED HYPOTHETICAL PROTEIN	10	VI	D
Rv3666c	dppA	PROBABLE PERIPLASMIC DIPEPTIDE-BINDING LIPOPROTEIN DPPA	3	III.A.1	C
Rv3688c	Rv3688c	CONSERVED HYPOTHETICAL PROTEIN	10	V	D
Rv3690	Rv3690	PROBABLE CONSERVED MEMBRANE PROTEIN	3	VI	D
Rv3852	Hns	POSSIBLE HISTONE-LIKE PROTEIN HNS	2	II.A.4	D
Rv3877	Rv3877	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	3	V	B
Rv3903c	Rv3903c	HYPOTHETICAL ALANINE AND PROLINE RICH PROTEIN	16	VI	B

[a] Cell Wall Protein fractions: A, 2% SDS; B, 6M GuHCl; C, 4% TX-114; D, 2DLC TX-114.

528 cell wall proteins identified were not reported in previous *Mtb* proteome publications,[16,19,22,24,27] including a comprehensive proteomic database http://web.mpiib-berlin.mpg.de/cgi-bin/pdbs/2d-page/extern/index.cgi,[19] and http://web.mpiib-berlin.mpg.de/cgi-bin/pdbs/2d-page/extern/index.cgi[42] (Table 1).

A majority of proteins were classified in either Category I, Small Molecule Metabolism (35%), or Category II, Macromolecule synthesis and degradation (25%) (Figure 3). Subclasses of category I showed 19% classified in Small molecule metabolism-other (I.X), which included the classes: Central intermediary metabolism and Amino acid biosynthesis (Figure 3; upper right); and subclasses of category II demonstrated an even distribution of proteins between synthesis/degradation of macromolecules (II.A, II.B, 12.5%) and cell envelope proteins (II.C, 12.1%) (Figure 3; lower right).

The secreted proteins of *Mtb* have traditionally been characterized as important antigens and immune-modulators. Prior to being exported, many of these proteins are resident within the cell wall where their function remains largely unknown. To find the putative secreted proteins identified within the cell wall proteome, all identified proteins were subject to interrogation against Neural Network (NN) and Hidden Markov model (HMM) algorithms (SignalP, http://www.cbs.dtu.dk/services/SignalP/) which revealed 18%, 19%, 27% putative secreted proteins in GuHCl, SDS, TX-114 CWP samples, respectively, and 13% in the 2DLC resolved TX-114 CWP fraction. Cumulatively, of the 528 proteins identified in this study, 87

Table 2. Secreted Cell Wall Proteins and Putative Lipoproteins[a]

		CWP fractions				secreted			
Rv no.	protein name	GuHCl	SDS	TX-114	2DLC	experimental[b]	predicted[c]	NetOglyc motif	reference
Rv0125	pepA	x				yes	yes		42, 16
Rv0129c	fbpC	x	x		x	yes	yes		42, 19, 16
Rv0169	mce1A				x				24
Rv0172	mce1D				x		yes		16
Rv0173	**lprK**				x		yes		16
Rv0174	mce1F				x				16, 19
Rv0194	Rv0194		x						24
Rv0202c	mmpL11	x	x	x					24
Rv0227c	Rv0227c				x				24, 46
Rv0237	**lpqI**				x		yes	yes	16, 46
Rv0291	mycP3				x		yes		16
Rv0292	Rv0292				x				46
Rv0309	Rv0309				x		yes		16
Rv0338c	Rv0338c	x							24, 46
Rv0402c	mmpL1	x					yes		16
Rv0431	Rv0431				x				
Rv0432	**sodC**			x	x			yes	45, 46, 19
Rv0480c	Rv0480c	x							24
Rv0507	mmpL2		x						24
Rv0512	hemB			x					
Rv0559c	Rv0559c				x		yes		42, 16
Rv0583c	**lpqN**				x		yes	yes	16, 19
Rv0622	Rv0622		x						
Rv0830	Rv0830	x							46
Rv0858c	Rv0858c	x							
Rv0892	Rv0892		x						
Rv0902c	prrB				x				
Rv0928	**pstS3**				x		yes	yes	16, 45
Rv0934	**pstS1**		x	x	x	yes	yes	yes	42, 19, 16, 46
Rv1006	**Rv1006**		x		x				24, 46
Rv1016c	**lpqT**				x			yes	46
Rv1029	kdpA	x							
Rv1075c	Rv1075c		x				yes		16, 19
Rv1096	Rv1096				x				
Rv1130	Rv1130	x							24
Rv1184c	Rv1184c				x				
Rv1209	Rv1209				x				
Rv1230c	Rv1230c		x						
Rv1244	**lpqZ**				x				46, 19
Rv1269c	Rv1269c				x		yes		42, 16
Rv1270c	**lprA**			x	x		yes	yes	16, 46, 19
Rv1273c	Rv1273c		x						
Rv1275	**lprC**				x				24, 46
Rv1307	atpH				x				24, 46
Rv1368	**lprF**			x	x			yes	46, 24, 22
Rv1411c	**lprG**		x		x			yes	46, 24, 22
Rv1450c	PE_PGRS27	x							
Rv1488	Rv1488				x		yes		16, 24, 22
Rv1522c	mmpL12		x						22
Rv1815	Rv1815		x				yes		16
Rv1886c	fbpB	x	x			yes	yes		42, 16, 22
Rv1911c	**lppC**	x					yes	yes	22, 46
Rv1926c	mpt63				x	yes	yes		42, 16
Rv1969	mce3D	x							
Rv1980c	mpt64				x	yes	yes		42, 16
Rv1984c	cfp21	x					yes		16
Rv2095c	Rv2095c				x				
Rv2171	**lppM**				x				
Rv2216	Rv2216				x				22
Rv2223c	Rv2223c		x						
Rv2307c	Rv2307c		x						
Rv2345	Rv2345				x				46, 22
Rv2376c	cfp2	x				yes	yes		19, 16
Rv2518c	**lppS**				x				
Rv2531c	Rv2531c		x						24
Rv2536	Rv2536				x				24
Rv2576c	Rv2576c	x							16
Rv2625c	Rv2625c		x						22
Rv2631	Rv2631		x						16

Table 2. Continued

		CWP fractions				secreted			
Rv no.	protein name	GuHCl	SDS	TX-114	2DLC	experimental[b]	predicted[c]	NetOglyc motif	reference
Rv2712c	Rv2712c	x							
Rv2721c	Rv2721c	x			x				16
Rv2756c	hsdM		x						24
Rv2864c	**Rv2864c**		x						
Rv2945c	**lppX**			x	x			yes	16, 46, 24, 22
Rv3006	**lppZ**				x			yes	16, 46, 22
Rv3106	fprA				x				22
Rv3193c	Rv3193c				x				16, 22
Rv3224	Rv3224	x			x	yes			42, 19, 24, 22
Rv3243c	Rv3243c	x							
Rv3244c	**lpqB**		x					yes	16, 46
Rv3584	**lpqE**				x				16, 22
Rv3623	**lpqG**			x				yes	24, 22
Rv3666c	**dppA**			x					
Rv3682	ponA2	x							16, 24, 22
Rv3725	Rv3725		x						16
Rv3763	**lpqH**			x	x	yes		yes	19, 46, 24, 22
Rv3804c	fbpA		x		x	yes	yes		42, 19, 16, 24, 22

[a] Proteins in bold indicate putative lipoprotein. [b] Experimental evidence reported for secretion. See reference. [c] Predicted Signal Petptidase I cleavage (SignalP, http://www.cbs.dtu.dk/services/SignalP/).

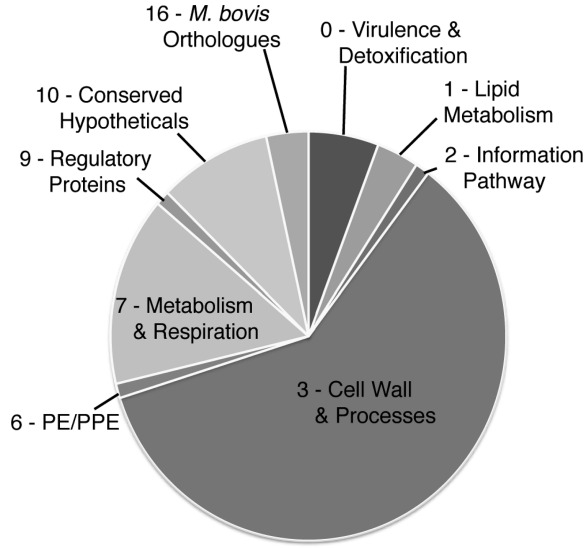

Figure 4. Functional classification for putative secreted proteins and lipoproteins. Eighty-seven cell wall proteins were predicted to contain secretion signals and putative lipoprotein motifs. Sixty percent of these proteins are classified within the cell wall and cell wall processes functional category.

proteins were predicted to contain secretion signals and included the identification of 23 proteins uniquely found in this study (Table 2). A majority (60%) of the CWP associated with secretion are indeed within the category of the cell wall and its processes (Figure 4). Further, many of these CWP are also membrane associated, either by description[22,24,25] or functional annotation. Of the 23 secreted CWP unique to this study, 11 have been described to be involved with small molecule and peptide binding. Additionally, 23 of the 87 secreted proteins are classified as hypothetical or unknown (Categories V and VI of Sanger Institute) illustrating that to a large extent, the functions of these exported proteins are poorly understood.

Next, we interrogated the CWP to identify those with putative lipoprotein motifs. Here, a total of 16 proteins were predicted by LipoP (LipoP, http://www.cbs.dtu.dk/services/LipoP/)[43] to contain a signal peptidase II cleavage motif, and an additional 8 proteins were identified when compared to the 99 putative lipoproteins (2.5% of the genome) that reportedly exist within the *Mtb* proteome.[7] A majority of the lipoproteins contain no additional functional classification; however, 4 proteins (Rv0928, Rv0934, Rv2864c, Rv3666c) are involved with substrate binding and transport within the periplasm. One of two superoxide dismutases in *Mtb*, SodC (Rv0432), was also identified within the lipoprotein-enriched TX-114 CWP fraction. This protein has recently been defined as a highly glycosylated putative lipoprotein and is a defined B-cell antigen.[44,45] Other putative lipoproteins identified in this study contain O-glycosylation motifs as predicted by NetOglyc and were found within the *Mtb* glycoproteome (Table 2).[46] The function of these dual-modified lipo-glycoproteins, as well as the nature of their modification remain largely undefined.

Discussion

In this study, we focused on the gel-based two-dimensional separation of the cell wall and complemented that separation by liquid chromatography of digested TX-114 CWP. Identified protein families included known antigens, Fbp A, B, and C (Rv 3804c, Rv1886c, and Rv0129c), CFP10 (Rv3874) and Esat-6 like proteins EsxJ and EsxL (Rv1038c, Rv1198) along with antigens lpqH (Rv3763), lprG (Rv1411c), and lprA (Rv1270c). Others included members of the MmpL and Mce protein family of lipid export machinery. In concordance with previous work,[26,47] a number of ribosomal proteins are reported here, which must indicate a high level of protein synthesis occurring at the cytosol−cell wall interface, most likely to facilitate the entry of proteins into the cellular envelope.

For over a decade, the advancement of techniques in 2DGE, mass spectrometry, and increased access to bioinformatic tools greatly enhanced proteomic studies of *Mtb*. Largely descriptive, these studies were undertaken to identify novel virulence factors and drug targets.[18,19,27] Subsequent studies set out to understand functional relationships between proteins and discover antigens responsible for adaptive T-cell responses.[23,24,26,48,49] The quantitative techniques

of mass spectrometric profiling, isotope coded affinity tag (iCAT) and isobaric tagging for relative and absolute quantification (iTRAQ), have afforded the accurate monitoring of up- and down-regulation of proteins on a global scale.[23,50] Previous work mining subcellular fractions of *Mtb*, focused on the cytosol and culture filtrate fractions, resulted in thorough characterization and the creation of 2D gel databases.[19,42,51−53] Efforts have been made more recently to resolve the insoluble cell wall fraction. As mentioned previously, Mawuenyega et al. was able to identify 300 proteins within the cell wall using 2DLC as a separation method. Their study demonstrated functional relationships among key protein families within the fatty-acid synthesis pathway. Using these studies as experimental platforms, we were able to focus on the *Mtb* cell wall and were successful in reporting over 100 additional proteins that had not been identified previously. The elucidation of the cell wall proteome can facilitate subsequent studies of a more targeted nature where specific proteins can be monitored in response to various environmental, nutritional, or drug pressures.

One of the most interesting yet poorly understood protein families residing within the cell wall of *Mtb* are the triacylated lipoproteins. As stated above, Sutcliffe et al. predicted 2.5% of the genome to encode for lipoproteins. Defining these proteins beyond their antigenic and immune modulatory potential remains quite elusive with only a few studies predicting functional roles for *Mtb* lipoproteins.[13,54−56] In this study, a vast majority of proteins found were involved in small molecule and macromolecule metabolism. Specifically, many of the CWP were annotated as small molecule and nutrient binding. This builds upon the evidence that clearly defines two *Mtb* CW lipoproteins, LppX and LprG, as aiding and/or facilitating the transport of lipids[57] and small molecules,[58] respectively. The presence of these lipoproteins and perhaps other secreted, CW resident proteins provides a conduit between various nutrient transportation pathways that are required for cellular survival and growth while allowing the mycobacterial cell wall to maintain its rigid hydrophobic integrity. While bacterial cell walls have historically been thought of as structural scaffolding with little active interplay between the extracellular milieu and inner cytosolic environment, this perception of a static bacterial structure is fading and we are beginning to understand the true complexity and dynamic nature of the biological processes being facilitated within the cell wall. For instance, these proteins may contribute to bacterial cell wall remodeling events in response to environmental stressors such as nutrient depletion, antibiotic pressures, the host immune response, and tissue remodeling events throughout the course of infection. For *Mtb*, the highly complex framework of the mAGP−outer lipid architecture can now be definitively complemented with a highly diverse protein population. This new perspective will provide further insight into cell wall remodeling processes during mycobacterial infection as well as give a more comprehensive list of potential drug targets and diagnostic candidates.

Conclusion

The cell wall proteins of *Mtb* were extracted and resolved by 2DGE or 2DLC. Several classes of proteins were identified, including putative secreted proteins, lipoproteins, and known T and B cell antigens. Many proteins have unknown function; however, their presence within the cell wall may be of biological relevance.

Acknowledgment. This work was supported by the NIH, NIAID contract HHSN266200400091c (K.M.D., PI). The authors thank Drs. John Belisle and Darragh Heaslip for their consultation and advice throughout the construction of this work and Dr. Carolina Mehaffy for technical advice in the writing of this manuscript.

Supporting Information Available: Two tables, S1 and S2, are provided for validation of MS/MS data quality and complete listing of all proteins identified in this study, respectively. Figure S1 provides representative images of the 2DGE protein profiles for each CWP fraction. Figure S2 depicts Western blot validation of some proteins identified by MS/MS. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Flynn, J. L.; Chan, J. Immunology of tuberculosis. *Annu. Rev. Immunol.* **2001**, *19*, 93–129.
(2) Szczepan, J.; Andrzej, S.; Katarzyna, K. How *Mycobacterium tuberculosis* subverts host immune responses. *BioEssays* **2008**, *30* (10), 943–954.
(3) Brennan, P. J. Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis. *Tuberculosis* **2003**, *83* (1−3), 91–97.
(4) Briken, V.; Porcelli, S. A.; Besra, G. S.; Kremer, L. Mycobacterial lipoarabinomannan and related lipoglycans: from biogenesis to modulation of the immune response. *Mol. Microbiol.* **2004**, *53* (2), 391–403.
(5) Jackson, M.; Stadthagen, G.; Gicquel, B. Long-chain multiple methyl-branched fatty acid-containing lipids of Mycobacterium tuberculosis: biosynthesis, transport, regulation and biological activities. *Tuberculosis* **2007**, *87* (2), 78–86.
(6) Neyrolles, O.; Gould, K.; Gares, M. P.; Brett, S.; Janssen, R.; O'Gaora, P.; Herrmann, J. L.; Prevost, M. C.; Perret, E.; Thole, J. E.; Young, D. Lipoprotein access to MHC class I presentation during infection of murine macrophages with live mycobacteria. *J. Immunol.* **2001**, *166* (1), 447–457.
(7) Sutcliffe, I. C.; Harrington, D. J. Lipoproteins of Mycobacterium tuberculosis: an abundant and functionally diverse class of cell envelope components. *FEMS Microbiol. Rev.* **2004**, *28* (5), 645–659.
(8) Young, D. B.; Garbe, T. R. Lipoprotein antigens of Mycobacterium tuberculosis. *Res. Microbiol.* **1991**, *142* (1), 55–65.
(9) Brightbill, H. D.; Libraty, D. H.; Krutzik, S. R.; Yang, R. B.; Belisle, J. T.; Bleharski, J. R.; Maitland, M.; Norgard, M. V.; Plevy, S. E.; Smale, S. T.; Brennan, P. J.; Bloom, B. R.; Godowski, P. J.; Modlin, R. L. Host defense mechanisms triggered by microbial lipoproteins through toll-like receptors. *Science* **1999**, *285* (5428), 732–736.
(10) Gehring, A. J.; Dobos, K. M.; Belisle, J. T.; Harding, C. V.; Boom, W. H. Mycobacterium tuberculosis LprG (Rv1411c): a novel TLR-2 ligand that inhibits human macrophage class II MHC antigen processing. *J. Immunol.* **2004**, *173* (4), 2660–2668.
(11) Pecora, N. D.; Gehring, A. J.; Canaday, D. H.; Boom, W. H.; Harding, C. V. Mycobacterium tuberculosis LprA is a lipoprotein agonist of TLR2 that regulates innate immunity and APC function. *J. Immunol.* **2006**, *177* (1), 422–4229.
(12) Hertz, C. J.; Kiertscher, S. M.; Godowski, P. J.; Bouis, D. A.; Norgard, M. V.; Roth, M. D.; Modlin, R. L. Microbial lipopeptides stimulate dendritic cell maturation via Toll-like receptor 2. *J. Immunol.* **2001**, *166* (4), 2444–2450.
(13) Sanchez, A.; Espinosa, P.; Esparza, M. A.; Colon, M.; Bernal, G.; Mancilla, R. Mycobacterium tuberculosis 38-kDa lipoprotein is apoptogenic for human monocyte-derived macrophages. *Scand. J. Immunol.* **2009**, *69* (1), 20–28.
(14) Malen, H.; Softeland, T.; Wiker, H. G. Antigen analysis of Mycobacterium tuberculosis H37Rv culture filtrate proteins. *Scand. J. Immunol.* **2008**, *67* (3), 245–252.
(15) Malen, H.; Berven, F. S.; Softeland, T.; Arntzen, M. O.; D'Santos, C. S.; De Souza, G. A.; Wiker, H. G. Membrane and membrane-associated proteins in Triton X-114 extracts of Mycobacterium

bovis BCG identified using a combination of gel-based and gel-free fractionation strategies. *Proteomics* **2008**, *8* (9), 1859–1870.
(16) Malen, H.; Berven, F. S.; Fladmark, K. E.; Wiker, H. G. Comprehensive analysis of exported proteins from Mycobacterium tuberculosis H37Rv. *Proteomics* **2007**, *7* (10), 1702–1718.
(17) Urquhart, B. L.; Atsalos, T. E.; Roach, D.; Basseal, D. J.; Bjellqvist, B.; Britton, W. L.; Humphery-Smith, I. 'Proteomic contigs' of Mycobacterium tuberculosis and Mycobacterium bovis (BCG) using novel immobilised pH gradients. *Electrophoresis* **1997**, *18* (8), 1384–1392.
(18) Urquhart, B. L.; Cordwell, S. J.; Humphery-Smith, I. Comparison of predicted and observed properties of proteins encoded in the genome of Mycobacterium tuberculosis H37Rv. *Biochem. Biophys. Res. Commun.* **1998**, *253* (1), 70–79.
(19) Rosenkrands, I.; King, A.; Weldingh, K.; Moniatte, M.; Moertz, E.; Andersen, P. Towards the proteome of Mycobacterium tuberculosis. *Electrophoresis* **2000**, *21* (17), 3740–3756.
(20) Jungblut, P. R.; Muller, E. C.; Mattow, J.; Kaufmann, S. H. Proteomics reveals open reading frames in Mycobacterium tuberculosis H37Rv not predicted by genomics. *Infect. Immun.* **2001**, *69* (9), 5905–5907.
(21) Mattow, J.; Jungblut, P. R.; Schaible, U. E.; Mollenkopf, H. J.; Lamer, S.; Zimny-Arndt, U.; Hagens, K.; Muller, E. C.; Kaufmann, S. H. Identification of proteins from Mycobacterium tuberculosis missing in attenuated Mycobacterium bovis BCG strains. *Electrophoresis* **2001**, *22* (14), 2936–2946.
(22) Gu, S.; Chen, J.; Dobos, K. M.; Bradbury, E. M.; Belisle, J. T.; Chen, X. Comprehensive proteomic profiling of the membrane constituents of a Mycobacterium tuberculosis strain. *Mol. Cell. Proteomics* **2003**, *2* (12), 1284–1296.
(23) Schmidt, F.; Donahoe, S.; Hagens, K.; Mattow, J.; Schaible, U. E.; Kaufmann, S. H.; Aebersold, R.; Jungblut, P. R. Complementary analysis of the Mycobacterium tuberculosis proteome by two-dimensional electrophoresis and isotope-coded affinity tag technology. *Mol. Cell. Proteomics* **2004**, *3* (1), 24–42.
(24) Mawuenyega, K. G.; Forst, C. V.; Dobos, K. M.; Belisle, J. T.; Chen, J.; Bradbury, E. M.; Bradbury, A. R.; Chen, X. Mycobacterium tuberculosis functional network analysis by global subcellular protein profiling. *Mol. Biol. Cell* **2005**, *16* (1), 396–404.
(25) Xiong, Y.; Chalmers, M. J.; Gao, F. P.; Cross, T. A.; Marshall, A. G. Identification of Mycobacterium tuberculosis H37Rv integral membrane proteins by one-dimensional gel electrophoresis and liquid chromatography electrospray ionization tandem mass spectrometry. *J. Proteome Res.* **2005**, *4* (3), 855–861.
(26) Sinha, S.; Kosalai, K.; Arora, S.; Namane, A.; Sharma, P.; Gaikwad, A. N.; Brodin, P.; Cole, S. T. Immunogenic membrane-associated proteins of Mycobacterium tuberculosis revealed by proteomics. *Microbiology* **2005**, *151* (Pt 7), 2411–2419.
(27) Rosenkrands, I.; Weldingh, K.; Jacobsen, S.; Hansen, C. V.; Florio, W.; Gianetri, I.; Andersen, P. Mapping and identification of Mycobacterium tuberculosis proteins by two-dimensional gel electrophoresis, microsequencing and immunodetection. *Electrophoresis* **2000**, *21* (5), 935–948.
(28) Takayama, K.; Schnoes, H. K.; Armstrong, E. L.; Boyle, R. W. Site of inhibitory action of isoniazid in the synthesis of mycolic acids in Mycobacterium tuberculosis. *J. Lipid Res.* **1975**, *16* (4), 308–317.
(29) Hirschfield, G. R.; McNeil, M.; Brennan, P. J. Peptidoglycan-associated polypeptides of Mycobacterium tuberculosis. *J. Bacteriol.* **1990**, *172* (2), 1005–1013.
(30) Konigsberg, W. H.; Henderson, L. Removal of sodium dodecyl sulfate from proteins by ion-pair extraction. *Methods Enzymol.* **1983**, *91*, 254–259.
(31) Radolf, J. D.; Chamberlain, N. R.; Clausell, A.; Norgard, M. V. Identification and localization of integral membrane proteins of virulent Treponema pallidum subsp. pallidum by phase partitioning with the nonionic detergent triton X-114. *Infect. Immun.* **1988**, *56* (2), 490–498.
(32) Hellman, U.; Wernstedt, C.; Gonez, J.; Heldin, C. H. Improvement of an "in-gel" digestion procedure for the micropreparation of internal protein fragments for amino acid sequencing. *Anal. Biochem.* **1995**, *224* (1), 451–455.
(33) Rosenfeld, J.; Capdevielle, J.; Guillemot, J. C.; Ferrara, P. In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Anal. Biochem.* **1992**, *203* (1), 173–179.
(34) Turner, J.; Dobos, K. M.; Keen, M. A.; Frank, A. A.; Ehlers, S.; Orme, I. M.; Belisle, J. T.; Cooper, A. M. A limited antigen-specific cellular response is sufficient for the early control of Mycobacterium tuberculosis in the lung but is insufficient for long-term survival. *Infect. Immun.* **2004**, *72* (7), 3759–3768.
(35) Keller, A.; Nesvizhskii, A. I.; Kolker, E.; Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **2002**, *74* (20), 5383–5392.
(36) Searle, B. C. Scaffold: a bioinformatic tool for validating MS/MS-based proteomic studies. *Proteomics* , *10* (6), 1265–1269.
(37) Bendtsen, J. D.; Nielsen, H.; von Heijne, G.; Brunak, S. Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **2004**, *340* (4), 783–795.
(38) Nielsen, H.; Engelbrecht, J.; Brunak, S.; von Heijne, G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **1997**, *10* (1), 1–6.
(39) Nielsen, H.; Krogh, A. Prediction of signal peptides and signal anchors by a hidden Markov model. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **1998**, *6*, 122–130.
(40) Mawuenyega, K. G.; Kaji, H.; Yamuchi, Y.; Shinkawa, T.; Saito, H.; Taoka, M.; Takahashi, N.; Isobe, T. Large-scale identification of *Caenorhabditis elegans* proteins by multidimensional liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* **2003**, *2* (1), 23–35.
(41) Isobe, T.; Uchida, K.; Taoka, M.; Shinkai, F.; Manabe, T.; Okuyama, T. Automated two-dimensional liquid chromatographic system for mapping proteins in highly complex mixtures. *J. Chromatogr.* **1991**, *588* (1−2), 115–123.
(42) Mollenkopf, H. J.; Jungblut, P. R.; Raupach, B.; Mattow, J.; Lamer, S.; Zimny-Arndt, U.; Schaible, U. E.; Kaufmann, S. H. A dynamic two-dimensional polyacrylamide gel electrophoresis database: the mycobacterial proteome via Internet. *Electrophoresis* **1999**, *20* (11), 2172–2180.
(43) Juncker, A. S.; Willenbrock, H.; Von Heijne, G.; Brunak, S.; Nielsen, H.; Krogh, A. Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci.* **2003**, *12* (8), 1652–1662.
(44) Sartain, M. J.; Slayden, R. A.; Singh, K. K.; Laal, S.; Belisle, J. T. Disease state differentiation and identification of tuberculosis biomarkers via native antigen array profiling. *Mol. Cell. Proteomics* **2006**, *5* (11), 2102–2113.
(45) Sartain, M. J.; Belisle, J. T. N-Terminal clustering of the O-glycosylation sites in the Mycobacterium tuberculosis lipoprotein SodC. *Glycobiology* **2009**, *19* (1), 38–51.
(46) Gonzalez-Zamorano, M.; Mendoza-Hernandez, G.; Xolalpa, W.; Parada, C.; Vallecillo, A. J.; Bigi, F.; Espitia, C. Mycobacterium tuberculosis glycoproteomics based on ConA-lectin affinity capture of mannosylated proteins. *J. Proteome Res.* **2009**, *8* (2), 721–733.
(47) Driessen, A. J.; Manting, E. H.; van der Does, C. The structural basis of protein targeting and translocation in bacteria. *Nat. Struct. Biol.* **2001**, *8* (6), 492–498.
(48) Mattow, J.; Schaible, U. E.; Schmidt, F.; Hagens, K.; Siejak, F.; Brestrich, G.; Haeselbarth, G.; Muller, E. C.; Jungblut, P. R.; Kaufmann, S. H. Comparative proteome analysis of culture supernatant proteins from virulent Mycobacterium tuberculosis H37Rv and attenuated M. bovis BCG Copenhagen. *Electrophoresis* **2003**, *24* (19−20), 3405–3420.
(49) Dobos, K. M.; Spencer, J. S.; Orme, I. M.; Belisle, J. T. Proteomic approaches to antigen discovery. *Methods Mol. Med.* **2004**, *94*, 3–17.
(50) Mehaffy, C.; Hess, A.; Prenni, J. E.; Mathema, B.; Kreiswirth, B.; Dobos, K. M. Descriptive proteomic analysis shows protein variability between closely related clinical isolates of Mycobacterium tuberculosis. *Proteomics* **2010**, *10* (10), 1966–1984.
(51) Young, D.; Garbe, T.; Lathigra, R.; Abou-Zeid, C.; Zhang, Y. Characterization of prominent protein antigens from mycobacteria. *Bull. Int. Union Tuberc. Lung Dis.* **1991**, *66* (1), 47–51.
(52) Sonnenberg, M. G.; Belisle, J. T. Definition of Mycobacterium tuberculosis culture filtrate proteins by two-dimensional polyacrylamide gel electrophoresis, N-terminal amino acid sequencing, and electrospray mass spectrometry. *Infect. Immun.* **1997**, *65* (11), 4515–24.
(53) Mattow, J.; Jungblut, P. R.; Muller, E. C.; Kaufmann, S. H. Identification of acidic, low molecular mass proteins of Mycobacterium tuberculosis strain H37Rv by matrix-assisted laser desorption/ionization and electrospray ionization mass spectrometry. *Proteomics* **2001**, *1* (4), 494–507.
(54) Ciaramella, A.; Cavone, A.; Santucci, M. B.; Garg, S. K.; Sanarico, N.; Bocchino, M.; Galati, D.; Martino, A.; Auricchio, G.; D'Orazio, M.; Stewart, G. R.; Neyrolles, O.; Young, D. B.; Colizzi, V.; Fraziano, M. Induction of apoptosis and release of interleukin-1 beta by cell wall-associated 19-kDa lipoprotein during the course of mycobacterial infection. *J. Infect. Dis.* **2004**, *190* (6), 1167–1176.
(55) Farrow, M. F.; Rubin, E. J. Function of a mycobacterial major facilitator superfamily pump requires a membrane-associated lipoprotein. *J. Bacteriol.* **2008**, *190* (5), 1783–1791.

(56) Drage, M. G.; Pecora, N. D.; Hise, A. G.; Fabbraio, M.; Silverstein, R. L.; Golenbock, D. T.; Boom, W. H.; Harding, C. V. TLR2 and its co-receptors determine responses of macrophages and dendritic cells to lipoproteins of *Mycobacterium tuberculosis*. *Cellular Immunology* **2009**, *258*, 29–37.
(57) Sulzenbacher, G.; Canaan, S.; Bordat, Y.; Neyrolles, O.; Stadthagen, G.; Roig-Zamboni, V.; Rauzier, J.; Maurin, D.; Laval, F.; Daffe, M.; Cambillau, C.; Gicquel, B.; Bourne, Y.; Jackson, M. LppX is a lipoprotein required for the translocation of phthiocerol dimycocerosates to the surface of Mycobacterium tuberculosis. *EMBO J.* **2006**, *25* (7), 1436–1444.
(58) Farrow, M. F.; Rubin, E. J. Function of a mycobacterial major facilitator superfamily pump requires a membrane-associated lipoprotein. *J. Bacteriol.* **2008**, *190* (5), 1783–1791.

PR1005873